A Solution Manual for: A First Course In Probability by Sheldon M. Ross.

John L. Weatherwax*

September 24, 2012

Introduction

Here you'll find some notes that I wrote up as I worked through this excellent book. I've worked hard to make these notes as good as I can, but I have no illusions that they are perfect. If you feel that that there is a better way to accomplish or explain an exercise or derivation presented in these notes; or that one or more of the explanations is unclear, incomplete, or misleading, please tell me. If you find an error of any kind – technical, grammatical, typographical, whatever – please tell me that, too. I'll gladly add to the acknowledgments in later printings the name of the first person to bring each problem to my attention.

Acknowledgments

Special thanks to (most recent comments are listed first): Luis Mendoza-Nativida, Francisco Perez, Jason Franklin, Chuyue Wei, CJ Turtoro, Ho Wei Hong, John Williams (several contributions to chapter 4), Ralph Lozano, Vincent Leung, Mark Chamness, Dale Peterson, Doug Edmunds, Marlene Miller, Timothy Alsobrooks, Konstantinos Stouras, William Howell, Robert Futyma, Waldo Arriagada, Atul Narang, Andrew Jones, Vincent Frost, and Gerardo Robert for helping improve these notes and solutions. It should be noted that Marlene Miller made several helpful suggestions on most of the material in Chapter 3. Her algebraic use of event "set" notation to solve probability problems has opened my eyes to this powerful technique. It is a tool that I wish to become more proficient with.

All comments (no matter how small) are much appreciated. In fact, if you find these notes useful I would appreciate a contribution in the form of a solution to a problem that is not yet

*wax@alum.mit.edu

worked in these notes. Sort of a "take a penny, leave a penny" type of approach. Remember: pay it forward.

Miscellaneous Problems

The Crazy Passenger Problem

The following is known as the "crazy passenger problem" and is stated as follows. A line of 100 airline passengers is waiting to board the plane. They each hold a ticket to one of the 100 seats on that flight. (For convenience, let's say that the k-th passenger in line has a ticket for the seat number k.) Unfortunately, the first person in line is *crazy*, and will ignore the seat number on their ticket, picking a random seat to occupy. All the other passengers are quite normal, and will go to their proper seat unless it is already occupied. If it is occupied, they will then find a free seat to sit in, at random. What is the probability that the last (100th) person to board the plane will sit in their proper seat (#100)?

If one tries to solve this problem with conditional probability it becomes very difficult. We begin by considering the following cases if the first passenger sits in seat number 1, then all the remaining passengers will be in their correct seats and certainly the #100'th will also. If he sits in the last seat #100, then certainly the last passenger cannot sit there (in fact he will end up in seat #1). If he sits in any of the 98 seats *between* seats #1 and #100, say seat k, then all the passengers with seat numbers $2, 3, \ldots, k-1$ will have empty seats and be able to sit in their respective seats. When the passenger with seat number k enters he will have as possible seating choices seat #1, one of the seats $k+1, k+2, \ldots, 99$, or seat #100. Thus the options available to this passenger are the *same* options available to the first passenger. That is if he sits in seat #1 the remaining passengers with seat labels $k+1, k+2, \ldots, 100$ can sit in their assigned seats and passenger #100 can sit in his seat, or he can sit in seat #100 in which case the passenger #100 is blocked, or finally he can sit in one of the seats between seat k and seat #99. The only difference is that this k-th passenger has fewer choices for the "middle" seats. This k passenger effectively becomes a new "crazy" passenger.

From this argument we begin to see a recursive structure. To fully specify this recursive structure lets generalize this problem a bit an assume that there are N total seats (rather than just 100). Thus at each stage of placing a k-th crazy passenger we can choose from

- seat #1 and the last or N-th passenger will then be able to sit in their assigned seat, since all intermediate passenger's seats are unoccupied.
- seat # N and the last or N-th passenger will be unable to sit in their assigned seat.
- any seat before the N-th and after the k-th. Where the k-th passenger's seat is taken by a crazy passenger from the previous step. In this case there are $N-1-(k+1)+1 = N-k-1$ "middle" seat choices.

If we let $p(n, 1)$ be the probability that given one crazy passenger and n total seats to select from that the last passenger sits in his seat. From the argument above we have a recursive structure give by

$$\begin{aligned} p(N,1) &= \frac{1}{N}(1) + \frac{1}{N}(0) + \frac{1}{N}\sum_{k=2}^{N-1} p(N-k,1) \\ &= \frac{1}{N} + \frac{1}{N}\sum_{k=2}^{N-1} p(N-k,1)\,. \end{aligned}$$

where the first term is where the first passenger picks the first seat (where the N will sit correctly with probability one), the second term is when the first passenger sits in the N-th seat (where the N will sit correctly with probability zero), and the remaining terms represent the first passenger sitting at position k, which will then require repeating this problem with the k-th passenger choosing among $N - k + 1$ seats.

To solve this recursion relation we consider some special cases and then apply the principle of mathematical induction to prove it. Lets take $N = 2$. Then there are only two possible arrangements of passengers $(1, 2)$ and $(2, 1)$ of which one (the first) corresponds to the second passenger sitting in his assigned seat. This gives

$$p(2,1) = \frac{1}{2}\,.$$

If $N = 3$, then from the $3! = 6$ possible choices for seating arrangements

$$(1,2,3)\,(1,3,2)\,(2,3,1)\,(2,1,3)\,(3,1,2)\,(3,2,1)$$

Only

$$(1,2,3)\,(2,1,3)\,(3,2,1)$$

correspond to admissible seating arrangements for this problem so we see that

$$p(3,1) = \frac{3}{6} = \frac{1}{2}\,.$$

If we hypothesis that $p(N, 1) = \frac{1}{2}$ for all N, placing this assumption into the recursive formulation above gives

$$p(N,1) = \frac{1}{N} + \frac{1}{N}\sum_{k=2}^{N-1}\frac{1}{2} = \frac{1}{2}\,.$$

Verifying that indeed this constant value satisfies our recursion relationship.

Chapter 1 (Combinatorial Analysis)

Chapter 1: Problems

Problem 1 (counting license plates)

Part (a): In each of the first two places we can put any of the 26 letters giving 26^2 possible letter combinations for the first two characters. Since the five other characters in the license plate must be numbers, we have 10^5 possible five digit letters their specification giving a total of

$$26^2 \cdot 10^5 = 67600000\,,$$

total license plates.

Part (b): If we can't repeat a letter or a number in the specification of a license plate then the number of license plates becomes

$$26 \cdot 25 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 = 19656000\,,$$

total license plates.

Problem 2 (counting die rolls)

We have six possible outcomes for each of the die rolls giving $6^4 = 1296$ possible total outcomes for all four rolls.

Problem 3 (assigning workers to jobs)

Since each job is different and each worker is unique we have 20! different pairings.

Problem 4 (creating a band)

If each boy can play each instrument we can have $4! = 24$ ordering. If Jay and Jack can play only two instruments then we will assign the instruments they play first with 2! possible orderings. The other two boys can be assigned the remaining instruments in 2! ways and thus we have

$$2! \cdot 2! = 4\,,$$

possible unique band assignments.

Problem 5 (counting telephone area codes)

In the first specification of this problem we can have $9 - 2 + 1 = 8$ possible choices for the first digit in an area code. For the second digit there are two possible choices. For the third digit there are 9 possible choices. So in total we have

$$8 \cdot 2 \cdot 9 = 144 \, ,$$

possible area codes. In the second specification of this problem, if we must start our area codes with the digit "four" we will only have $2 \cdot 9 = 18$ area codes.

Problem 6 (counting kittens)

The traveler would meet $7^4 = 2401$ kittens.

Problem 7 (arranging boys and girls)

Part (a): Since we assume that each person is unique, the total number of ordering is given by $6! = 720$.

Part (b): We have 3! orderings of each group of the three boys and girls. Since we can put these groups of boys and girls in 2! different ways (either the boys first or the girls first) we have

$$(2!) \cdot (3!) \cdot (3!) = 2 \cdot 6 \cdot 6 = 72 \, ,$$

possible orderings.

Part (c): If the boys must sit together we have $3! = 6$ ways to arrange the block of boys. This block of boys can be placed either at the ends or in between any of the individual 3! orderings of the girls. This gives four locations where our block of boys can be placed we have

$$4 \cdot (3!) \cdot (3!) = 144 \, ,$$

possible orderings.

Part (d): The only way that no two people of the same sex can sit together is to have the two groups interleaved. Now there are 3! ways to arrange each group of girls and boys, and to interleave we have two different choices for interleaving. For example with three boys and girls we could have

$$g_1 b_1 g_2 b_2 g_3 b_3 \quad \text{vs.} \quad b_1 g_1 b_2 g_2 b_3 g_3 \, ,$$

thus we have

$$2 \cdot 3! \cdot 3! = 2 \cdot 6^2 = 72 \, ,$$

possible arrangements.

Problem 8 (counting arrangements of letters)

Part (a): Since "Fluke" has five unique letters we have $5! = 120$ possible arrangements.

Part (b): Since "Propose" has seven letters of which four (the "o"'s and the "p"'s) repeat we have

$$\frac{7!}{2! \cdot 2!} = 1260\,,$$

arrangements.

Part (c): Now "Mississippi" has eleven characters with the "i" repeated four times, the "s" repeated four times and the "p" repeated two times, so we have

$$\frac{11!}{4! \cdot 4! \cdot 2!} = 34650\,,$$

possible rearranges.

Part (d): "Arrange" has seven characters with a double "a" and a double "r" so it has

$$\frac{7!}{2! \cdot 2!} = 1260\,,$$

different arrangements.

Problem 9 (counting colored blocks)

Assuming each block is unique we have 12! arrangements, but since the six black and the four red blocks are not distinguishable we have

$$\frac{12!}{6! \cdot 4!} = 27720\,,$$

possible arrangements.

Problem 10 (seating people in a row)

Part (a): We have $8! = 40320$ possible seating arrangements.

Part (b): We have 6! ways to place the people (not including A and B). We have 2! ways to order A and B. Once the pair of A and B is determined, they can be placed in between any ordering of the other six. For example, any of the "x"'s in the expression below could be replaced with the A B pair

$$x\,P_1\,x\,P_2\,x\,P_3\,x\,P_4\,x\,P_5 x\,P_6\,x\,.$$

Giving seven possible locations for the A,B pair. Thus the total number of orderings is given by

$$2! \cdot 6! \cdot 7 = 10080 .$$

Part (c): To place the men and women according to the given rules, the men and women must be interleaved. We have 4! ways to arrange the men and 4! ways to arrange the women. We can start our sequence of eight people with a woman or a man (giving two possible choices). We thus have

$$2 \cdot 4! \cdot 4! = 1152 ,$$

possible arrangements.

Part (d): Since the five men must sit next to each other their ordering can be specified in $5! = 120$ ways. This block of men can be placed in between any of the three women, or at the end of the block of women, who can be ordered in 3! ways. Since there are four positions we can place the block of men we have

$$5! \cdot 4 \cdot 3! = 2880 ,$$

possible arrangements.

Part (e): The four couple have 2! orderings within each pair, and then 4! orderings of the pairs giving a total of

$$(2!)^4 \cdot 4! = 384 ,$$

total orderings.

Problem 11 (counting arrangements of books)

Part (a): We have $(3 + 2 + 1)! = 6! = 720$ arrangements.

Part (b): The mathematics books can be arranged in 2! ways and the novels in 3! ways. Then the block ordering of mathematics, novels, and chemistry books can be arranged in 3! ways resulting in

$$(3!) \cdot (2!) \cdot (3!) = 72 ,$$

possible arrangements.

Part (c): The number of ways to arrange the novels is given by $3! = 6$ and the other three books can be arranged in 3! ways with the blocks of novels in any of the four positions in between giving

$$4 \cdot (3!) \cdot (3!) = 144 ,$$

possible arrangements.

Problem 12 (counting awards)

Part (a): We have 30 students to choose from for the first award, and 30 students to choose from for the second award, etc. So the total number of different outcomes is given by

$$30^5 = 24300000$$

Part (b): We have 30 students to choose from for the first award, 29 students to choose from for the second award, etc. So the total number of different outcomes is given by

$$30 \cdot 29 \cdot 28 \cdot 27 \cdot 26 = 17100720$$

Problem 13 (counting handshakes)

With 20 people the number of pairs is given by

$$\left(\begin{array}{c} 20 \\ 2 \end{array} \right) = 190\,.$$

Problem 14 (counting poker hands)

A deck of cards has four suits with thirteen cards each giving in total 52 cards. From these 52 cards we need to select five to form a poker hand thus we have

$$\left(\begin{array}{c} 52 \\ 5 \end{array} \right) = 2598960\,,$$

unique poker hands.

Problem 15 (pairings in dancing)

We must first choose five women from ten in $\left(\begin{array}{c} 10 \\ 5 \end{array} \right)$ possible ways, and five men from 12 in $\left(\begin{array}{c} 12 \\ 5 \end{array} \right)$ ways. Once these groups are chosen then we have 5! pairings of the men and women. Thus in total we will have

$$\left(\begin{array}{c} 10 \\ 5 \end{array} \right) \left(\begin{array}{c} 12 \\ 5 \end{array} \right) 5! = 252 \cdot 792 \cdot 120 = 23950080\,,$$

possible pairings.

Problem 16 (forced selling of books)

Part (a): We have to select a subject from three choices. If we choose math we have $\left(\begin{array}{c}6\\2\end{array}\right) = 15$ choices of books to sell. If we choose science we have $\left(\begin{array}{c}7\\2\end{array}\right) = 21$ choices of books to sell. If we choose economics we have $\left(\begin{array}{c}4\\2\end{array}\right) = 6$ choices of books to sell. Since each choice is mutually exclusive in total we have $15 + 21 + 6 = 42$, possible choices.

Part (b): We must pick two subjects from $\left(\begin{array}{c}3\\2\end{array}\right) = 3$ choices. If we denote the letter "M" for the choice math the letter "S" for the choice science, and the letter "E" for the choice economics then the three choices are

$$(M, S) \quad (M, E) \quad (S, E) \,.$$

For each of the choices above we have $6 \cdot 7 + 6 \cdot 4 + 7 \cdot 4 = 94$ total choices.

Problem 17 (distributing gifts)

We can choose seven children to give gifts to in $\left(\begin{array}{c}10\\7\end{array}\right)$ ways. Once we have chosen the seven children, the gifts can be distributed in 7! ways. This gives a total of

$$\left(\begin{array}{c}10\\7\end{array}\right) \cdot 7! = 604800 \,,$$

possible gift distributions.

Problem 18 (selecting political parties)

We can choose two Republicans from the five total in $\left(\begin{array}{c}5\\2\end{array}\right)$ ways, we can choose two Democrats from the six in $\left(\begin{array}{c}6\\2\end{array}\right)$ ways, and finally we can choose three Independents from the four in $\left(\begin{array}{c}4\\3\end{array}\right)$ ways. In total, we will have

$$\left(\begin{array}{c}5\\2\end{array}\right) \cdot \left(\begin{array}{c}6\\2\end{array}\right) \cdot \left(\begin{array}{c}4\\3\end{array}\right) = 600 \,,$$

different committees.

Problem 19 (counting committees with constraints)

Part (a): We select three men from six in $\binom{6}{3}$, but since two men won't serve together we need to compute the number of these pairings of three men that have the two that won't serve together. The number of committees we can form (with these two together) is given by

$$\binom{2}{2} \cdot \binom{4}{1} = 4 \, .$$

So we have

$$\binom{6}{3} - 4 = 16 \, ,$$

possible groups of three men. Since we can choose $\binom{8}{3} = 56$ different groups of women, we have in total $16 \cdot 56 = 896$ possible committees.

Part (b): If two women refuse to serve together, then we will have $\binom{2}{2} \cdot \binom{6}{1}$ groups with these two women in them from the $\binom{8}{3}$ ways to draw three women from eight. Thus we have

$$\binom{8}{3} - \binom{2}{2} \cdot \binom{6}{1} = 56 - 6 = 50 \, ,$$

possible groupings of woman. We can select three men from six in $\binom{6}{3} = 20$ ways. In total then we have $50 \cdot 20 = 1000$ committees.

Part (c): We have $\binom{8}{3} \cdot \binom{6}{3}$ total committees, and

$$\binom{1}{1} \cdot \binom{7}{2} \cdot \binom{1}{1} \cdot \binom{5}{2} = 210 \, ,$$

committees containing the man and women who refuse to serve together. So we have

$$\binom{8}{3} \cdot \binom{6}{3} - \binom{1}{1} \cdot \binom{7}{2} \cdot \binom{1}{1} \cdot \binom{5}{2} = 1120 - 210 = 910 \, ,$$

total committees.

Problem 20 (counting the number of possible parties)

Part (a): There are a total of $\left(\begin{array}{c}8\\5\end{array}\right)$ possible groups of friends that could attend (assuming no feuds). We have $\left(\begin{array}{c}2\\2\end{array}\right)\cdot\left(\begin{array}{c}6\\3\end{array}\right)$ sets with our two feuding friends in them, giving

$$\left(\begin{array}{c}8\\5\end{array}\right)-\left(\begin{array}{c}2\\2\end{array}\right)\cdot\left(\begin{array}{c}6\\3\end{array}\right)=36$$

possible groups of friends

Part (b): If two fiends must attend together we have that $\left(\begin{array}{c}2\\2\end{array}\right)\left(\begin{array}{c}6\\3\end{array}\right)$ if the *do* attend the party together and $\left(\begin{array}{c}6\\5\end{array}\right)$ if they *don't* attend at all, giving a total of

$$\left(\begin{array}{c}2\\2\end{array}\right)\left(\begin{array}{c}6\\3\end{array}\right)+\left(\begin{array}{c}6\\5\end{array}\right)=26\,.$$

Problem 21 (number of paths on a grid)

From the hint given that we must take four steps to the right and three steps up, we can think of any possible path as an arraignment of the letters "U" for up and "R" for right. For example the string

$$UUURRRR\,,$$

would first step up three times and then right four times. Thus our problem becomes one of counting the number of unique arrangements of three "U"'s and four "R"'s, which is given by

$$\frac{7!}{4!\cdot 3!}=35\,.$$

Problem 22 (paths on a grid through a specific point)

One can think of the problem of going through a specific point (say P) as counting the number of paths from the start A to P and then counting the number of paths from P to the end B. To go from A to P (where P occupies the $(2,2)$ position in our grid) we are looking for the number of possible unique arrangements of two "U"'s and two "R"'s, which is given by

$$\frac{4!}{2!\cdot 2!}=6\,,$$

possible paths. The number of paths from the point P to the point B is equivalent to the number of different arrangements of two "R"'s and one "U" which is given by

$$\frac{3!}{2!\cdot 1!}=3\,.$$

From the basic principle of counting then we have $6 \cdot 3 = 18$ total paths.

Problem 23 (assignments to beds)

Assuming that twins sleeping in different bed in the same room counts as a different arraignment, we have $(2!) \cdot (2!) \cdot (2!) = 8$ possible assignments of each set of twins to a room. Since there are 3! ways to assign the pair of twins to individual rooms we have $6 \cdot 8 = 48$ possible assignments.

Problem 24 (practice with the binomial expansion)

This is given by

$$(3x^2 + y)^5 = \sum_{k=0}^{5} \left(\begin{array}{c} 5 \\ k \end{array} \right) (3x^2)^k y^{5-k} \,.$$

Problem 25 (bridge hands)

We have 52! unique permutations, but since the different arrangements of cards within a given hand do not matter we have

$$\frac{52!}{(13!)^4} \,,$$

possible bridge hands.

Problem 26 (practice with the multinomial expansion)

This is given by the multinomial expansion

$$(x_1 + 2x_2 + 3x_3)^4 = \sum_{n_1+n_2+n_3=4} \left(\begin{array}{c} 4 \\ n_1 \,, n_2 \,, n_3 \end{array} \right) x_1^{n_1} (2x_2)^{n_2} (3x_3)^{n_3}$$

The number of terms in the above summation is given by

$$\left(\begin{array}{c} 4+3-1 \\ 3-1 \end{array} \right) = \left(\begin{array}{c} 6 \\ 2 \end{array} \right) = \frac{6 \cdot 5}{2} = 15 \,.$$

Problem 27 (counting committees)

This is given by the multinomial coefficient

$$\binom{12}{3\,,4\,,5} = 27720$$

Problem 28 (divisions of teachers)

If we decide to send n_1 teachers to school one and n_2 teachers to school two, etc. then the total number of unique assignments of (n_1, n_2, n_3, n_4) number of teachers to the four schools is given by

$$\binom{8}{n_1\,,n_2\,,n_3\,,n_4}.$$

Since we want the total number of divisions, we must sum this result for all possible combinations of n_i, or

$$\sum_{n_1+n_2+n_3+n_4=8} \binom{8}{n_1\,,n_2\,,n_3\,,n_4} = (1+1+1+1)^8 = 65536\,,$$

possible divisions.

If each school must receive two in each school, then we are looking for

$$\binom{8}{2\,,2\,,2\,,2} = \frac{8!}{(2!)^4} = 2520\,,$$

orderings.

Problem 29 (dividing weight lifters)

We have 10! possible permutations of all weight lifters but the permutations of individual countries (contained within this number) are irrelevant. Thus we can have

$$\frac{10!}{3!\cdot 4!\cdot 2!\cdot 1!} = \binom{10}{3\,,4\,,2\,,1} = 12600\,,$$

possible divisions. If the united states has one competitor in the top three and two in the bottom three. We have $\binom{3}{1}$ possible positions for the US member in the first three positions and $\binom{3}{2}$ possible positions for the two US members in the bottom three positions, giving a total of

$$\binom{3}{1}\binom{3}{2} = 3\cdot 3 = 9\,,$$

combinations of US members in the positions specified. We also have to place the other countries participants in the remaining $10-3=7$ positions. This can be done in $\binom{7}{4,2,1} = \frac{7!}{4!\cdot 2!\cdot 1!} = 105$ ways. So in total then we have $9\cdot 105 = 945$ ways to position the participants.

Problem 30 (seating delegates in a row)

If the French and English delegates are to be seated next to each other, they can be can be placed in 2! ways. Then this pair constitutes a new "object" which we can place anywhere among the remaining eight people, i.e. there are 9! arrangements of the eight remaining people and the French and English pair. Thus we have $2\cdot 9! = 725760$ possible combinations. Since in some of these the Russian and US delegates are next to each other, this number over counts the true number we are looking for by $2\cdot 2\cdot 8! = 161280$ (the first two is for the number of arrangements of the French and English pair). Combining these two criterion we have

$$2\cdot(9!) - 4\cdot(8!) = 564480\,.$$

Problem 31 (distributing blackboards)

Let x_i be the number of black boards given to school i, where $i=1,2,3,4$. Then we must have $\sum_i x_i = 8$, with $x_i \geq 0$. The number of solutions to an equation like this is given by

$$\binom{8+4-1}{4-1} = \binom{11}{3} = 165\,.$$

If each school must have at least one blackboard then the constraints change to $x_i \geq 1$ and the number of such equations is give by

$$\binom{8-1}{4-1} = \binom{7}{3} = 35\,.$$

Problem 32 (distributing people)

Assuming that the elevator operator can only tell the number of people getting off at each floor, we let x_i equal the number of people getting off at floor i, where $i=1,2,3,4,5,6$. Then the constraint that all people are off at the sixth floor means that $\sum_i x_i = 8$, with $x_i \geq 0$. This has

$$\binom{n+r-1}{r-1} = \binom{8+6-1}{6-1} = \binom{13}{5} = 1287\,,$$

possible distribution people. If we have five men and three women, let m_i and w_i be the number of men and women that get off at floor i. We can solve this problem as the combination of two problems. That of tracking the men that get off on floor i and that of tracking the women that get off on floor i. Thus we must have

$$\begin{aligned}\sum_{i=1}^{6} m_i &= 5 \quad m_i \geq 0 \\ \sum_{i=1}^{6} w_i &= 3 \quad w_i \geq 0 \,.\end{aligned}$$

The number of solutions to the first equation is given by

$$\left(\begin{array}{c} 5+6-1 \\ 6-1 \end{array} \right) = \left(\begin{array}{c} 10 \\ 5 \end{array} \right) = 252 \,,$$

while the number of solutions to the second equation is given by

$$\left(\begin{array}{c} 3+6-1 \\ 6-1 \end{array} \right) = \left(\begin{array}{c} 8 \\ 5 \end{array} \right) = 56 \,.$$

So in total then (since each number is exclusive) we have $252 \cdot 56 = 14114$ possible elevator situations.

Problem 33 (possible investment strategies)

Let x_i be the number of investments made in opportunity i. Then we must have

$$\sum_{i=1}^{4} x_i = 20$$

with constraints that $x_1 \geq 2$, $x_2 \geq 2$, $x_3 \geq 3$, $x_4 \geq 4$. Writing this equation as

$$x_1 + x_2 + x_3 + x_4 = 20$$

we can subtract the lower bound of each variable to get

$$(x_1 - 2) + (x_2 - 2) + (x_3 - 3) + (x_4 - 4) = 20 - 2 - 2 - 3 - 4 = 9 \,.$$

Then defining $v_1 = x_1 - 2$, $v_2 = x_2 - 2$, $v_3 = x_3 - 3$, and $v_4 = x_4 - 4$, then our equation becomes $v_1 + v_2 + v_3 + v_4 = 9$, with the constraint that $v_i \geq 0$. The number of solutions to equations such as these is given by

$$\left(\begin{array}{c} 9+4-1 \\ 4-1 \end{array} \right) = \left(\begin{array}{c} 12 \\ 3 \end{array} \right) = 220 \,.$$

Part (b): First we pick the three investments from the four possible in $\left(\begin{array}{c} 4 \\ 3 \end{array} \right) = 4$ possible ways. The four choices are denoted in table 1, where a one denotes that we invest in that option. Then investment choice number one requires the equation $v_2+v_3+v_4 = 20-2-3-4 =$

choice	$v_1 = x_1 - 2 \geq 0$	$v_2 = x_2 - 2 \geq 0$	$v_3 = x_3 - 3 \geq 0$	$v_4 = x_4 - 4 \geq 0$
1	0	1	1	1
2	1	0	1	1
3	1	1	0	1
4	1	1	1	0

Table 1: All possible choices of three investments.

11, and has $\binom{11+3-1}{3-1} = \binom{13}{2} = 78$ possible solutions. Investment choice number two requires the equation $v_1 + v_3 + v_4 = 20 - 2 - 3 - 4 = 11$, and again has $\binom{11+3-1}{3-1} = \binom{13}{2} = 78$ possible solutions. Investment choice number three requires the equation $v_1 + v_2 + v_4 = 20 - 2 - 2 - 4 = 12$, and has $\binom{12+3-1}{3-1} = \binom{14}{2} = 91$ possible solutions. Finally, investment choice number four requires the equation $v_1 + v_2 + v_3 = 20 - 2 - 2 - 3 = 13$, and has $\binom{13+3-1}{3-1} = \binom{15}{2} = 105$ possible solutions. Of course we could also invest in all four opportunities which has the same number of possibilities as in part (a) or 220. Then in total since we can do any of these choices we have $220 + 105 + 91 + 78 + 78 = 572$ choices.

Chapter 1: Theoretical Exercises

Problem 1 (the generalized counting principle)

This can be proved by recursively applying the basic principle of counting.

Problem 2 (counting dependent experimental outcomes)

We have m choices for the outcome of the first experiment. If the first experiment returns i as an outcome, then there are n_i possible outcomes for the second experiment. Thus if the experiment returns "one" we have n_1 possible outcomes, if it returns "two" we have n_2 possible outcomes, etc. To count the number of possible experimental outcomes we can envision a tree like structure representing the totality of possible outcomes, where we have m branches leaving the root node indicating the m possible outcomes from the first experiment. From the first of these branches we have n_1 additional branches representing the outcome of the second experiment when the first experimental outcome was a one. From the second branch we have n_2 additional branches representing the outcome of the second experiment when the first experimental outcome was a two. We can continue this process, with the m-th branch from the root node having n_m leaves representing the outcome of the second experiment when the first experimental outcome was a m. Counting all of these outcomes we have

$$n_1 + n_2 + n_3 + \cdots + n_m \,,$$

total experimental outcomes.

Problem 3 (selecting r objects from n)

To select r objects from n, we will have n choices for the first object, $n-1$ choices for the second object, $n-2$ choices for the third object, etc. Continuing we will have $n-r+1$ choices for the selection of the r-th object. Giving a total of $n(n-1)(n-2)\cdots(n-r+1)$ total choices if the order of selection matters. If it does not then we must divide by the number of ways to rearrange the r selected objects i.e. $r!$ giving

$$\frac{n(n-1)(n-2)\cdots(n-r+1)}{r!} \,,$$

possible ways to select r objects from n when the order of selection of the r object does not matter.

Problem 4 (combinatorial explanation of $\left(\begin{array}{c} n \\ k \end{array}\right)$)

If all balls are distinguishable then there are $n!$ ways to arrange all the balls. With in this arrangement there are $r!$ ways to uniquely arrange the black balls and $(n-r)!$ ways to uniquely arranging the white balls. These arrangements don't represent new patterns since the balls with the same color are in fact indistinguishable. Dividing by these repeated patterns gives

$$\frac{n!}{r!(n-r)!} \,,$$

gives the unique number of permutations.

Problem 5 (the number of binary vectors who's sum is greater than k)

To have the sum evaluate to exactly k, we must select at k components from the vector x to have the value one. Since there are n components in the vector x, this can be done in $\left(\begin{array}{c} n \\ k \end{array}\right)$ ways. To have the sum exactly equal $k+1$ we must select $k+1$ components from x to have a value one. This can be done in $\left(\begin{array}{c} n \\ k+1 \end{array}\right)$ ways. Continuing this pattern we see that the number of binary vectors x that satisfy

$$\sum_{i=1}^{n} x_i \geq k$$

is given by

$$\sum_{l=k}^{n} \binom{n}{l} = \binom{n}{n} + \binom{n}{n-1} + \binom{n}{n-2} + \ldots + \binom{n}{k+1} + \binom{n}{k} .$$

Problem 6 (counting the number of increasing vectors)

If the first component x_1 were to equal n, then there is no possible vector that satisfies the inequality $x_1 < x_2 < x_3 < \ldots < x_k$ constraint. If the first component x_1 equals $n-1$ then again there are no vectors that satisfy the constraint. The first largest value that the component x_1 can take on and still result in a complete vector satisfying the inequality constraints is when $x_1 = n-k+1$ For that value of x_1, the other components are determined and are given by $x_2 = n-k+2$, $x_3 = n-k+3$, up to the value for x_k where $x_k = n$. This assignment provides *one* vector that satisfies the constraints. If $x_1 = n-k$, then we can construct an inequality satisfying vector x by assigning the $k-1$ other components x_2, x_3, up to x_k by assigning the integers $n-k+1\,, n-k+2\,, \ldots n-1\,, n$ to the $k-1$ components. This can be done in $\binom{k}{1}$ ways. Continuing if $x_1 = n-k-1$, then we can obtain a valid vector x by assign the integers $n-k\,, n-k+1\,, \ldots n-1\,, n$ to the $k-1$ other components of x. This can be seen as an equivalent problem to that of specifying two blanks from $n-(n-k)+1 = k+1$ spots and can be done in $\binom{k+1}{2}$ ways. Continuing to decrease the value of the x_1 component, we finally come to the case where we have n locations open for assignment with k assignments to be made (or equivalently $n-k$ blanks to be assigned) since this can be done in $\binom{n}{n-k}$ ways. Thus the total number of vectors is given by

$$1 + \binom{k}{1} + \binom{k+1}{2} + \binom{k+2}{3} + \ldots + \binom{n-1}{n-k-1} + \binom{n}{n-k} .$$

Problem 7 (choosing r from n by drawing subsets of size $r-1$)

Equation 4.1 from the book is given by

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} .$$

Considering the right hand side of this expression, we have

$$\begin{aligned}\binom{n-1}{r-1}+\binom{n-1}{r} &= \frac{(n-1)!}{(n-1-r+1)!(r-1)!}+\frac{(n-1)!}{(n-1-r)!r!}\\ &= \frac{(n-1)!}{(n-r)!(r-1)!}+\frac{(n-1)!}{(n-1-r)!r!}\\ &= \frac{n!}{(n-r)!r!}\left(\frac{r}{n}+\frac{n-r}{n}\right)\\ &= \binom{n}{r},\end{aligned}$$

and the result is proven.

Problem 8 (selecting r people from from n men and m women)

We desire to prove

$$\binom{n+m}{r}=\binom{n}{0}\binom{m}{r}+\binom{n}{1}\binom{m}{r-1}+\ldots+\binom{n}{r}\binom{m}{0}.$$

We can do this in a combinatorial way by considering subgroups of size r from a group of n men and m women. The left hand side of the above represents one way of obtaining this identity. Another way to count the number of subsets of size r is to consider the number of possible groups can be found by considering a subproblem of how many men chosen to be included in the subset of size r. This number can range from zero men to r men. When we have a subset of size r with zero men we must have all women. This can be done in $\binom{n}{0}\binom{m}{r}$ ways. If we select one man and $r-1$ women the number of subsets that meet this criterion is given by $\binom{n}{1}\binom{m}{r-1}$. Continuing this logic for all possible subset of the men we have the right hand side of the above expression.

Problem 9 (selecting n from $2n$)

From problem 8 we have that when $m = n$ and $r = n$ that

$$\binom{2n}{n}=\binom{n}{0}\binom{n}{n}+\binom{n}{1}\binom{n}{n-1}+\ldots+\binom{n}{n}\binom{n}{0}.$$

Using the fact that $\binom{n}{k}=\binom{n}{n-k}$ the above is becomes

$$\binom{2n}{n}=\binom{n}{0}^2+\binom{n}{1}^2+\ldots+\binom{n}{n}^2,$$

which is the desired result.

Problem 10 (committees with a chair)

Part (a): We can select a committee with k members in $\binom{n}{k}$ ways. Selecting a chairperson from the k committee members gives

$$k\binom{n}{k}$$

possible choices.

Part (b): If we choose the non chairperson members first this can be done in $\binom{n}{k-1}$ ways. We then choose the chairperson based on the remaining $n-k+1$ people. Combining these two we have

$$(n-k+1)\binom{n}{k-1}$$

possible choices.

Part (c): We can first pick the chair of our committee in n ways and then pick $k-1$ committee members in $\binom{n-1}{k-1}$. Combining the two we have

$$n\binom{n-1}{k-1},$$

possible choices.

Part (d): Since all expressions count the same thing they must be equal and we have

$$k\binom{n}{k} = (n-k+1)\binom{n}{k-1} = n\binom{n-1}{k-1}.$$

Part (e): We have

$$\begin{aligned} k\binom{n}{k} &= k\frac{n!}{(n-k)!k!} \\ &= \frac{n!}{(n-k)!(k-1)!} \\ &= \frac{n!(n-k+1)}{(n-k+1)!(k-1)!} \\ &= (n-k+1)\binom{n}{k-1} \end{aligned}$$

Factoring out n instead we have

$$\begin{aligned} k\binom{n}{k} &= k\frac{n!}{(n-k)!k!} \\ &= n\frac{(n-1)!}{(n-1-(k-1))!(k-1)!} \\ &= n\binom{n-1}{k-1} \end{aligned}$$

Problem 11 (Fermat's combinatorial identity)

We desire to prove the so called Fermat's combinatorial identity

$$\begin{aligned} \binom{n}{k} &= \sum_{i=k}^{n}\binom{i-1}{k-1} \\ &= \binom{k-1}{k-1}+\binom{k}{k-1}+\cdots+\binom{n-2}{k-1}+\binom{n-1}{k-1}. \end{aligned}$$

Following the hint, consider the integers $1, 2, \cdots, n$. Then consider subsets of size k from n elements as a sum over i where we consider i to be the largest entry in all the given subsets of size k. The smallest i can be is k of which there are $\binom{k-1}{k-1}$ subsets where when we add the element k we get a complete subset of size k. The next subset would have $k+1$ as the largest element of which there are $\binom{k}{k-1}$ of these. There are $\binom{k+1}{k-1}$ subsets with $k+2$ as the largest element etc. Finally, we will have $\binom{n-1}{k-1}$ sets with n the largest element. Summing all of these subsets up gives $\binom{n}{k}$.

Problem 12 (moments of the binomial coefficients)

Part (a): Consider n people from which we want to count the total number of committees of any size with a chairman. For a committee of size $k=1$ we have $1\cdot\binom{n}{1}=n$ possible choices. For a committee of size $k=2$ we have $\binom{n}{2}$ subsets of two people and two choices for the person who is the chair. This gives $2\binom{n}{2}$ possible choices. For a committee of size $k=3$ we have $3\binom{n}{3}$, etc. Summing all of these possible choices we find that the total number of committees with a chair is

$$\sum_{k=1}^{n}k\binom{n}{k}.$$

Another way to count the total number of all committees with a chair, is to consider first selecting the chairperson from which we have n choices and then considering all possible subsets of size $n-1$ (which is 2^{n-1}) from which to construct the remaining committee members. The product then gives $n2^{n-1}$.

Part (b): Consider again n people where now we want to count the total number of committees of size k with a chairperson and a secretary. We can select all subsets of size k in $\left(\begin{array}{c} n \\ k \end{array} \right)$ ways. Given a subset of size k, there are k choices for the chairperson and k choices for the secretary giving $k^2 \left(\begin{array}{c} n \\ k \end{array} \right)$ committees of size k with a chair and a secretary. The total number of these is then given by summing this result or

$$\sum_{k=1}^{n} k^2 \left(\begin{array}{c} n \\ k \end{array} \right) .$$

Now consider first selecting the chair which can be done in n ways. Then selecting the secretary which can either be the chair or one of the $n-1$ other people. If we select the chair and the secretary to be the same person we have $n-1$ people to choose from to represent the committee. All possible subsets from as set of $n-1$ elements is given by 2^{n-1}, giving in total $n2^{n-1}$ possible committees with the chair and the secretary the same person. If we select a different person for the secretary this chair/secretary selection can be done in $n(n-1)$ ways and then we look for all subsets of a set with $n-2$ elements (i.e. 2^{n-2}) so in total we have $n(n-1)2^{n-2}$. Combining these we obtain

$$n2^{n-1} + n(n-1)2^{n-2} = n2^{n-2}(2+n-1) = n(n+1)2^{n-2} .$$

Equating the two we have

$$\sum_{k=1}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) k^2 = 2^{n-2} n(n+1) .$$

Part (c): Consider now selecting all committees with a chair a secretary and a stenographer, where each can be the same person. Then following the results of Part (b) this total number is given by $\sum_{k=1}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) k^3$. Now consider the following situations and a count of how many cases they provide.

- If the same person is the chair, the secretary, and the stenographer, then this combination gives $n2^{n-1}$ total committees.
- If the same person is the chair and the secretary, but not the stenographer, then this combination gives $n(n-1)2^{n-2}$ total committees.
- If the same person is the chair and the stenographer, but not the secretary, then this combination gives $n(n-1)2^{n-2}$ total committees.
- If the same person is the secretary and the stenographer, but not the chair, then this combination gives $n(n-1)2^{n-2}$ total committees.

- Finally, if no person has more than one job, then this combination gives $n(n-1)(n-2)2^{n-3}$ total committees.

Adding all of these possible combinations up we find that

$$n(n-1)(n-2)2^{n-3} + 3n(n-1)2^{n-2} + n2^{n-1} = n^2(n+3)2^{n-3}.$$

Problem 13 (an alternating series of binomial coefficients)

From the binomial theorem we have

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

If we select $x = -1$ and $y = 1$ then $x + y = 0$ and the sum above becomes

$$0 = \sum_{k=0}^{n} \binom{n}{k} (-1)^k,$$

as we were asked to prove.

Problem 14 (committees and subcommittees)

Part (a): Pick the committee of size j in $\binom{n}{j}$ ways. The subcommittee of size i from these j can be selected in $\binom{j}{i}$ ways, giving a total of $\binom{j}{i}\binom{n}{j}$ committees and subcommittee. Now assume that we pick the subcommittee first. This can be done in $\binom{n}{i}$ ways. We then pick the committee in $\binom{n-i}{j-i}$ ways resulting in a total $\binom{n}{i}\binom{n-i}{j-i}$.

Part (b): I think that the lower index on this sum should start at i (the smallest subcommittee size). If so then we have

$$\begin{aligned}
\sum_{j=i}^{n} \binom{n}{j}\binom{j}{i} &= \sum_{j=i}^{n} \binom{n}{i}\binom{n-i}{j-i} \\
&= \binom{n}{i} \sum_{j=i}^{n} \binom{n-i}{j-i} \\
&= \binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} = \binom{n}{i} 2^{n-i}.
\end{aligned}$$

Part (c): Consider the following manipulations of a binomial like sum

$$
\begin{aligned}
\sum_{j=i}^{n} \binom{n}{j} \binom{j}{i} x^{j-i} y^{n-i-(j-i)} &= \sum_{j=i}^{n} \binom{n}{i} \binom{n-i}{j-i} x^{j-i} y^{n-j} \\
&= \binom{n}{i} \sum_{j=i}^{n} \binom{n-i}{j-i} x^{j-i} y^{n-j} \\
&= \binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} x^{j} y^{n-(j+i)} \\
&= \binom{n}{i} \sum_{j=0}^{n-i} \binom{n-i}{j} x^{j} y^{n-i-j} \\
&= \binom{n}{i} (x+y)^{n-i} .
\end{aligned}
$$

In summary we have shown that

$$
\sum_{j=i}^{n} \binom{n}{j} \binom{j}{i} x^{j-i} y^{n-j} = \binom{n}{i} (x+y)^{n-i} \quad \text{for} \quad i \leq n
$$

Now let $x = 1$ and $y = -1$ so that $x + y = 0$ and using these values in the above we have

$$
\sum_{j=i}^{n} \binom{n}{j} \binom{j}{i} (-1)^{n-j} = 0 \quad \text{for} \quad i \leq n .
$$

Problem 15 (the number of ordered vectors)

As stated in the problem we will let $H_k(n)$ be the number of vectors with components $x_1, x_2, \cdots, x_k$ for which each x_i is a positive integer such that $1 \leq x_i \leq n$ and the x_i are ordered i.e. $x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$

Part (a): Now $H_1(n)$ is the number of vectors with one component (with the restriction on its value of $1 \leq x_1 \leq n$). Thus there are n choices for x_1 so $H_1(n) = n$.

We can compute $H_k(n)$ by considering how many vectors there can be when the last component i.e. x_k has value of j. This would be the expression $H_{k-1}(j)$, since we know the value of the k-th component. Since j can range from 1 to n the total number of vectors with k components (i.e. $H_k(n)$) is given by the sum of all the previous $H_{k-1}(j)$. That is

$$
H_k(n) = \sum_{j=1}^{n} H_{k-1}(j) .
$$

Part (b): We desire to compute $H_3(5)$. To do so we first note that from the formula above the points at level k (the subscript) depends on the values of H at level $k-1$. To evaluate this expression when $n = 5$, we need to evaluate $H_k(n)$ for $k = 1$ and $k = 2$. We have that

$$\begin{aligned} H_1(n) &= n \\ H_2(n) &= \sum_{j=1}^{n} H_1(j) = \sum_{j=1}^{n} j = \frac{n(n+1)}{2} \\ H_3(n) &= \sum_{j=1}^{n} H_2(j) = \sum_{j=1}^{n} \frac{j(j+1)}{2}. \end{aligned}$$

Thus we can compute the first few values of $H_2(\cdot)$ as

$$\begin{aligned} H_2(1) &= 1 \\ H_2(2) &= 3 \\ H_2(3) &= 6 \\ H_2(4) &= 10 \\ H_2(5) &= 15. \end{aligned}$$

So that we find that

$$\begin{aligned} H_3(5) &= H_2(1) + H_2(2) + H_2(3) + H_2(4) + H_2(5) \\ &= 1 + 3 + 6 + 10 + 15 = 35. \end{aligned}$$

Problem 16 (the number of tied tournaments)

Part (a): See Table 2 for the enumerations used in computing $N(3)$. We have denoted A, B, and C by the people all in the first place.

Part (b): To argue the given sum, we consider how many outcomes there are when i-players tie for last place. To determine this we have to choose the i players from n that will tie (which can be done in $\binom{n}{i}$ ways). We then have to distributed the remaining $n - i$ players in winning combinations (with ties allowed). This can be done recursively in $N(n - i)$ ways. Summing up all of these terms we find that

$$N(n) = \sum_{i=1}^{n} \binom{n}{i} N(n-i).$$

Part (c): In the above expression let $j = n - i$, then our limits on the sum above change as follows

$$\begin{aligned} i = 1 &\rightarrow j = n - 1 \quad \text{and} \\ i = n &\rightarrow j = 0, \end{aligned}$$

so that the above sum for $N(n)$ becomes

$$N(n) = \sum_{j=0}^{n-1} \binom{n}{j} N(j).$$

First Place	Second Place	Third Place
A,B,C		
A,B	C	
A,C	B	
C,B	A	
A	B,C	
B	C,A	
C	A,B	
A	B	C
B	C	A
C	A	B
A	C	B
$\vdots$	$\vdots$	$\vdots$
B	A	C
C	B	A

Table 2: Here we have enumerated many of the possible ties that can happen with three people. The first row corresponds to all three in first place. The next three rows corresponds to two people in first place and the other in second place. The third row corresponds to two people in second place and one in first. The remaining rows correspond to one person in each position. The ellipses ($\vdots$) denotes thirteen possible outcomes.

Part (d): For the specific case of $N(3)$ we find that

$$\begin{aligned} N(3) &= \sum_{j=0}^{2} \binom{3}{j} N(j) \\ &= \binom{3}{0} N(0) + \binom{3}{1} N(1) + \binom{3}{2} N(2) \\ &= N(0) + 3N(1) + 3N(2) = 1 + 3(1) + 3(3) = 13\,. \end{aligned}$$

We also find for $N(4)$ that

$$\begin{aligned} N(4) &= \sum_{j=0}^{3} \binom{4}{j} N(j) \\ &= \binom{4}{0} N(0) + \binom{4}{1} N(1) + \binom{4}{2} N(2) + \binom{4}{3} N(3) \\ &= N(0) + 4N(1) + \frac{3 \cdot 4}{2} N(2) + 4N(3) = 1 + 4(1) + 6(3) + 4(13) = 75\,. \end{aligned}$$

Problem 17 (why the binomial equals the multinomial)

The expression $\binom{n}{r}$ is the number of ways to choose r objects from n, leaving another group of $n-r$ objects. The expression $\binom{n}{r, n-r}$ is the number of divisions of n distinct objects into two groups of size r and of size $n-r$ respectively. As these are the same thing the numbers are equivalent.

Problem 18 (a decomposition of the multinomial coefficient)

To compute $\left(\begin{array}{c} n \\ n_1, n_2, n_3, \cdots, n_r \end{array}\right)$ we consider fixing one particular object from the n. Then this object can end up in any of the r individual groups. If it appears in the first one then we have $\left(\begin{array}{c} n-1 \\ n_1-1, n_2, n_3, \cdots, n_r \end{array}\right)$, possible arrangements for the other objects. If it appears in the second group then the remaining objects can be distributed in $\left(\begin{array}{c} n-1 \\ n_1, n_2-1, n_3, \cdots, n_r \end{array}\right)$ ways, etc. Repeating this argument for all of the r groups we see that the original multinomial coefficient can be written as sums of these individual multinomial terms as

$$\begin{aligned} \left(\begin{array}{c} n \\ n_1, n_2, n_3, \cdots, n_r \end{array}\right) &= \left(\begin{array}{c} n-1 \\ n_1-1, n_2, n_3, \cdots, n_r \end{array}\right) \\ &+ \left(\begin{array}{c} n-1 \\ n_1, n_2-1, n_3, \cdots, n_r \end{array}\right) \\ &+ \cdots \\ &+ \left(\begin{array}{c} n-1 \\ n_1, n_2, n_3, \cdots, n_r-1 \end{array}\right). \end{aligned}$$

Problem 19 (the multinomial theorem)

The multinomial therm is

$$(x_1 + x_2 + \cdots + x_r)^n = \sum_{n_1+n_2+\cdots+n_r=n} \left(\begin{array}{c} n \\ n_1, n_2, \cdots, n_r \end{array}\right) x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r},$$

which can be proved by recognizing that the product of $(x_1 + x_2 + \cdots + x_r)^n$ will contain products of the type $x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$, and recognizing that the number of such terms, i.e. the coefficient in front of this term is a count of the number of times we can select n_1 of the variable x_1's, and n_2 of the variable x_2, etc from the n variable choices. Since this number equals the multinomial coefficient we have proven the multinomial theorem.

Problem 20 (the number of ways to fill bounded urns)

Let x_i be the number of balls in the ith urn. We must have $x_i \geq m_i$ and we are distributing the n balls so that $\sum_{i=1}^{r} x_i = n$. To solve this problem lets shift our variables so that each must be greater than or equal to zero. Our constraint then becomes (by subtracting the lower bound on x_i)

$$\sum_{i=1}^{r}(x_i - m_i) = n - \sum_{i=1}^{r} m_i\,.$$

This expression motivates us to define $v_i = x_i - m_i$. Then $v_i \geq 0$ so we are looking for the number of solutions to the equation

$$\sum_{i=1}^{r} v_i = n - \sum_{i=1}^{r} m_i\,,$$

where v_i must be greater than or equal to zero. This number is given by

$$\left(\begin{array}{c} n - \sum_{i=1}^{r} m_i + r - 1 \\ r - 1 \end{array}\right).$$

Problem 21 (**k zeros in an integer equation**)

To find the number of solutions to

$$x_1 + x_2 + \cdots + x_r = n\,,$$

where exactly k of the x_r's are zero, we can select k of the x_i's to be zero in $\left(\begin{array}{c} r \\ k \end{array}\right)$ ways and then count the number of solutions with positive (greater than or equal to one solutions) for the remaining $r - k$ variables. The number of solutions to the remaining equation is $\left(\begin{array}{c} n-1 \\ r-k-1 \end{array}\right)$ ways so that the total number is the product of the two or

$$\left(\begin{array}{c} r \\ k \end{array}\right)\left(\begin{array}{c} n-1 \\ r-k-1 \end{array}\right).$$

Problem 22 (the number of partial derivatives)

Let n_i be the number of derivatives taken of the x_ith variable. Then a total order of n derivatives requires that these componentwise derivatives satisfy $\sum_{i=1}^{n} n_i = n$, with $n_i \geq 0$. The number of such is given by

$$\left(\begin{array}{c} n + n - 1 \\ n - 1 \end{array}\right) = \left(\begin{array}{c} 2n - 1 \\ n - 1 \end{array}\right).$$

Problem 23 (counting discrete wedges)

We require that $x_i \geq 1$ and that they sum to a value less than k, i.e.

$$\sum_{i=1}^{n} x_i \leq k\,.$$

To count the number of solutions to this equation consider the number of equations with $x_i \geq 1$ and $\sum_{i=1}^n x_i = \hat{k}$, which is

$$\left(\begin{array}{c} \hat{k}-1 \\ n-1 \end{array} \right)$$

so to calculate the number of equations to the requested problem we add these up for all $\hat{k} < k$. The number of solutions is given by

$$\sum_{\hat{k}=n}^{k} \left(\begin{array}{c} \hat{k}-1 \\ n-1 \end{array} \right) \quad \text{with} \quad k > n\,.$$

Chapter 1: Self-Test Problems and Exercises

Problem 1 (counting arrangements of letters)

Part (a): Consider the pair of A with B as one object. Now there are two orderings of this "fused" object i.e. AB and BA. The remaining letters can be placed in 4! orderings and once an ordering is specified the fused A/B block can be in any of the five locations around the permutation of the letters $CDEF$. Thus we have $2 \cdot 4! \cdot 5 = 240$ total orderings.

Part (b): We want to enforce that A must be before B. Lets begin to construct a valid sequence of characters by first placing the other letters $CDEF$, which can be done in $4! = 24$ possible ways. Now consider an arbitrary permutation of $CDEF$ such as $DFCE$. Then if we place A in the left most position (such as as in $ADFCE$), we see that there are five possible locations for the letter B. For example we can have $ABDFCE$, $ADBFCE$, $ADFBCE$, $ADFCBE$, or $ADFCEB$. If A is located in the second position from the left (as in $DAFCE$) then there are four possible locations for B. Continuing this logic we see that we have a total of $5+4+3+2+1 = \frac{5(5+1)}{2} = 15$ possible ways to place A and B such that they are ordered with A before B in each permutation. Thus in total we have $15 \cdot 4! = 360$ total orderings.

Part (c): Lets solve this problem by placing A, then placing B and then placing C. Now we can place these characters at any of the six possible character locations. To explicitly specify their locations lets let the integer variables n_0, n_1, n_2, and n_3 denote the number of blanks (from our total of six) that are before the A, between the A and the B, between the B and the C, and after the C. By construction we must have each n_i satisfy

$$n_i \geq 0 \quad \text{for} \quad i = 0, 1, 2, 3\,.$$

In addition the sum of the n_i's plus the three spaces occupied by A, B, and C must add to six or

$$n_0 + n_1 + n_2 + n_3 + 3 = 6\,,$$

or equivalently

$$n_0 + n_1 + n_2 + n_3 = 3\,.$$

The number of solutions to such integer equalities is discussed in the book. Specifically, there are

$$\binom{3+4-1}{4-1} = \binom{6}{3} = 20\,,$$

such solutions. For each of these solutions, we have $3! = 6$ ways to place the three other letters giving a total of $6 \cdot 20 = 120$ arrangements.

Part (d): For this problem A must be before B and C must be before D. Let begin to construct a valid ordering by placing the letters E and F first. This can be done in two ways EF or FE. Next lets place the letters A and B, which if A is located at the left most position as in AEF, then B has three possible choices. As in Part (b) from this problem there are a total of $3+2+1=6$ ways to place A and B such that A comes before B. Following the same logic as in Part (b) above when we place C and D there are $5+4+3+2+1=15$ possible placements. In total then we have $15 \cdot 6 \cdot 2 = 180$ possible orderings.

Part (e): There are $2!$ ways of arranging A and B, $2!$ ways of arranging C and D, and $2!$ ways of arranging the remaining letters E and F. Lets us first place the blocks of letters consisting of the pair A and B which can be placed in any of the positions around E and F. There are three such positions. Next lets us place the block of letters consisting of C and D which can be placed in any of the four positions (between the E, F individual letters, or the A and B block). This gives a total number of arrangements of

$$2! \cdot 2! \cdot 2! \cdot 3 \cdot 4 = 96\,.$$

Part (f): E can be placed in any of five choices, first, second, third, fourth or fifth. Then the remaining blocks can be placed in $5!$ ways to get in total $5(5!) = 600$ arrangement's.

Problem 2 (counting seatings of people)

We have $4!$ arrangements of the Americans, $3!$ arrangements of the French, and $3!$ arrangements of the Britch and then $3!$ arrangements of these groups giving

$$4! \cdot 3! \cdot 3! \cdot 3!\,,$$

possible arrangements.

Problem 3 (counting presidents)

Part (a): With no restrictions we must select three people from ten. This can be done in $\binom{10}{3}$ ways. Then with these three people there are $3!$ ways to specify which person is the president, the treasurer, etc. Thus in total we have

$$\binom{10}{3} \cdot 3! = \frac{10!}{7!} = 720\,,$$

possible choices.

Part (b): If A and B will not searve together we can construct the total number of choices by considering clubs consisting of instances with A included but no B, B included by no A, and finally neither A or B included. This can be represented as

$$1 \cdot \left(\begin{array}{c} 8 \\ 2 \end{array}\right) + 1 \cdot \left(\begin{array}{c} 8 \\ 2 \end{array}\right) + \cdot \left(\begin{array}{c} 8 \\ 3 \end{array}\right) = 112 \,.$$

This result needs to again be multipled by 3! as in Part (a) of this problem. When we do so we find we obtain 672.

Part (c): In the same way as in Part (b) of this problem lets count first the number of clubs with C and D in them and second the number of clubs without C and D in them. This number is

$$\left(\begin{array}{c} 8 \\ 1 \end{array}\right) + \left(\begin{array}{c} 8 \\ 3 \end{array}\right) = 64 \,.$$

Again multiplying by 3! we find a total number of $3! \cdot 64 = 384$ clubs.

Part (d): For E to be an officer means that E must be selected as a club member. The number of other members that can be selected is given by $\left(\begin{array}{c} 9 \\ 2 \end{array}\right) = 36$. Again multiplying this by 3! gives a total of 216 clubs.

Part (e): If for F to serve F must be a president we have two cases. The first is where F serves and is the president and the second where F does not serve. When F is the president we have two permutations for the jobs of the other two selected members. When F does not serve, we have $3! = 6$ possible permutations in assigning titles amoung the selected people. In total then we have

$$2 \left(\begin{array}{c} 9 \\ 2 \end{array}\right) + 6 \left(\begin{array}{c} 9 \\ 3 \end{array}\right) = 576 \,,$$

possible clubs.

Problem 4 (anwsering questions)

She must select seven questions from ten, which can be done in $\left(\begin{array}{c} 10 \\ 7 \end{array}\right) = 120$ ways. If she must select at least three from the first five then she can choose to anwser three, four or all five of the questions. Counting each of these choices in tern, we find that she has

$$\left(\begin{array}{c} 5 \\ 3 \end{array}\right)\left(\begin{array}{c} 5 \\ 4 \end{array}\right) + \left(\begin{array}{c} 5 \\ 4 \end{array}\right)\left(\begin{array}{c} 5 \\ 3 \end{array}\right) + \left(\begin{array}{c} 5 \\ 5 \end{array}\right)\left(\begin{array}{c} 5 \\ 2 \end{array}\right) = 110 \,.$$

possible ways.

Problem 5 (dividing gifts)

We have $\left(\begin{array}{c}7\\3\end{array}\right)$ ways to select three gifts for the first child, then $\left(\begin{array}{c}4\\2\end{array}\right)$ ways to select two gifts for the second, and finally $\left(\begin{array}{c}2\\2\end{array}\right)$ for the third child. Giving a total of

$$\left(\begin{array}{c}7\\3\end{array}\right)\cdot\left(\begin{array}{c}4\\2\end{array}\right)\cdot\left(\begin{array}{c}2\\2\end{array}\right)=210\,,$$

arrangements.

Problem 6 (license plates)

We can pick the location of the three letters in $\left(\begin{array}{c}7\\3\end{array}\right)$ ways. Once these positions are selected we have 26^3 different combinations of letters that can be placed in the three spots. From the four remaining slots we can place 10^4 different digits giving in total

$$\left(\begin{array}{c}7\\3\end{array}\right)\cdot 26^3\cdot 10^4\,,$$

possible seven place license plates.

Problem 7 (a simple combinatorial argument)

Remember that the expression $\left(\begin{array}{c}n\\r\end{array}\right)$ counts the number of ways we can select r items from n. Notice that once we have specified a particular selection of r items, by construction we have also specified a particular selection of $n-r$ items, i.e. the remaining ones that are unselected. Since for each specification of r items we have an equivalent selection of $n-r$ items, the number of both i.e. $\left(\begin{array}{c}n\\r\end{array}\right)$ and $\left(\begin{array}{c}n\\n-r\end{array}\right)$ must be equal.

Problem 8 (counting n-digit numbers)

Part (a): To have no to consecutive digits equal, we can select the first digit in one of ten possible ways. The next digit in one of nine possible ways (we can't use the digit we selected for the first position). For the third digit we have three possible choices, etc. Thus in total we have

$$10\cdot 9\cdot 9\cdots 9=10\cdot 9^{n-1}\,,$$

possible digits.

Part (b): We now want to count the number of n-digit numbers where the digit 0 appears i times. Lets pick the locations where we want to place the zeros. This can be done in $\left(\begin{array}{c} n \\ i \end{array} \right)$ ways. We then have nine choices for the other digits to place in the other $n - i$ locations. This gives 9^{n-i} possible enoumerations for non-zero digits. In total then we have

$$\left(\begin{array}{c} n \\ i \end{array} \right) 9^{n-i} ,$$

n digit numbers with i zeros in them.

Problem 9 (selecting three students from three classes)

Part (a): To choose three students from $3n$ total students can be done in $\left(\begin{array}{c} 3n \\ 3 \end{array} \right)$ ways.

Part (b): To pick three students from the same class we must first pick the class to draw the student from. This can be done in $\left(\begin{array}{c} 3 \\ 1 \end{array} \right) = 3$ ways. Once the class has been picked we have to pick the three students in from the n in that class. This can be done in $\left(\begin{array}{c} n \\ 3 \end{array} \right)$ ways. Thus in total we have

$$3 \left(\begin{array}{c} n \\ 3 \end{array} \right) ,$$

possible selections of three students all from one class.

Part (c): To get two students in the same class and another in a different class, we must first pick the class from which to draw the two students from. This can be done in $\left(\begin{array}{c} 3 \\ 1 \end{array} \right) = 3$ ways. Next we pick the other class from which to draw the singleton student from. Since there are two possible classes to select this student from this can be done in two ways. Once both of these classes are selected we pick the individual two and one students from their respective classes in $\left(\begin{array}{c} n \\ 2 \end{array} \right)$ and $\left(\begin{array}{c} n \\ 1 \end{array} \right)$ ways respectively. Thus in total we have

$$3 \cdot 2 \cdot \left(\begin{array}{c} n \\ 2 \end{array} \right) \left(\begin{array}{c} n \\ 1 \end{array} \right) = 6n \frac{n(n-1)}{2} = 3n^2(n-1) ,$$

ways.

Part (d): Three students (all from a different class) can be picked in $\left(\begin{array}{c} n \\ 1 \end{array} \right)^3 = n^3$ ways.

Part (e): As an identity we have then that

$$\left(\begin{array}{c} 3n \\ 3 \end{array} \right) = 3 \left(\begin{array}{c} n \\ 3 \end{array} \right) + 3n^2(n-1) + n^3 .$$

We can check that this expression is correct by expanding each side. Expanding the left hand side we find that

$$\left(\begin{array}{c} 3n \\ 3 \end{array} \right) = \frac{3n!}{3!(3n-3)!} = \frac{3n(3n-1)(3n-2)}{6} = \frac{9n^3}{2} - \frac{9n^2}{2} + n \,.$$

While expanding the right hand side we find that

$$\begin{aligned} 3\left(\begin{array}{c} n \\ 3 \end{array} \right) + 3n^2(n-1) + n^3 &= 3\frac{n!}{3!(n-3)!} + 3n^3 - 3n^2 + n^3 \\ &= \frac{n(n-1)(n-2)}{2} + 4n^3 - 3n^2 \\ &= \frac{n(n^2-3n+2)}{2} + 4n^3 - 3n^2 \\ &= \frac{n^3}{2} - \frac{3n^2}{2} + n + 4n^3 - 3n^2 \\ &= \frac{9n^3}{2} - \frac{9n^2}{2} + n \,, \end{aligned}$$

which is the same, showing the equivalence.

Problem 10 (counting five digit numbers with no triple counts)

Lets first enumerate the number of five digit numbers that can be constructed with no repeated digits. Since we have nine choices for the first digit, eight choices for the second digit, seven choices for the third digit etc. We find the number of five digit numbers with no repeated digits given by $9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = \frac{9!}{4!} = 15120$.

Now lets count the number of five digit numbers where *one* of the digits $1, 2, 3, \cdots, 9$ repeats. We can pick the digit that will repeat in nine ways and select its position in the five digits in $\left(\begin{array}{c} 5 \\ 2 \end{array} \right)$ ways. To fill the remaining three digit location can be done in $8 \cdot 7 \cdot 6$ ways. This gives in total

$$9 \cdot \left(\begin{array}{c} 5 \\ 2 \end{array} \right) \cdot 8 \cdot 7 \cdot 6 = 30240 \,.$$

Lets now count the number five digit numbers with two repeated digits. To compute this we might argue as follows. We can select the first digit and its location in $9 \cdot \left(\begin{array}{c} 5 \\ 2 \end{array} \right)$ ways. We can select the second repeated digit and its location in $8 \cdot \left(\begin{array}{c} 3 \\ 2 \end{array} \right)$ ways. The final digit can be selected in seven ways, giving in total

$$9 \left(\begin{array}{c} 5 \\ 2 \end{array} \right) \cdot 8 \left(\begin{array}{c} 3 \\ 2 \end{array} \right) \cdot 7 = 15120 \,.$$

We note, however, that this analysis (as it stands) double counts the true number of five digits numbers with two repeated digits. This is because in first selecting the first digit from nine classes and then selecting the second digit from eight choices the total two digits chosen can actually be selected in the opposite order but placed in same spots from among our five digits. Thus we have to divide the above number by two giving

$$\frac{15120}{2} = 7560\,.$$

So in total we have by summing up all these mutually exclusive events we find that the total number of five digit numbers allowing repeated digits is given by

$$9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 + 9 \left(\begin{array}{c} 5 \\ 2 \end{array} \right) \cdot 8 \cdot 7 \cdot 6 + \frac{1}{2} \cdot 9 \cdot \left(\begin{array}{c} 5 \\ 2 \end{array} \right) 8 \left(\begin{array}{c} 3 \\ 2 \end{array} \right) \cdot 7 = 52920\,.$$

Problem 11 (counting first round winners)

Lets consider a simple case first and then generalize this result. Consider some symbolic players denoted by A, B, C, D, E, F. Then we can construct a pairing of players by first selecting three players and then ordering the remaining three players with respect to the first chosen three. For example, lets first select the players B, E, and F. Then if we want A to play E, C to play F, and D to play B we can represent this graphically by the following

$$\begin{array}{c} B\,E\,F \\ D\,A\,C\,, \end{array}$$

where the players in a given fixed column play each other. From this we can select three different winners by selecting who wins each match. This can be done in 2^3 total ways. Since we have two possible choices for the winner of the first match, two possible choices for the winner of the second match, and finally two possible choices for the winner of the third match. Thus two generalize this procedure to $2n$ people we must first select n players from the $2n$ to for the "template" first row. This can be done in $\left(\begin{array}{c} 2n \\ n \end{array} \right)$ ways. We then must select one of the $n!$ orderings of the remaining n players to form matches with. Finally, we must select winners of each match in 2^n ways. In total we would then conclude that we have

$$\left(\begin{array}{c} 2n \\ n \end{array} \right) \cdot n! \cdot 2^n = \frac{(2n)!}{n!} \cdot 2^n\,,$$

total first round results. The problem with this is that it will double count the total number of pairings. It will count the pairs AB and BA as distinct. To remove this over counting we need to divide by the total number of ordered n pairs. This number is 2^n. When we divide by this we find that the total number of first round results is given by

$$\frac{(2n)!}{n!}\,.$$

Problem 12 (selecting committees)

Since we must select a total of six people consisting of at least three women and two men, we could select a committee with four women and two mean *or* a committee with three woman and three men. The number of ways of selecting this first type of committee is given by $\binom{8}{4}\binom{7}{2}$. The number of ways to select the second type of committee is given by $\binom{8}{3}\binom{7}{3}$. So the total number of ways to select a committee of six people is

$$\binom{8}{4}\binom{7}{2} + \binom{8}{3}\binom{7}{3}$$

Problem 13 (the number of different art sales)

Let D_i be the number of Dalis collected/bought by the i-th collector, G_i be the number of van Goghs collected by the i-th collector, and finally P_i the number of Picassos' collected by the i-th collector when $i = 1, 2, 3, 4, 5$. Then since all paintings are sold we have the following constraints on D_i, G_i, and P_i,

$$\sum_{i=1}^{5} D_i = 4\,, \quad \sum_{i=1}^{5} G_i = 5\,, \quad \sum_{i=1}^{5} P_i = 6\,.$$

Along with the requirements that $D_i \geq 0$, $G_i \geq 0$, and $P_i \geq 0$. Remembering that the number of solutions to an equation like

$$x_1 + x_2 + \cdot + x_r = n\,,$$

when $x_i \geq 0$ is given by $\binom{n+r-1}{r-1}$. Thus the number of solutions to the first equation above is given by $\binom{4+5-1}{5-1} = \binom{8}{4} = 70$, the number of solutions to the second equation is given by $\binom{5+5-1}{5-1} = \binom{9}{4} = 126$, and finally the number of solutions to the third equation is given by $\binom{6+5-1}{5-1} = \binom{10}{4} = 210$. Thus the total number of solutions is given by the product of these three numbers. We find that

$$\binom{8}{4}\binom{9}{4}\binom{10}{4} = 1852200\,,$$

See the Matlab file `chap_1_st_13.m` for these calculations.

Problem 14 (counting vectors that sum to less than k)

We want to evaluate the number of solutions to $\sum_{i=1}^{n} x_i \leq k$ for $k \geq n$, and x_i a positive integer. Now since the smallest value that $\sum_{i=1}^{n} x_i$ can be under these conditions is given when $x_i = 1$ for all i and gives a resulting sum of n. Now we note that for this problem the sum $\sum_{i=1}^{n} x_i$ take on any value greater than n up to and including k. Consider the number of solutions to $\sum_{i=1}^n x_i = j$ when j is fixed such that $n \le j \le k$. This number is given by $\left(\begin{array}{c} j-1 \\ n-1 \end{array} \right)$. So the total number of solutions is given by summing this expression over j for j ranging from n to k. We then find the total number of vectors $(x_1, x_2, \cdots, x_n)$ such that each x_i is a positive integer and $\sum_{i=1}^n x_i \le k$ is given by

$$\sum_{j=n}^{k} \left(\begin{array}{c} j-1 \\ n-1 \end{array} \right) .$$

Problem 15 (all possible passing students)

With n total students, lets assume that k people pass the test. These k students can be selected in $\left(\begin{array}{c} n \\ k \end{array} \right)$ ways. All possible orderings or rankings of these k people is given by $k!$ so that the we have

$$\left(\begin{array}{c} n \\ k \end{array} \right) k! \,,$$

different possible orderings when k people pass the test. Then the total number of possible test postings is given by

$$\sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) k! \,.$$

Problem 16 (subsets that contain at least one number)

There are $\left(\begin{array}{c} 20 \\ 4 \end{array} \right)$ subsets of size four. The number of subsets that contain at least one of the elements $1, 2, 3, 4, 5$ is the *complement* of the number of subsets that don't contain any of the elements $1, 2, 3, 4, 5$. This number is $\left(\begin{array}{c} 15 \\ 4 \end{array} \right)$, so the total number of subsets that contain at least one of $1, 2, 3, 4, 5$ is given by

$$\left(\begin{array}{c} 20 \\ 4 \end{array} \right) - \left(\begin{array}{c} 15 \\ 4 \end{array} \right) = 4845 - 1365 = 3480 \,.$$

Problem 17 (a simple combinatorial identity)

To show that

$$\left(\begin{array}{c} n \\ 2 \end{array} \right) = \left(\begin{array}{c} k \\ 2 \end{array} \right) + k(n-k) + \left(\begin{array}{c} n-k \\ 2 \end{array} \right) \quad \text{for} \quad 1 \le k \le n \,,$$

is true, begin by expanding the right hand side (RHS) of this expression. Using the definition of the binomial coefficients we obtain

$$
\begin{aligned}
\text{RHS} &= \frac{k!}{2!(k-2)!} + k(n-k) + \frac{(n-k)!}{2!(n-k-2)!} \\
&= \frac{k(k-1)}{2} + k(n-k) + \frac{(n-k)(n-k-1)}{2} \\
&= \frac{1}{2}\left(k^2 - k + kn - k^2 + n^2 - nk - n - kn + k^2 + k\right) \\
&= \frac{1}{2}\left(n^2 - n\right) .
\end{aligned}
$$

Which we can recognize as equivalent to $\left(\begin{array}{c} n \\ 2 \end{array}\right)$ since from its definition we have that

$$
\left(\begin{array}{c} n \\ 2 \end{array}\right) = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2} .
$$

proving the desired equivalence. A combinatorial argument for this expression can be given in the following way. The left hand side $\left(\begin{array}{c} n \\ 2 \end{array}\right)$ represents the number of ways to select two items from n. Now for any k (with $1 \le k \le n$) we can think about the entire set of n items as being divided into two parts. The first part will have k items and the second part will have the remaining $n-k$ items. Then by considering all possible halves the two items selected could come from will yield the decomposition shown on the right hand side of the above. For example, we can draw our two items from the initial k in the first half in $\left(\begin{array}{c} k \\ 2 \end{array}\right)$ ways, from the second half (which has $n-k$ elements) in $\left(\begin{array}{c} n-k \\ 2 \end{array}\right)$ ways, or by drawing one element from the set with k elements and another element from the set with $n-k$ elements, in $k(n-k)$ ways. Summing all of these terms together gives

$$
\left(\begin{array}{c} k \\ 2 \end{array}\right) + k(n-k) + \left(\begin{array}{c} n-k \\ 2 \end{array}\right) \quad \text{for} \quad 1 \le k \le n ,
$$

as an equivalent expression for $\left(\begin{array}{c} n \\ 2 \end{array}\right)$.

Chapter 2 (Axioms of Probability)

Chapter 2: Problems

Problem 1 (the sample space)

The sample space consists of the possible experimental outcomes, which in this case is given by

$$\{(R,R),(R,G),(R,B),(G,R),(G,G),(G,B),(B,R),(B,G),(B,B)\}\,.$$

If the first marble is not replaced then our sample space loses all "paired" terms in the above (i.e. terms like (R,R)) and it becomes

$$\{(R,G),(R,B),(G,R),(G,B),(B,R),(B,G)\}\,.$$

Problem 2 (the sample space of continually rolling a die)

The sample space consists of all possible die rolls to obtain a six. For example we have

$$\{(6),(1,6),(2,6),(3,6),(4,6),(5,6),(1,1,6),(1,2,6),\cdots,(2,1,6),(2,2,6)\cdots\}$$

The points in E_n are all sequences of rolls with n elements in them, so that $\cup_1^\infty E_n$ is all possible sequences ending with a six. Since a six must happen eventually, we have $(\cup_1^\infty E_n)^c = \phi$.

Problem 8 (mutually exclusive events)

Since A and B are mutually exclusive then $P(A\cup B) = P(A) + P(B)$.

Part (a): To calculate the probability that either A or B occurs we evaluate $P(A\cup B) = P(A) + P(B) = 0.3 + 0.5 = 0.8$

Part (b): To calculate the probability that A occurs but B does not we want to evaluate $P(A\backslash B)$. This can be done by considering

$$P(A\cup B) = P(B\cup(A\backslash B)) = P(B) + P(A\backslash B)\,,$$

where the last equality is due to the fact that B and $A\backslash B$ are mutually independent. Using what we found from part (a) $P(A\cup B) = P(A) + P(B)$, the above gives

$$P(A\backslash B) = P(A) + P(B) - P(B) = P(A) = 0.3\,.$$

Part (c): To calculate the probability that both A and B occurs we want to evaluate $P(A \cap B)$, which can be found by using

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \, .$$

Using what we know in the above we have that

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.8 - 0.3 - 0.5 = 0 \, ,$$

Problem 9 (accepting credit cards)

Let A be the event that a person carries the American Express card and B be the event that a person carries the VISA card. Then we want to evaluate $P(\cup B)$, the probability that a person carries the American Express card or the person carries the VISA card. This can be calculated as

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.24 + 0.61 - 0.11 = 0.74 \, .$$

Problem 10 (wearing rings and necklaces)

Let $P(A)$ be the probability that a student wears a ring. Let $P(B)$ be the probability that a student wears a necklace. Then from the information given we have that

$$\begin{aligned} P(A) &= 0.2 \\ P(B) &= 0.3 \\ P((A \cup B)^c) &= 0.3 \, . \end{aligned}$$

Part (a): We desire to calculate for this subproblem $P(A \cup B)$, which is given by

$$P(A \cup B) = 1 - P((A \cup B)^c) = 1 - 0.6 = 0.4 \, ,$$

Part (b): We desire to calculate for this subproblem $P(AB)$, which can be calculated by using the inclusion/exclusion identity for two sets which is

$$P(A \cup B) = P(A) + P(B) - P(AB) \, .$$

so solving for $P(AB)$ in the above we find that

$$P(AB) = P(A) + P(B) - P(A \cup B) = 0.2 + 0.3 - 0.4 = 0.1 \, .$$

Problem 11 (smoking cigarettes v.s cigars)

Let A be the event that a male smokes cigarettes and let B be the event that a male smokes cigars. Then the data given is that $P(A) = 0.28$, $P(B) = 0.07$, and $P(AB) = 0.05$.

Part (a): We desire to calculate for this subproblem $P((A \cup B)^c)$, which is given by (using the inclusion/exclusion identity for two sets)

$$\begin{aligned} P((A \cup B)^c) &= 1 - P(A \cup B) \\ &= 1 - (P(A) + P(B) - P(AB)) \\ &= 1 - 0.28 - 0.07 + 0.05 = 0.7 \,. \end{aligned}$$

Part (b): We desire to calculate for this subproblem $P(B \cap A^c)$ We will compute this from the identity

$$P(B) = P((B \cap A^c) \cup (B \cap A)) = P(B \cap A^c) + P(B \cap A) \,,$$

since the events $B \cap A^c$ and $B \cap A$ are mutually exclusive. With this identity we see that the event that we desire the probability of $(B \cap A^c)$ is given by

$$P(B \cap A^c) = P(B) - P(A \cap B) = 0.07 - 0.05 = 0.02 \,.$$

Problem 12 (language probabilities)

Let S be the event that a student is in a Spanish class, let F be the event that a student is in a French class and let G be the event that a student is in a German class. From the data given we have that

$$\begin{aligned} P(S) &= 0.28 \,, \quad P(F) = 0.26 \,, \quad P(G) = 0.16 \\ P(S \cap F) &= 0.12 \,, \quad P(S \cap G) = 0.04 \,, \quad P(F \cap G) = 0.06 \\ P(S \cap F \cap G) &= 0.02 \,. \end{aligned}$$

Part (a): We desire to compute

$$P(\neg(S \cup F \cup G)) = 1 - P(S \cup F \cup G) \,.$$

Define the event A to be $A = S \cup F \cup G$, then we will use the inclusion/exclusion identity for three sets which expresses $P(A) = P(S \cup F \cup G)$ in terms of set intersections as

$$\begin{aligned} P(A) &= P(S) + P(F) + P(G) - P(S \cap F) - P(S \cap G) - P(F \cap G) + P(S \cap F \cap G) \\ &= 0.28 + 0.26 + 0.16 - 0.12 - 0.04 - 0.06 + 0.02 = 0.5 \,. \end{aligned}$$

So that we have that $P(\neg(S \cup F \cup G)) = 1 - 0.5 = 0.5$.

Part (b): Using the definitions of the events above for this subproblem we want to compute

$$P(S \cap (\neg F) \cap (\neg G)) \,, \quad P((\neg S) \cap F \cap (\neg G)) \,, \quad P((\neg S) \cap (\neg F) \cap G) \,.$$

As these are all of the same form, lets first consider $P(S \cap (\neg F) \cap (\neg G))$, which equals $P(S \cap (\neg(F \cup G)))$. Now decomposing S into two disjoint sets $S \cap (\neg(F \cup G))$ and $S \cap (F \cup G)$ we see that $P(S)$ can be written as

$$P(S) = P(S \cap (\neg(F \cup G))) + P(S \cap (F \cup G)) \,.$$

Now since we know $P(S)$ if we knew $P(S \cap (F \cup G))$ we can compute the desired probability. Distributing the intersection in $S \cap (F \cup G)$, we see that we can write this set as

$$S \cap (F \cup G) = (S \cap F) \cup (S \cap G) .$$

So that $P(S \cap (F \cup G))$ can be computed (using the inclusion/exclusion identity) as

$$\begin{aligned} P(S \cap (F \cup G)) &= P((S \cap F) \cup (S \cap G)) \\ &= P(S \cap F) + P(S \cap G) - P((S \cap F) \cap (S \cap G)) \\ &= P(S \cap F) + P(S \cap G) - P(S \cap F \cap G) \\ &= 0.12 + 0.04 - 0.02 = 0.14 . \end{aligned}$$

Thus

$$\begin{aligned} P(S \cap (\neg(F \cup G))) &= P(S) - P(S \cap (F \cup G)) \\ &= 0.28 - 0.14 = 0.14 . \end{aligned}$$

In the same way we find that

$$\begin{aligned} P((\neg S) \cap F \cap (\neg G)) &= P(F) - P(F \cap (S \cup G)) \\ &= P(F) - (P(F \cap S) + P(F \cap G) - P(F \cap S \cap G) \\ &= 0.26 - 0.12 - 0.06 + 0.02 = 0.1 . \end{aligned}$$

and that

$$\begin{aligned} P((\neg S) \cap (\neg F) \cap G) &= P(G) - P(G \cap (S \cup F)) \\ &= P(G) - (P(G \cap S) + P(G \cap F) - P(S \cap F \cap G) \\ &= 0.16 - 0.04 - 0.06 + 0.02 = 0.08 . \end{aligned}$$

With all of these intermediate results we can compute that the probability that a student is taking exactly one language class is given by the sum of the probabilities of the three events introduced at the start of this subproblem. We find that this sum is given by

$$0.14 + 0.1 + 0.08 = 0.32 .$$

Part (c): If two students are chosen randomly the probability that at least one of them is taking a language class is the complement of the probability that neither is taking a language class. From part a of this problem we know that fifty students are not taking a language class, from the one hundred students at the school. Therefore the probability that we select two students *both* not in a language class is given by

$$\frac{\binom{50}{2}}{\binom{100}{2}} = \frac{1225}{4950} = \frac{49}{198} ,$$

thus the probability of drawing two students at least one of which is in a language class is given by

$$1 - \frac{49}{198} = \frac{149}{198} .$$

Problem 13 (the number of paper readers)

Before we begin to solve this problem lets take the given probabilities of *intersections* of events and convert them into probabilities of *unions* of events. Then if we need these values later in the problem we will have them. This can be done with the inclusion-exclusion identity. For two general sets A and B the inclusion-exclusion identity is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) .$$

Using this we can evaluate the probabilities of union of events.

$$\begin{aligned}
P(\text{II} \cup \text{III}) &= P(\text{II}) + P(\text{III}) - P(\text{II} \cap \text{III}) = 0.3 + 0.05 - 0.04 = 0.31 \\
P(\text{I} \cup \text{II}) &= P(\text{I}) + P(\text{II}) - P(\text{I} \cap \text{II}) = 0.1 + 0.3 - 0.08 = 0.32 \\
P(\text{I} \cup \text{III}) &= P(\text{I}) + P(\text{III}) - P(\text{I} \cap \text{III}) = 0.1 + 0.05 - 0.02 = 0.13 \\
P(\text{I} \cup \text{II} \cup \text{III}) &= P(\text{I}) + P(\text{II}) + P(\text{III}) - P(\text{I} \cap \text{II}) - P(\text{I} \cap \text{III}) \\
&- P(\text{II} \cap \text{III}) + P(\text{I} \cap \text{II} \cap \text{III}) \\
&= 0.1 + 0.3 + 0.05 - 0.08 - 0.02 - 0.04 + 0.01 = 0.32 .
\end{aligned}$$

We will now use these results in the following wherever needed.

Part (a): The requested proportion of people who read only one paper can be represented from three disjoint probabilities/proportions:

1. $P(\text{I} \cap \neg\text{II} \cap \neg\text{III})$ which represents the proportion of people who only read paper I.
2. $P(\neg\text{I} \cap \text{II} \cap \neg\text{III})$ which represents the proportion of people who only read paper II.
3. $P(\neg\text{I} \cap \neg\text{II} \cap \text{III})$ which represents the proportion of people who only read paper III.

The sum of these three probabilities will be the total number of people who read only one newspaper. To compute the first probability ($P(\text{I} \cap \neg\text{II} \cap \neg\text{III})$) we begin by noting that

$$P(\text{I} \cap \neg\text{II} \cap \neg\text{III}) + P(\text{I} \cap \neg(\neg\text{II} \cap \neg\text{III})) = P(\text{I}) ,$$

which is true since we can obtain the event I by intersecting it with two sets that union to the entire sample space i.e. $\neg\text{II} \cap \neg\text{III}$, and its negation $\neg(\neg\text{II} \cap \neg\text{III})$. With this expression we can evaluate our desired probability $P(\text{I}\cap\neg\text{II}\cap\neg\text{III})$ using the above. Simple subtraction gives

$$\begin{aligned}
P(\text{I} \cap \neg\text{II} \cap \neg\text{III}) &= P(\text{I}) - P(\text{I} \cap \neg(\neg\text{II} \cap \neg\text{III})) \\
&= P(\text{I}) - P(\text{I} \cap (\text{II} \cup \text{III})) \\
&= P(\text{I}) - P((\text{I} \cap \text{II}) \cup (\text{I} \cap \text{III})) .
\end{aligned}$$

Where the last two equations follows from the first by some simple set theory. Since the problem statement gives the probabilities of the events $\text{I} \cap \text{II}$ and $\text{I} \cap \text{III}$, to be able to further evaluate the right hand side of the expression above requires the ability to compute probabilities of unions of such sets. This can be done with the inclusion-exclusion identity which for two general sets A and B is given by $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Thus the above desired probability then becomes

$$\begin{aligned} P(\text{I} \cap \neg\text{II} \cap \neg\text{III}) &= P(\text{I}) - P(\text{I} \cap \text{II}) - P(\text{I} \cap \text{III}) + P((\text{I} \cap \text{II}) \cap (\text{I} \cap \text{III})) \\ &= P(\text{I}) - P(\text{I} \cap \text{II}) - P(\text{I} \cap \text{III}) + P(\text{I} \cap \text{II} \cap \text{III}) \\ &= 0.1 - 0.08 - 0.02 + 0.01 = 0.01 \,, \end{aligned}$$

using the numbers provided. For the probability $P(\neg\text{I} \cap \text{II} \cap \neg\text{III})$ of we can use the work earlier with the substitutions

$$\begin{aligned} \text{I} &\to \text{II} \\ \text{II} &\to \text{I} \,. \end{aligned}$$

Since in the first probability we computed the event *not* negated is event I, while in the second probability this is event II. This substitution gives

$$\begin{aligned} P(\neg\text{I} \cap \text{II} \cap \neg\text{III}) &= P(\text{II}) - P(\text{II} \cap \text{I}) - P(\text{II} \cap \text{III}) + P(\text{II} \cap \text{I} \cap \text{III}) \\ &= 0.3 - 0.08 - 0.04 + 0.01 = 0.19 \,, \end{aligned}$$

For the probability $P(\neg\text{I} \cap \neg\text{II} \cap \text{III})$ of we can use the work earlier with the substitutions

$$\begin{aligned} \text{I} &\to \text{III} \\ \text{III} &\to \text{I} \,. \end{aligned}$$

To give

$$\begin{aligned} P(\neg\text{I} \cap \neg\text{II} \cap \text{III}) &= P(\text{III}) - P(\text{III} \cap \text{II}) - P(\text{III} \cap \text{I}) + P(\text{I} \cap \text{II} \cap \text{III}) \\ &= 0.05 - 0.04 - 0.02 + 0.01 = 0.00 \,. \end{aligned}$$

Finally the number of people who read only one newspaper is given by

$$0.01 + 0.19 + 0.00 = 0.2 \,,$$

so the number of people who read only one newspaper is given by $0.2 \times 10^5 = 20,000$.

Part (b): The requested proportion of people who read at least two newspapers can be represented from three disjoint probabilities/proportions:

1. $P(\text{I} \cap \text{II} \cap \neg\text{III})$
2. $P(\text{I} \cap \neg\text{II} \cap \text{III})$
3. $P(\neg\text{I} \cap \text{II} \cap \text{III})$
4. $P(\text{I} \cap \text{II} \cap \text{III})$

We can compute each in the following ways. For the first probability we note that

$$\begin{aligned} P(\neg \text{I} \cap \text{II} \cap \text{III}) + P(\text{I} \cap \text{II} \cap \text{III}) &= P(\text{II} \cap \text{III}) \\ &= P(\text{II}) + P(\text{III}) - P(\text{II} \cup \text{III}) \\ &= 0.3 + 0.5 - 0.31 = 0.04\,. \end{aligned}$$

So that $P(\neg \text{I} \cap \text{II} \cap \text{III}) = 0.04 - P(\text{I} \cap \text{II} \cap \text{III}) = 0.04 - 0.01 = 0.03$. Using this we find that

$$\begin{aligned} P(\text{I} \cap \neg \text{II} \cap \text{III}) &= P(\text{I} \cap \text{III}) - P(\text{I} \cap \text{II} \cap \text{III}) \\ &= P(\text{I}) + P(\text{III}) - P(\text{I} \cup \text{III}) - P(\text{I} \cap \text{II} \cap \text{III}) \\ &= 0.1 + 0.5 - 0.13 - 0.01 = 0.01\,, \end{aligned}$$

and that

$$\begin{aligned} P(\text{I} \cap \text{II} \cap \neg \text{III}) &= P(\text{I} \cap \text{II}) - P(\text{I} \cap \text{II} \cap \text{III}) \\ &= P(\text{I}) + P(\text{II}) - P(\text{I} \cup \text{II}) - P(\text{I} \cap \text{II} \cap \text{III}) \\ &= 0.1 + 0.3 - 0.32 - 0.01 = 0.07\,. \end{aligned} \tag{1}$$

We also have $P(\text{I} \cap \text{II} \cap \text{III}) = 0.01$, from the problem statement. Combining all of this information the total percentage of people that read at least two newspapers is given by

$$0.03 + 0.01 + 0.07 + 0.01 = 0.12\,,$$

so the total number of people is given by $0.12 \times 10^5 = 12000$.

Part (c): For this part we to compute $P((\text{I} \cap \text{II}) \cup (\text{III} \cap \text{II}))$, which gives

$$\begin{aligned} P((\text{I} \cap \text{II}) \cup (\text{III} \cap \text{II})) &= P(\text{I} \cap \text{II}) + P(\text{III} \cap \text{II}) - P(\text{I} \cap \text{II} \cap \text{III}) \\ &= 0.08 + 0.04 - 0.01 = 0.11\,, \end{aligned}$$

so the number of people read at least one morning paper and one evening paper is $0.11 \times 10^5 = 11000$.

Part (d): To not read any newspaper we are looking for

$$1 - P(\text{I} \cup \text{II} \cup \text{III}) = 1 - 0.32 = 0.68\,,$$

so the number of people is 68000.

Part (e): To read only one morning paper and one evening paper is expressed as

$$P(\text{I} \cap \text{II} \cap \neg \text{III}) + P(\neg \text{I} \cap \text{II} \cap \text{III})\,.$$

The first expression has been calculated in Equation 1 as 0.07, while the second expansion has been calculated as 0.03 giving a total 0.01 giving a total of 10000 people who read I as their morning paper and II as their evening paper or who read III as their morning paper and II as their evening paper.

Problem 14 (an inconsistent study)

Following the hint given in the book, we let M denote the set of people who are married, W the set of people who are working professionals, and G the set of people who are college graduates. If we choose a random person and ask what the probability that he/she is either married or working or a graduate we are looking to compute $P(M \cup W \cup G)$. By the inclusion/exclusion theorem we have that the probability of this event is given by

$$\begin{aligned} P(M \cup W \cup G) &= P(M) + P(W) + P(G) \\ &- P(M \cap W) - P(M \cap G) - P(W \cap G) \\ &+ P(M \cap W \cap G) . \end{aligned}$$

From the given data each individual event probability can be estimated as

$$P(M) = \frac{470}{1000}, \quad P(G) = \frac{525}{1000}, \quad P(W) = \frac{312}{1000}$$

and each pairwise event probability can be estimated as

$$P(M \cap G) = \frac{147}{1000}, \quad P(M \cap W) = \frac{86}{1000}, \quad P(W \cap G) = \frac{42}{1000}$$

Finally the three-way event probability can be estimated as

$$P(M \cap W \cap G) = \frac{25}{1000} .$$

Using these numbers in the inclusion/exclusion formula above we find that

$$\begin{aligned} P(M \cup W \cup G) &= 0.47 + 0.525 + 0.312 - 0.147 - 0.086 - 0.042 + 0.025 \\ &= 1.057 > 1 , \end{aligned}$$

in contradiction to the rules of probability.

Problem 15 (probabilities of various poker hands)

Part (a): We must count the number of ways to obtain five cards of the same suit. We can first pick the suit in $\binom{4}{1} = 4$ ways after which we must pick five cards in $\binom{13}{5}$ ways. So in total we have

$$4 \binom{13}{5} = 5148 ,$$

ways to pick cards in a flush giving a probability of

$$\frac{4 \binom{13}{5}}{\binom{52}{5}} = 0.00198 .$$

Part (b): We can select the first denomination "a" in thirteen ways with $\binom{4}{2}$ ways to obtain the faces for these two cards. We can select the second denomination "b" in twelve ways with $\binom{4}{1}$ possible faces, the third denomination in eleven ways with four faces, the fourth denomination in ten ways again with four possible faces. The selection of the cards "b", "c", and "d" can be permuted in any of the 3! ways and the same hand results. Thus we have in total for the number of paired hands the following count

$$\frac{13\binom{4}{2}\cdot 12\binom{4}{1}\cdot 11\binom{4}{1}\cdot 10\binom{4}{1}}{3!} = 1098240\,.$$

Giving a probability of 0.42256.

Part (c): To calculate the number of hands with two pairs we have $\binom{13}{1}\binom{4}{2}$ ways to select the "a" pair. Then $\binom{12}{1}\binom{4}{2}$ ways to select the "b" pair. Since first selecting the "a" pair and then the "b" pair results in the same hand as selecting the "b" pair and then the "a" pair this direct product over counts the total number of "a" and "b" pairs by $2! = 2$. Finally, we have $\binom{11}{1}\binom{4}{1}$ ways to pick the last card in the hand. Thus we have

$$\frac{\binom{13}{1}\binom{4}{2}\cdot\binom{12}{1}\binom{4}{2}}{2!}\cdot\binom{11}{1}\binom{4}{1} = 123552\,,$$

total number of hands. Giving a probability of 0.04754.

Part (d): We have $\binom{13}{1}\binom{4}{3}$ ways to pick the "a" triplet. We can then pick "b" in $\binom{12}{1}\cdot 4$ and pick "c" in $\binom{11}{1}\cdot 4$. This combination over counts by two so that the total number of three of a kind hands is given by

$$\binom{13}{1}\cdot\binom{4}{3}\frac{\binom{12}{1}\cdot 4\binom{11}{1}\cdot 4}{2!} = 54912\,,$$

giving a probability of 0.021128.

Part (e): We have $13\cdot\binom{4}{4}$ ways to pick the "a" denomination and twelve ways to pick the second card with a possible four faces, giving in total $13\cdot 12\cdot 4 = 624$ possible hands. This gives a probability of 0.00024.

Problem 16 (poker dice probabilities)

Part (a): We can select the five numbers that will show on the face of the 5 die in $\left(\begin{array}{c} 6 \\ 5 \end{array}\right) = 6$ ways. We then have 5! ways to order these five selected numbers. This gives for a probability

$$\frac{6 \cdot 5!}{6^5} = \frac{6!}{6^5} = 0.09259 \, .$$

Another way to compute this is using the results from parts (b)-(g) for this problem our probability of interest is

$$1 - P_b - P_c - P_d - P_e - P_f - P_g \, ,$$

where P_i is the probability computed during part "i" of this problem. Using the values provided in the problem we can evaluate the above to 0.0925.

Part (b): So solve this problem we will think of the die's outcome as being a numerical specifications (one through six) of five "slots". In this specification there are 6^5 total outcomes for a trial with the five dice. To determine the number of one pair "hands", we note that we can pick the number in the pair in six ways and their locations from the five bins in $\left(\begin{array}{c} 5 \\ 2 \end{array}\right)$ ways. Another number in the hand can be chosen from the five remaining numbers and placed in any of the remaining bins in $\left(\begin{array}{c} 3 \\ 1 \end{array}\right)$ ways. Continuing this line of reasoning for the values and placements of the remaining two dice, we have

$$6 \cdot \left(\begin{array}{c} 5 \\ 2 \end{array}\right) \cdot 5 \left(\begin{array}{c} 3 \\ 1 \end{array}\right) \cdot 4 \left(\begin{array}{c} 2 \\ 1 \end{array}\right) \cdot 3 \left(\begin{array}{c} 1 \\ 1 \end{array}\right) ,$$

as the number of *ordered* placements of our four distinct numbers. Since the ordered placement of the three different singleton numbers does not matter we must divide this result by 3!, which results in a value of 3600. Then the probability of one pair is given by

$$\frac{3600}{6^5} = 0.4629 \, .$$

Part (c): We specify the two numerical values to use in each of the two pairs in $\left(\begin{array}{c} 6 \\ 2 \end{array}\right)$ ways. Then the location of the first pair in $\left(\begin{array}{c} 5 \\ 2 \end{array}\right)$, the location of the second pair in $\left(\begin{array}{c} 3 \\ 2 \end{array}\right)$ ways, and finally the $\left(\begin{array}{c} 4 \\ 1 \end{array}\right)$ to select the third number. When we multiply these we get

$$\left(\begin{array}{c} 6 \\ 2 \end{array}\right) \left(\begin{array}{c} 5 \\ 2 \end{array}\right) \left(\begin{array}{c} 3 \\ 2 \end{array}\right) \left(\begin{array}{c} 4 \\ 1 \end{array}\right) = 1800 \, .$$

Combined this gives a probability of obtaining two pair of

$$\frac{1800}{6^5} = 0.2315 \, .$$

Part (d): We can pick the number for the digit that is repeated three times in six ways, another digit in five ways and the final digit in four ways. The number of ways we can place the three dice with the same numeric value is given by $\binom{5}{3}$ ways. So the number of permutations of these three numbers is given by

$$6 \cdot 5 \cdot 4 \cdot \binom{5}{3} = 1200\,.$$

This gives a probability of $\frac{1200}{6^5} = 0.154$.

Part (e): Recall that a full house is five dice, three and two of which have the same numeric value. We can choose the number shown on three die in 6 ways and their locations in the five rolls in $\binom{5}{3}$ ways. We then choose the number shown on the remaining two die in 5 ways. Thus the probability of a full houses is thus given by

$$\frac{6 \cdot 5 \cdot \binom{5}{3}}{6^5} = 0.0386\,.$$

Part (f): To get four dice with the same numeric value we must pick one special number out of six in $\binom{6}{1}$ ways representing the four common die. We then pick one more number from the remaining five in $\binom{5}{1}$ ways representing the number on the lone die. Thus we have $\binom{6}{1} \cdot \binom{5}{1}$ ways to pick the two numbers to use in the selection of this hand. We have $\binom{5}{1} = 5$ places in which we can place the lone die after which the location of the common four is determined. Using this the count of the number of arrangements is given by

$$\binom{6}{1} \cdot \binom{5}{1} \cdot 5 = 150\,.$$

This gives a requested probability of $\frac{150}{6^5} = 0.01929$.

Part (g): If all five dice are the same then there are one of six possibilities (the six numbers on a die). The total number of possible die throws is $6^5 = 7776$ giving a probability to throw this hand of

$$\frac{6}{6^5} = \frac{1}{6^4} = 0.0007716\,.$$

Problem 17 (randomly placing rooks)

A possible placement of a rook on the chess board can be obtained by specifying the row and column at which we will locate our rook. Since there are eight rows and eight columns there are $8^2 = 64$ possible placements for a given rook. After we place each rook we obviously have one less position where we can place the additional rooks. So the total number of possible locations where we can place eight rooks is given by

$$64 \cdot 63 \cdot 62 \cdot 61 \cdot 60 \cdot 59 \cdot 58 \cdot 57\,,$$

since the order of placement does not matter we must divide this number by 8! to get

$$\frac{64!}{8!(64-8)!} = \binom{64}{8} = 4426165368\,.$$

The number of locations where eight rooks can be placed who won't be able to capture any of the other is given by

$$8^2 \cdot 7^2 \cdot 6^2 \cdot 5^2 \cdot 4^2 \cdot 3^2 \cdot 2^2 \cdot 1^2\,,$$

Which can be reasoned as follows. The first rook can be placed in 64 different places. Once this rook is located we cannot place the next rook in the same row or column that the first rook holds. This leaves seven choices for a row and seven choices for a column giving a total of $7^2 = 49$ possible choices. Since the order of these choices does not matter we will need to divide this product by 8! giving a total probability of

$$\frac{\frac{8!^2}{8!}}{\binom{64}{8}} = 9.109\,10^{-6}\,,$$

in agreement with the book.

Problem 18 (randomly drawing blackjack)

The total number of possible two card hands is given by $\binom{52}{2}$. We can draw an ace in one of four possible ways i.e. in $\binom{4}{1}$ ways. For blackjack the other card must be a ten or a jack or a queen or a king (of any suite) and can be drawn in $\binom{4+4+4+4}{1} = \binom{16}{1}$ possible ways. Thus the number of possible ways to draw blackjack is given by

$$\frac{\binom{4}{1}\binom{16}{1}}{\binom{52}{2}} = 0.048265\,.$$

Problem 19 (symmetric dice)

We can solve this problem by considering the disjoint events that both dice land on colors given by red, black, yellow, or white. For the first die to land on red will happen with probability 2/6, the same for the second die. Thus the probability that both die land on red is given by

$$\left(\frac{2}{6}\right)^2 .$$

Summing up all the probabilities for all the possible colors, we have a total probability of obtaining the same color on both dice given by

$$\left(\frac{2}{6}\right)\left(\frac{2}{6}\right)+\left(\frac{2}{6}\right)\left(\frac{2}{6}\right)+\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)+\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)=\frac{5}{18} .$$

Problem 20 (blackjack against a dealer)

We assume that blackjack means the player gets an ace and a king, queen, a jack or a ten on the initial draw, and ignore the cases where the ace is used with a value of one and the player may draw another card. In that case, the probability that either the player or the dealer gets blackjack (independent of the other player) is just

$$\frac{\binom{4}{1}\binom{16}{1}}{\binom{52}{2}} = 0.048265 .$$

Let A and B be the events that player A or B gets blackjack. In the above we calculated $P(A)$ and $P(B)$. We want to calculate $P((A \cup B)^c) = 1 - P(A \cup B)$. This last event is

$$P(A \cup B) = P(A) + P(B) - P(AB) .$$

Thus we need to calculate $P(AB)$. This can be done as

$$P(AB) = P(B|A)P(A) = \frac{\binom{3}{1}\binom{15}{1}}{\binom{50}{2}} P(A) = 0.001773 .$$

We thus find that $P((A \cup B)^c) = 1 - (2(0.048265) - 0.00177) = 0.9052$.

Problem 21 (the number of children)

Part (a): Let P_i be the probability that the family chosen has i children. Then we see from the numbers provided that $P_1 = \frac{4}{20} = \frac{1}{5}$, $P_2 = \frac{8}{20} = \frac{2}{5}$, $P_3 = \frac{5}{20} = \frac{1}{4}$, and $P_4 = \frac{1}{20}$, assuming a uniform probability of selecting any given family.

Part (b): We have

$$4(1) + 8(2) + 5(3) + 2(4) + 1(5) = 4 + 16 + 15 + 8 + 5 = 48 ,$$

total children. Then the probability a random child comes from a family with i children is given by (and denoted by P_i) is $P_1 = \frac{4}{48}$, $P_2 = \frac{16}{48}$, $P_3 = \frac{15}{48}$, $P_4 = \frac{8}{48}$, and $P_5 = \frac{5}{48}$.

	1	2	3	4	5	6
1	0	1	1	1	1	1
2	0	0	1	1	1	1
3	0	0	0	1	1	1
4	0	0	0	0	1	1
5	0	0	0	0	0	1
6	0	0	0	0	0	0

Table 3: The elements of the sample space where the second die is strictly larger in value than the first.

Problem 22 (shuffling a deck of cards)

To have the ordering exactly the same we must have k heads in a row (which leave the first k cards unmoved) followed by $n-k$ tails in a row (which will move the cards $k+1, k+2, \ldots n$ to the end sequentially). We can do this for any $k=0$ to $k=n$. The probability of getting k heads followed by $n-k$ tails is

$$\left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \left(\frac{1}{2}\right)^n$$

Now since each of these outcomes is mutually exclusive to compute the total probability we can sum this result for $k=0$ to $k=n$ to get

$$\sum_{k=0}^{n} \left(\frac{1}{2}\right)^n = \frac{n+1}{2^n}.$$

Problem 23 (a larger roll than the first)

We begin by constructing the sample space of possible outcomes. These numbers are computed in table 3, where the row corresponds to the outcome of the first die through and the column corresponds to the outcome of the second die through. In each square we have placed a one if the number on the second die is strictly larger than the first. Since each element of our sample space has a probability of $1/36$, by enumeration we find that

$$\frac{15}{36} = \frac{5}{12},$$

is our desired probability.

Problem 24 (the probability the sum of the dice is i)

As in Problem 23 we can explicitly enumerate these probabilities by counting the number of times each occurrence happens, in Table 4 we have placed the sum of the two dice in the

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Table 4: The possible values for the sum of the values when two dice are rolled.

center of each square. Then by counting the number of squares where are sum equals each number from two to twelve, we have

$$\begin{aligned}
P_2 &= \frac{1}{36}, \quad P_7 = \frac{6}{36} = \frac{1}{6} \\
P_3 &= \frac{2}{36} = \frac{1}{18}, \quad P_8 = \frac{5}{36} \\
P_4 &= \frac{3}{36} = \frac{1}{12}, \quad P_9 = \frac{4}{36} = \frac{1}{9} \\
P_5 &= \frac{4}{36} = \frac{1}{9}, \quad P_{10} = \frac{3}{36} = \frac{1}{12} \\
P_6 &= \frac{5}{36}, \quad P_{11} = \frac{2}{36} = \frac{1}{18}, \quad P_{12} = \frac{1}{36}. \qquad (2)
\end{aligned}$$

Problem 25 (rolling a five before a seven)

A sum of five has a probability of $P_5 = \frac{2}{18} = \frac{1}{9}$ of occurring. A sum of seven has a probability of $P_7 = \frac{1}{6}$ of occurring, so the probability that neither a five or a seven is given by $1 - \frac{1}{9} - \frac{1}{6} = \frac{13}{18}$. Following the hint we let E_n be the event that a five occurs on the n-th roll and no five or seven occurs on the $n-1$-th rolls up to that point. Then

$$P(E_n) = \left(\frac{13}{18}\right)^{n-1} \frac{1}{9},$$

since we want the probability that a five comes first, this can happen at roll number one ($n = 1$), at roll number two ($n = 2$) or any subsequent roll. Thus the probability that a five comes first is given by

$$\begin{aligned}
\sum_{n=1}^{\infty} \left(\frac{13}{18}\right)^{n-1} \frac{1}{9} &= \frac{1}{9} \sum_{n=0}^{\infty} \left(\frac{13}{18}\right)^{n} \\
&= \frac{1}{9} \frac{1}{\left(1 - \frac{13}{18}\right)} = \frac{2}{5} = 0.4\,.
\end{aligned}$$

Problem 26 (winning at craps)

From Problem 24 we have computed the individual probabilities for various sum of two random dice. Following the hint, let E_i be the event that the initial dice sum to i and that the player wins. We can compute some of these probabilities immediately $P(E_2) = P(E_3) = P(E_{12}) = 0$, and $P(E_7) = P(E_{11}) = 1$. We now need to compute $P(E_i)$ for $i = 4, 5, 6, 8, 9, 10$. Again following the hint define $E_{i,n}$ to be the event that the player initial sum is i and wins on the n-th **subsequent** roll. Then

$$P(E_i) = \sum_{n=1}^{\infty} P(E_{i,n}) \,,$$

since if we win, it must be either on the first, or second, or third, etc roll *after the initial roll*. We now need to calculate the $P(E_{i,n})$ probabilities for each n. As an example of this calculation first lets compute $P(E_{4,n})$ which means that we initially roll a sum of four and the player wins on the n-th subsequent roll. We will win if we roll a sum of a four or loose if we roll a sum of a seven, while if roll anything else we continue, so to win when $n = 1$ we see that

$$P(E_{4,1}) = \frac{1+1+1}{36} = \frac{1}{12} \,,$$

since to get a sum of four we can roll pairs consisting of $(1, 3)$, $(2, 2)$, and $(3, 1)$.

To compute $P(E_{4,2})$ the rules of craps state that we will win if a sum of four comes up (with probability $\frac{1}{12}$) and loose if a sum of a seven comes up (with probability $\frac{6}{36} = \frac{1}{6}$) and continue playing if anything else is rolled. This last event (continued play) happens with probability

$$1 - \frac{1}{12} - \frac{1}{6} = \frac{3}{4} \,.$$

Thus $P(E_{4,2}) = \left(\frac{3}{4}\right)\frac{1}{12} = \frac{1}{16}$. Here the first $\frac{3}{4}$ is the probability we don't roll a four or a seven on the $n = 1$ roll and the second $\frac{1}{12}$ comes from rolling a sum of a four on the second roll (where $n = 2$). In the same way we have for $P(E_{4,3})$ the following

$$P(E_{4,3}) = \left(\frac{3}{4}\right)^2 \frac{1}{12} \,.$$

Here the first two factors of $\frac{3}{4}$ are from the two rolls that "keep us in the game", and the factor of $\frac{1}{12}$, is the roll that allows us to win. Continuing in this in this manner we see that

$$P(E_{4,4}) = \left(\frac{3}{4}\right)^3 \frac{1}{12} \,,$$

and in general we find that

$$P(E_{4,n}) = \left(\frac{3}{4}\right)^{n-1} \frac{1}{12} \quad \text{for} \quad n \geq 1 \,.$$

To compute $P(E_{i,n})$ for other i, the derivations just performed, only change in the probabilities required to roll the initial sum. We thus find that for other initial rolls (heavily using the results of Problem 24) that

$$
\begin{aligned}
P(E_{5,n}) &= \frac{1}{9}\left(1-\frac{1}{9}-\frac{1}{6}\right)^{n-1} = \frac{1}{9}\left(\frac{13}{18}\right)^{n-1} \\
P(E_{6,n}) &= \frac{5}{36}\left(1-\frac{5}{36}-\frac{1}{6}\right)^{n-1} = \frac{5}{36}\left(\frac{25}{36}\right)^{n-1} \\
P(E_{8,n}) &= \frac{5}{36}\left(1-\frac{5}{36}-\frac{1}{6}\right)^{n-1} = \frac{5}{36}\left(\frac{25}{36}\right)^{n-1} \\
P(E_{9,n}) &= \frac{1}{9}\left(1-\frac{1}{9}-\frac{1}{6}\right)^{n-1} = \frac{1}{9}\left(\frac{13}{18}\right)^{n-1} \\
P(E_{10,n}) &= \frac{1}{12}\left(1-\frac{1}{12}-\frac{1}{6}\right)^{n-1} = \frac{1}{12}\left(\frac{3}{4}\right)^{n-1}.
\end{aligned}
$$

To compute $P(E_4)$ we need to sum the results above. We have that

$$
\begin{aligned}
P(E_4) &= \frac{1}{12}\sum_{n\geq 1}\left(\frac{3}{4}\right)^{n-1} = \frac{1}{12}\sum_{n\geq 0}\left(\frac{3}{4}\right)^{n} \\
&= \frac{1}{12}\frac{1}{\left(1-\frac{3}{4}\right)} = \frac{1}{3}.
\end{aligned}
$$

Note that this also gives the probability for $P(E_{10})$. For $P(E_5)$ we find $P(E_5) = \frac{2}{5}$, which also equals $P(E_9)$. For $P(E_6)$ we find that $P(E_6) = \frac{5}{11}$, which also equals $P(E_8)$. Then our probability of winning craps is given by summing all of the above probabilities weighted by the associated priors of rolling the given initial roll. We find by defining I_i to be the event that the initial roll is i and W the event that we win at craps that

$$
\begin{aligned}
P(W) &= 0\,P(I_2) + 0\,P(I_3) + \frac{1}{3}\,P(I_4) + \frac{4}{9}\,P(I_5) + \frac{5}{9}\,P(I_6) \\
&+ 1\,P(I_7) + \frac{5}{9}\,P(I_8) + \frac{4}{9}\,P(I_9) + \frac{1}{3}\,P(I_{10}) + 1\,P(I_{11}) + 0\,P(I_{12}).
\end{aligned}
$$

Using the results of Exercise 25 to evaluate $P(I_i)$ for each i we find that the above summation gives

$$P(W) = \frac{244}{495} = 0.49292\,.$$

These calculations are performed in the Matlab file `chap_2_prob_26.m`.

Problem 27 (drawing the first red ball)

We want the probability that A selects the first red ball. Since A draws first he will select a red ball on the first draw with probability $\frac{3}{10}$. If he does not select a red ball B will draw next and he must not draw a red ball (or the game will stop). The probability that A draws a red ball on the *third* total draw is then

$$P_3 = \left(1-\frac{3}{10}\right)\left(1-\frac{3}{9}\right)\left(\frac{3}{8}\right).$$

Continuing this pattern we see that for A to draw a ball on the *fifth* total draw will happen with probability

$$P_5 = \left(1 - \frac{3}{10}\right)\left(1 - \frac{3}{9}\right)\left(1 - \frac{3}{8}\right)\left(1 - \frac{3}{7}\right)\left(\frac{3}{6}\right),$$

and finally on the *seventh* total draw with probability

$$P_7 = \left(1 - \frac{3}{10}\right)\left(1 - \frac{3}{9}\right)\left(1 - \frac{3}{8}\right)\left(1 - \frac{3}{7}\right)\left(1 - \frac{3}{6}\right)\left(1 - \frac{3}{5}\right)\left(\frac{3}{4}\right).$$

If player A does not get a red ball after seven draws he will not draw a red ball before player B. The total probability that player A draws a red ball first is given by the sum of all these individual probabilities of these mutually exclusive events. In the Matlab code `chap_2_prob_27.m` we evaluate this sum and find the probability that A wins given by

$$P(A) = \frac{7}{12}.$$

So the corresponding probability that B wins is $1 - \frac{7}{12} = \frac{5}{12}$ showing the benefit to being the first "player" in a game like this.

Problem 28 (sampling colored balls from an urn)

Part (a): We want the probability that each ball will be of the same color. This is given by

$$\frac{\binom{5}{3} + \binom{6}{3} + \binom{8}{3}}{\binom{5+6+8}{3}} = 0.08875.$$

Part (b): The probability that all three balls are of different colors is given by

$$\frac{\binom{5}{1}\binom{6}{1}\binom{8}{1}}{\binom{19}{3}} = 0.247.$$

If we replace the ball after drawing it, then the probabilities that each ball is the same color is now given by

$$\left(\frac{5}{19}\right)^3 + \left(\frac{6}{19}\right)^3 + \left(\frac{8}{19}\right)^3 = 0.124.$$

while if we want three balls of different colors, then this happens with probability given by

$$3!\left(\frac{5}{19}\right)\left(\frac{6}{19}\right)\left(\frac{8}{19}\right) = 0.2099.$$

Problem 29

Warning: Here are some notes I had on this problem. I've not had the time to check these in as much detail as I would have liked. Caveat emptor.

Part (a): The probability we obtain two while balls is given by

$$\frac{n}{n+m}\left(\frac{n-1}{m+n-1}\right) .$$

The probability that we obtain two black balls is given by

$$\frac{m}{m+n}\left(\frac{m-1}{m+n-1}\right) ,$$

so the probability of two balls of the same color then is

$$\frac{n(n-1)}{(m+n)(m+n-1)} + \frac{m(m-1)}{(m+n)(m+n-1)} = \frac{n(n-1)+m(m-1)}{(m+n)(m+n-1)} ,$$

Part (b): Now we replace the balls after we draw them so the probability we draw two white balls is then

$$\frac{n}{m+n}\left(\frac{n}{m+n}\right) ,$$

and for black balls we have

$$\frac{m}{m+n}\left(\frac{m}{m+n}\right) ,$$

So in total then we have

$$\frac{n^2+m^2}{(m+n)^2} = \frac{n^2+m^2}{m^2+2mn+n^2} .$$

Part (c): We expect to have a better chance of getting two balls of the same color in Part (b) of this problem since we have an additional white or black ball in the pot to draw on the second draw. Thus we want to show that

$$\frac{n^2+m^2}{(m+n)^2} \geq \frac{n(n-1)+m(m-1)}{(m+n)(m+n-1)} .$$

We will perform reversible manipulations to derive an equivalent expression. If the reduced expression is true, then the original expression is true. We begin by canceling the factor $\frac{1}{m+n}$ to give

$$\frac{n^2+m^2}{m+n} \geq \frac{n(n-1)+m(m-1)}{m+n-1} .$$

multiplying by the common denominator we obtain the following sequence of transformations

$$\begin{aligned}
(m^2+n^2)(m+n-1) &\geq (m+n)(n(n-1)+m(m-1)) \\
m^3+m^2n-m^2+n^2m+n^3-n^2 &\geq n(nm-m+n^2-n)+m(m^2-m+nm-n) \\
m^3+m^2n-m^2+n^2m+n^3-n^2 &\geq mn^2-mn+n^3-n^2+m^3-m^2+nm^2-nm \\
n^2m &\geq -mn+nm^2-nm ,
\end{aligned}$$

by dividing by mn we get $n \geq -1+n-1$ or the inequality $0 \geq -2$ which is true showing that the original inequality is true.

Problem 30 (the chess club)

Part (a): For Rebecca and Elise to be paired they must first be selected onto their respected schools chess teams and then be paired in the tournament. Thus if S is the event that the sisters play each other then

$$P(S) = P(R)P(E)P(\text{Paired}|R, E)\,,$$

where R is the event that that Rebecca is selected for her schools chess team and E is the event that Elise is selected for her schools team and Paired is the event that the two sisters play each other. Computing these probabilities we have

$$P(R) = \frac{\binom{1}{1}\binom{7}{3}}{\binom{8}{4}} = \frac{1}{2}\,,$$

and

$$P(E) = \frac{\binom{1}{1}\binom{8}{3}}{\binom{9}{4}} = \frac{4}{9}\,,$$

and finally

$$P(\text{Paired}) = \frac{1 \cdot 3!}{4!} = \frac{1}{4}\,.$$

so that $P(S) = \frac{1}{2} \cdot \frac{4}{9} \cdot \frac{1}{4} = \frac{1}{18}$.

Part (b): The event that Rebecca and Elise are chosen and then do not play each other will occur with a probability of

$$P(R)P(E)P(\text{Paired}^c|R, E) = \frac{1}{2} \cdot \frac{4}{9}\left(1 - \frac{1}{4}\right) = \frac{1}{6}\,.$$

Part (c): For this part we can have either (and these events are mutually exclusive) Rebecca picked to represent her school or Elise picked to represent her school but not both and not neither. Since $\binom{1}{1}\binom{7}{3}$ is the number of ways to choose the team A with Rebecca as a member and $\binom{8}{4}$ are the number of ways to choose team B without having Elise as a member, their product is the number of ways of choosing the first option above. This given a probability of

$$\frac{\binom{1}{1}\binom{7}{3}}{\binom{8}{4}} \cdot \frac{\binom{8}{4}}{\binom{9}{4}} = \frac{5}{18}\,.$$

In the same way the other probability is given by

$$\frac{\binom{7}{4}}{\binom{8}{4}} \cdot \frac{\binom{1}{1}\binom{8}{3}}{\binom{9}{4}} = \frac{2}{9}.$$

Thus the probability we are after is the sum of the two probabilities above and is given by $\frac{9}{18} = \frac{1}{2}$.

Problem 31 (selecting basketball teams)

Part (a): On the first draw we will certainly get one of the team members. Then on the second draw we must get any team member *but* the one that we just drew. This happens with probability $\frac{2}{3}$. Finally, we must get the team member we have not drawn in the first two draws. This happens with probability $\frac{1}{3}$. In total then, the probability to draw an entire team is given by

$$1 \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}.$$

Part (b): The probability the second player plays the same position as the first drawn player is given by $\frac{1}{3}$, while the probability that the third player plays the same position as the first two is given by $\frac{1}{3}$. Thus this event has a probability of

$$\frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}.$$

Problem 32 (a girl in the i-th position)

We can compute all permutations of the $b+g$ people that have a girl in the i-th spot as follows. We have g choices for the specific girl we place in the i-th spot. Once this girl is selected we have $b+g-1$ other people to place in the $b+g-1$ slots around this i-th spot. This can be done in $(b+g-1)!$ ways. So the total number of ways to place a girl at position i is $g(b+g-1)!$. Thus the probability of finding a girl in the i-th spot is given by

$$\frac{g(b+g-1)!}{(b+g)!} = \frac{g}{b+g}.$$

Problem 33 (a forest of elk)

After tagging the initial elk we have 5 tagged elk from 20. When we capture four more elk the probability we get two tagged elk is the number of ways we can select two tagged elks (from 5) and two untagged elks (from $20-5=15$) divided by the number of ways to select four elk from 20. This probability p is given by

$$p = \frac{\binom{5}{2}\binom{15}{2}}{\binom{20}{4}} = \frac{70}{323}\,.$$

Problem 34 (the probability of a Yarborough)

We must not have a ten, a jack, a queen, a king, or an ace (a total of 5 face cards) in our hand of thirteen cards. The number of ways to select a hand that does not have any of these cards is equivalent to selecting thirteen cards from among a set that does not contain any of the cards mentioned above. Specifically this number is

$$\frac{\binom{52-4-4-4-4-4}{13}}{\binom{52}{13}} = \frac{\binom{32}{13}}{\binom{52}{13}} = 0.000547\,,$$

a relatively small probability.

Problem 35 (selecting psychiatrists for a conference)

The probability that at least one psychologist is choose is given by considering all selections of sets of psychologists that contain at least one

$$\frac{\binom{30}{2}\binom{24}{1}+\binom{30}{1}\binom{24}{2}+\binom{30}{0}\binom{24}{3}}{\binom{54}{3}} = 0.8363\,.$$

Where in the numerator we have enumerated all possible selections of three people such that at least one psychologist is chosen.

Problem 36 (choosing two identical cards)

Part (a): We have $\binom{52}{2}$ possible ways to draw two cards from the 52 total. For us to draw two aces, this can be done in $\binom{4}{2}$ ways. Thus our probability is given by

$$\frac{\binom{4}{2}}{\binom{52}{2}} = 0.00452\,.$$

Part (b): For the two cards to have the same value we can pick the value to represent in thirteen ways and the two cards in $\binom{4}{2}$ ways. Thus our probability is given by

$$\frac{13\binom{4}{2}}{\binom{52}{2}} = 0.0588\,.$$

Problem 37 (solving enough problems on an exam)

Part (a): In this part of the problem imagine that we label the 10 questions as "known" or "unknown". Since the student knows how to solve 7 of the 10 problems, we have 7 known problems and 3 unknown questions. If we imagine the teacher selecting the 5 exam questions randomly then the probability that the student answers all 5 selected problems correctly is the probability that we draw 5 known questions from a "set" of 7 known and 3 unknown questions. This later probability is given by

$$\frac{\binom{7}{5}\binom{3}{0}}{\binom{10}{5}} = \frac{1}{12} = 0.083333\,.$$

Part (b): To answer at least four of the questions correctly will happen if the student answers 5 questions correctly (with probability given above) or 4 questions correctly. In the same way as above this later probability is given by

$$\frac{\binom{7}{4}\binom{3}{1}}{\binom{10}{5}} = \frac{5}{12}\,.$$

Thus the probability that the student answers at least four of the problems correctly is the sum of these two probabilities or

$$\frac{5}{12} + \frac{1}{12} = \frac{1}{2}\,.$$

Problem 38 (two red socks)

We are told that three of the socks are red so that $n - 3$ are not red. When we select two socks, the probability that they are both red is given by

$$\frac{3}{n} \cdot \frac{2}{n-1}\,.$$

If we want this to be equal to $\frac{1}{2}$ we must solve for n in the following expression

$$\frac{3}{n} \cdot \frac{2}{n-1} = \frac{1}{2} \quad \Rightarrow \quad n^2 - n = 12\,.$$

Using the quadratic formula this has a solution given by

$$n = \frac{1 \pm \sqrt{1 + 4(1)(12)}}{2(1)} = \frac{1 \pm 7}{2}\,.$$

Taking the positive solution we have that $n = 4$.

Problem 39 (five different hotels)

When the first person checks into the hotel, the next person will check into a different hotel with probability $\frac{4}{5}$. The next person will check into a different hotel with probability $\frac{3}{5}$. Thus the probability that we check into three different hotels is given by

$$\frac{4}{5} \cdot \frac{3}{5} = \frac{12}{25} = 0.48\,.$$

Problem 41 (obtaining a six at least once)

This is the complement of the probability that a six never appears or

$$1 - \left(\frac{5}{6}\right)^4 = 0.5177\,.$$

Problem 42 (double sixes)

The probability that a double six appear at least once is the complement of the probability that a double six never appears. The probability of not seeing a double six is given by $1 - \frac{1}{36} = \frac{35}{36}$, so the probability that a double six appears at least once in n throws is given by

$$1 - \left(\frac{35}{36}\right)^n .$$

To make this probability at least $1/2$ we need to have

$$1 - \left(\frac{35}{36}\right)^n \geq \frac{1}{2} .$$

which gives when we solve for n

$$n \geq \frac{\ln(\frac{1}{2})}{\ln(\frac{35}{36})} \approx 24.6 \, ,$$

so we should take $n = 25$.

Problem 43 (the probability you are next to me)

Part (a): The number of ways to arrange N people is $N!$. To count the number of permutation of the other people and the "pair" A and B consider A and B as fused together as one unit (say AB) to be taken with the other $N - 2$ people. So in total we have $N - 2 + 1$ things to order. This can be done in $(N - 1)!$ ways. Note that for every permutation we also have two orderings of A and B i.e. AB and BA so we have $2(N - 1)!$ orderings where A and B are fused together. The the probability we have A and B fused together is given by $\frac{2(N-1)!}{N!} = \frac{2}{N}$.

Part (b): If the people are arraigned in a circle there are $(N-1)!$ unique arrangements of the total people. The number of arrangement as in part (a) is given by $2(N-2+1-1)! = 2(N-2)!$ so our probability is given by

$$\frac{2(N-2)!}{(N-1)!} = \frac{2}{N-1} .$$

Problem 44 (people between A and B)

Note that we have $5!$ orderings of the five individual people.

Part (a): The number of permutations that have one person between A and B can be determined as follows. First pick the person to put between A and B from our three choices C, D, and E. Then pick the ordering of A and B i.e AB or BA. Then considering this AB object as *one* object we have to place it with two other people in $3!$ ways. Thus the number of orderings with one person between A and B is given by $3 \cdot 2 \cdot 3!$, giving a probability of this event of

$$\frac{3 \cdot 2 \cdot 3!}{5!} = 0.3 \,.$$

Part (b): Following Part (a) we can pick the two people from the three remaining in $\binom{3}{2} = 3$ (ignoring order) ways. Since the people can be ordered in two different ways and A and B on the outside can be ordered in two different ways, we have $3 \cdot 2 \cdot 2 = 12$ ways to create the four person "object" with A and B on the outside. This can ordered with the remaining single person in two ways. Thus our probability is given by

$$\frac{2 \cdot 12}{5!} = \frac{1}{5} \,.$$

Part (c): To have three people between A and B, A and B must be on the ends with $3! = 6$ possible ordering of the remaining people. Thus with two orderings of A and B we have a probability of

$$\frac{2 \cdot 6}{5!} = \frac{1}{10} \,.$$

Problem 45 (trying keys at random)

Part (a): If unsuccessful keys are removed as we try them, then the probability that the k-th attempt opens the door can be computed by recognizing that all attempts up to (but not including) the k-th have resulted in failures. Specifically, if we let N be the random variable denoting the attempt that opens the door we see that

$$\begin{aligned}
P\{N = 1\} &= \frac{1}{n} \\
P\{N = 2\} &= \left(1 - \frac{1}{n}\right) \frac{1}{n-1} \\
P\{N = 3\} &= \left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{n-1}\right) \frac{1}{n-2} \\
&\vdots \\
P\{N = k\} &= \left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{n-1}\right) \cdots \left(1 - \frac{1}{n-(k-2)}\right) \frac{1}{n-(k-1)} \,.
\end{aligned}$$

We can check that this result is a valid expression to represent a probability by selecting a value for n and verifying that when we sum the above over k for $1 \le k \le n$ we sum to one. A verification of this can be found in the Matlab file `chap_2_prob_45.m`, along with explicit calculations of the mean and variance of N. A *much* simpler expression, making the above Matlab script rather silly, is obtained if we simplify the above expressions by multiplying all factors together. When we do that we see that we obtain

$$P\{N = k\} = \frac{1}{n} \,.$$

Part (b): If unsuccessful keys are not removed then the probability that the correct key is selected at draw k is a geometric random with parameter $p = 1/n$. Thus our probabilities are given by $P\{N = k\} = (1-p)^{k-1}p$, and we have for a geometric random variable an expectation and a variance given by

$$\begin{aligned} E[N] &= \frac{1}{p} = n \\ \text{Var}(N) &= \frac{1-p}{p^2} = n(n-1)\,. \end{aligned}$$

The above expression of $(1-p)^{k-1}p$ represents the probability that we fail to find the correct key (with probability $1-p$) in the trials $1, 2, \cdots, k-1$ and then on the k trial find the correct key (with probability p).

Problem 46 (the birthdays of people in a room)

The probability that at least two people share the same birthday is the complement of the probability that no two people have the same birthday (or that all people have distinct birthdays). Let E_n the the event that n people have at least one birthday in common. Based on the above if $n = 2$ then the probability that at least two people share the same birthday is

$$P(E_2) = 1 - \frac{11}{12}\,.$$

That three people share the same birthday is

$$P(E_3) = 1 - \frac{11}{12}\frac{10}{11} = 1 - \frac{10}{12}\,.$$

That 4 people share the same birthday is

$$P(E_4) = 1 - \frac{11}{12}\frac{10}{11}\frac{9}{10} = 1 - \frac{9}{12}\,.$$

It seems the pattern for general n is

$$P(E_n) = 1 - \frac{12-(n-1)}{12} = 1 - \frac{13-n}{12}\,.$$

We want to pick n such that $P(E_n) \geq \frac{1}{2}$. From the above when we solve for n this means that $n \geq 5$.

Problem 47 (strangers in a room)

There are 12^{12} possible ways to distribute birthdays to all people. We next want to count the number of ways to distribute birthdays so that no two are the same. This can be done in 12! ways. Thus we get a probability of

$$P = \frac{12!}{12^{12}} = \frac{11!}{12^{11}}\,.$$

Problem 48 (certain birthdays)

As each person can have his birthday assigned to one of the twelve calender months we have $T = 12^{20}$ possible ways to assign birthday months to people. This will be the denominator in the probability we seek. We now need to compute the number of ways we can get the desired distribution of months and people requested in the problem. We can select the four months that are to have two birthdays in $\left(\begin{array}{c}12\\4\end{array}\right)$ ways and after this the four months are to have three birthdays in $\left(\begin{array}{c}8\\4\end{array}\right)$ ways. Thus the number of selections of months M can be done in

$$M = \left(\begin{array}{c}12\\4\end{array}\right)\left(\begin{array}{c}8\\4\end{array}\right) = \frac{12!}{8!4!}\cdot\frac{8!}{4!4!} = \frac{12!}{4!^3} = 34650\,,$$

ways. Once the months are specified we need to select the people that will have their birthdays in these selected months. Since we need to put two men in the first selected four months and then three men in the second selected four months we can do that in N ways where N is given by

$$\begin{aligned}
N &= \left(\begin{array}{c}20\\2\end{array}\right)\left(\begin{array}{c}18\\2\end{array}\right)\left(\begin{array}{c}16\\2\end{array}\right)\left(\begin{array}{c}14\\2\end{array}\right)\times\left(\begin{array}{c}12\\3\end{array}\right)\left(\begin{array}{c}9\\3\end{array}\right)\left(\begin{array}{c}6\\3\end{array}\right)\left(\begin{array}{c}3\\3\end{array}\right)\\
&= \frac{20!}{18!2!}\cdot\frac{18!}{16!2!}\cdot\frac{16!}{14!2!}\cdot\frac{14!}{12!2!}\times\frac{12!}{9!3!}\cdot\frac{9!}{6!3!}\cdot\frac{6!}{3!3!}\cdot\frac{3!}{3!0!}\\
&= \frac{20!}{2!^4 3!^4} = 1.173274\,10^{14}\,.
\end{aligned}$$

Using these results we get a probability of

$$P = \frac{NM}{T} \approx 0.0010604\,,$$

the same as in the back of the book.

Problem 49 (men and women)

The only way to have equal numbers of men in each group is to have three men in each group (and thus three women in each group). We have $\left(\begin{array}{c}6\\3\end{array}\right)$ ways to select the men (and the same number of ways to select the women). The probability is then given by

$$P = \frac{\left(\begin{array}{c}6\\3\end{array}\right)\left(\begin{array}{c}6\\3\end{array}\right)}{\left(\begin{array}{c}12\\6\end{array}\right)} = \frac{20^2}{924} = 0.4329\,.$$

Problem 50 (hands of bridge with with spades)

We have $\binom{52}{13}$ ways to draw the first hand. If we want to have 5 spades we can select these in $\binom{13}{5}$ ways and then the additional cards for this hand in $\binom{52-13}{13-5} = \binom{39}{8}$ ways. The hand with five spades is then drawn with a probability

$$\frac{\binom{13}{5}\binom{39}{8}}{\binom{52}{13}} = 0.12469\,.$$

After this hand is drawn we need do draw the second hand. We want this hand to have the remaining 8 spades and can be drawn with probability

$$\frac{\binom{8}{8}\binom{31}{5}}{\binom{39}{13}} = 2.0918\,10^{-5}\,.$$

Then the probability that both of these events happen simultaneously is then given by their product or $2.6084\,10^{-6}$.

Problem 51 (n balls in N compartments)

If we put m balls in the first component we have to place the remaining $n-m$ balls in $N-1$ compartments. This can be done in $(N-1)^{n-m}$ ways. We can select our m balls to place in the first compartment in $\binom{n}{m}$ ways. Combining these two gives a probability of

$$\frac{\binom{n}{m}(N-1)^{n-m}}{N^n}\,.$$

Another way to view this problem is to consider the event that a given one of our n balls land in the first (or any specific compartment) as a success that happens with probability $p = \frac{1}{N}$. Then the probability we have m success from our n trials is a binomial random variable giving

$$P(M=m) = \binom{n}{m}\left(\frac{1}{N}\right)^m\left(1-\frac{1}{N}\right)^{n-m},$$

the same as earlier.

Problem 52 (a closet with shoes)

Part (a): We have $\left(\begin{array}{c} 20 \\ 8 \end{array}\right)$ ways of selecting our eight shoes. Since we don't want any matching pairs for this part we can select 8 pairs from the 10 pairs available in $\left(\begin{array}{c} 10 \\ 8 \end{array}\right)$ ways. In each pair we can select either the right or the left shoe. This gives

$$P = \frac{\left(\begin{array}{c} 10 \\ 8 \end{array}\right) 2^8}{\left(\begin{array}{c} 20 \\ 8 \end{array}\right)} = 0.09 \,.$$

Part (b): We select one pair in include in $\left(\begin{array}{c} 10 \\ 1 \end{array}\right)$ ways. Then the pairs the other shoes will come from in $\left(\begin{array}{c} 9 \\ 6 \end{array}\right)$ ways and the left-right shoe in 2^6 ways giving

$$P = \frac{\left(\begin{array}{c} 10 \\ 1 \end{array}\right)\left(\begin{array}{c} 9 \\ 6 \end{array}\right) 2^6}{\left(\begin{array}{c} 20 \\ 8 \end{array}\right)} = 0.4267 \,.$$

Problem 53 (four married couples in a row)

This problem is very much like the Example 5n from the book. We let E_i be the event that couple i sits next to each other. The event that at least one couple sits next to each other is $\cup_i E_i$. The probability that no couple sits next to each other is then $1 - P(\cup_i E_i)$. To evaluate $P(\cup_i E_i)$ we will use the inclusion-exclusion lemma which for our four couples is given by

$$P(\cup_{i=1}^4 E_i) = \sum_{i=1}^{4} P(E_i) - \sum_{i<j} P(E_i E_j) + \sum_{i<j<k} P(E_i E_j E_k) - \sum_{i<j<k<l} P(E_i E_j E_k E_l) \,. \quad (3)$$

We now need to compute each of these joint probabilities. To do that first consider $P(E_i)$. First given the 8 total people there are 8! ways of arranging all the people in a row. We want to count the number of these that have couple i sitting next to each other. If we consider this couple "fused" together there are then 7 objects that can be placed in a line such that the couple is sitting together (the 6 other people and the one couple object). This gives 7! ways of arranging this 7 objects. We have then two ways to permute the husband and wife in the couple giving

$$P(A_i) = \frac{2 \cdot 7!}{8!} \,.$$

Next consider the evaluation of $P(A_i A_j)$. We again have 8! for the denominator of this probability. To compute the numerator we again imagine fuse two couples together giving

$8 - 4 + 2 = 6$ objects to place. This can be done in $6!$ way. We can permute the husband and wife in each pair in $2 \cdot 2 = 2^2$ ways. Thus we find

$$P(A_i A_j) = \frac{2^2 \cdot 6!}{8!}.$$

In general, following the same logic we have for r couples

$$P(A_i A_j \cdots A_k) = \frac{2^r (8-r)!}{8!}.$$

Now by symmetry all of the probabilities in the individual sums are the same and that there are $\binom{4}{r}$ for $r = 1, 2, 3, 4$ terms respectively in each of the sums in Equation 3 above. Thus using what we have so far the probability that at least one couple sits together is given by

$$\begin{aligned} P(\cup_i E_i) &= \binom{4}{1} \frac{2 \cdot 7!}{8!} - \binom{4}{2} \frac{2^2 \cdot 6!}{8!} + \binom{4}{3} \frac{2^3 \cdot 5!}{8!} - \binom{4}{4} \frac{2^4 4!}{8!} \\ &= 1 - \frac{12}{35}, \end{aligned}$$

Thus the probability we seek is given by $1 - (1 - \frac{12}{35}) = \frac{12}{35}$.

Problem 54 (a bridge hand that is void in at least one suit)

We want the probability that a given hand of bridge is void in *at least one* suit which means the hand could be void in more than one suit. The error in the suggested calculation probability given is that it gives the probability that the hand is void in one (and only one suit), thus it underestimates the probability of interest. Let E_i be the event that the hand is void in the suit i for $i = 1, 2, 3, 4$. Then the probability we want is $P(\cup_{i=1}^4 E_i)$ which we can calculate by using the inclusion-exclusion identity given in this case by

$$P(\cup_{i=1}^4 E_i) = \sum_{i=1}^4 P(E_i) - \sum_{i<j} P(E_i E_j) + \sum_{i<j<k} P(E_i E_j E_k). \tag{4}$$

To do this we need to be able to evaluate the joint probabilities $P(E_i)$, $P(E_i E_j)$, and $P(E_i E_j E_k)$ for $i = 1, 2, 3, 4$. Note there is no terms $P(E_i E_j E_k E_l)$ since we must be dealt some cards. We start with $P(E_i)$ where we fix the value of i and

$$P(E_i) = \frac{\binom{39}{13}}{\binom{52}{13}} = 0.01279.$$

Next we have

$$P(E_i E_j) = \frac{\binom{26}{13}}{\binom{52}{13}} = 1.63785\, 10^{-5}.$$

and finally

$$P(E_iE_jE_k) = \frac{\binom{13}{13}}{\binom{52}{13}} = 1.57476\,10^{-12}\,.$$

Now by symmetry all of the probabilities in the individual sums are the same and that there are $\binom{4}{r}$ for $r = 1, 2, 3$ terms respectively in each of the sums in Equation 4 above. Thus we get

$$P(\cup_{i=1}^4 E_i) = \binom{4}{1}(0.01279) - \binom{4}{2}(1.63785\,10^{-5}) + \binom{4}{3}(1.57476\,10^{-12}) = 0.0510655208\,.$$

Problem 55 (hands of cards)

Part (a): We want the probability that a given hand of bridge has the ace and king in *at least one* suit. Let E_i be the event that the hand has an ace and a king in the suit i for $i = 1, 2, 3, 4$. Then the probability we want is $P(\cup_{i=1}^4 E_i)$ which we can calculate by using the inclusion-exclusion identity given in this case by

$$P(\cup_{i=1}^4 E_i) = \sum_{i=1}^4 P(E_i) - \sum_{i<j} P(E_iE_j) + \sum_{i<j<k} P(E_iE_jE_k) - \sum_{i<j<k<l} P(E_iE_jE_kE_l)\,. \quad (5)$$

To do this we need to be able to evaluate the joint probabilities $P(E_i)$, $P(E_iE_j)$, $P(E_iE_jE_k)$, and $P(E_iE_jE_kE_l)$ for i, j, k, and l for $1, 2, 3, 4$. We start with $P(E_i)$ where we fix the value of i and

$$P(E_i) = \frac{\binom{50}{11}}{\binom{52}{13}} = 0.0588235\,.$$

Next we have

$$P(E_iE_j) = \frac{\binom{48}{9}}{\binom{52}{13}} = 0.002641\,.$$

and

$$P(E_iE_jE_k) = \frac{\binom{46}{7}}{\binom{52}{13}} = 8.4289\,10^{-5}\,,$$

and finally

$$P(E_iE_jE_kE_l) = \frac{\binom{44}{5}}{\binom{52}{13}} = 1.7102\,10^{-6}\,,$$

A wins (Y/N)	spinner a outcome	spinner b outcome
Y	9	7
Y	9	6
Y	9	2
N	5	7
N	5	6
Y	5	2
N	1	7
N	1	6
N	1	2

Table 5: The possible values spinners a and b in the game for Problem 2-56.

Now by symmetry all of the probabilities in the individual sums are the same and that there are $\binom{4}{r}$ for $r = 1, 2, 3$ terms respectively in each of the sums in Equation 5 above. Thus we get

$$\begin{aligned} P(\cup_{i=1}^{4} E_i) &= \binom{4}{1}(0.0588235) - \binom{4}{2}(0.002641) + \binom{4}{3}(8.4289\,10^{-5}) - \binom{4}{4}(1.7102\,10^{-6}) \\ &= 0.0808910\,. \end{aligned}$$

Part (b): In the same way as before we let E_i be the event that the hand is missing all four suits from the ith denominator $1 \leq i \leq 13$. We then when we fix the indices i, j, k, and l

$$P(E_i) = \frac{\binom{48}{9}}{\binom{52}{13}} = 0.00264$$

$$P(E_iE_j) = \frac{\binom{52-2(4)}{13-2(4)}}{\binom{52}{13}} = 1.71021\,10^{-6}$$

$$P(E_iE_jE_k) = \frac{\binom{52-3(4)}{13-3(4)}}{\binom{52}{13}} = 6.29907\,10^{-11}\,,$$

Now by symmetry all of the probabilities in the individual sums are the same and that there are $\binom{13}{r}$ for $r = 1, 2, 3 \ldots 13$ terms respectively in each of the sums in the inclusion-exclusion identity. Thus we get

$$P(\cup_{i=1}^{13} E_i) = \binom{13}{1}(0.00264) - \binom{13}{2}(1.71021\,10^{-6}) + \binom{13}{3}(6.29907\,10^{-11}) = 0.034200\,.$$

	B then picks a	B then picks b	B then picks c
A first picks a		5/9	4/9
A first picks b	4/9		5/9
A first picks c	5/9	4/9	

Table 6: The probabilities that A will win the game proposed in Problem 2-56, when A selects first and takes the spinner indicated by the row and B selects his spinner second and takes the spinner indicated by the given column.

Problem 56 (a game with spinners)

Since the player A goes first lets calculate the probability he wins when he picks spinners a, b, and c and player B picks from the other two choices. Let $A = a$ be the event that player A picks spinner a, $A = b$ the event that A picks spinner b and all other notation should be interpreted in the same way. Then if $A = a$ and $B = b$ say we have the possible outcome from a set of spins given by Table 5. From this table we see that the probability that A wins in this case is given by $\frac{4}{9}$. Other probabilities can be computed from this one. For example $P(A\,\text{wins}|A = b, B = a) = 1 - P(A\,\text{wins}|A = a, B = b) = 1 - \frac{4}{9} = \frac{5}{9}$. When we compute the probability that A wins under each of his initial choices and then under each of B subsequent choices we get Table 6. In that table we see that no matter what spinner player A selects player B can select a spinner such that he has a higher probability of winning. For example, if player A selects b then player B should pick spinner a since that gives him a $1 - \frac{4}{9} = \frac{5}{9} > \frac{1}{2}$ chance of winning. Thus we see that it is better to be player B and have the second choice.

Chapter 2: Theoretical Exercises

Problem 1 (set identities)

To prove this let $x \in E \cap F$ then by definition $x \in E$ and therefore $x \in E \cup F$. Thus $E \cap F \subset E \cup F$.

Problem 2 (more set identities)

If $E \subset F$ then $x \in E$ implies that $x \in F$. If $y \in F^c$, then this implies that $y \notin F$ which implies that $y \in E$, for if y was in E then it would have to be in F which we know it is not.

Problem 3 (more set identities)

We want to prove that $F = (F \cap E) \cup (F \cap E^c)$. We will do this using the standard proof where we show that each set in the above is a subset of the other. We begin with $x \in F$. Then if $x \in E$, x will certainly be in $F \cap E$, while if $x \notin E$ then x will be in $F \cap E^c$. Thus in either case ($x \in E$ or $x \notin E$) x will be in the set $(F \cap E) \cup (F \cap E^c)$.

If $x \in (F \cap E) \cup (F \cap E^c)$ then x is in either $F \cap E$, $F \cap E^c$, or both by the definition of the union operation. Now x cannot be in both sets or else it would simultaneously be in E and E^c, so x must be in one of the two sets only. Being in either set means that $x \in F$ and we have that the set $(F \cap E) \cup (F \cap E^c)$ is a subset of F. Since each side is a subset of the other we have shown set equality.

To prove that $E \cup F = E \cup (E^c \cap F)$, we will begin by letting $x \in E \cup F$, thus x is an element of E or an element of F or of both. If x is in E at all then it is in the set $E \cup (E^c \cap F)$. If $x \notin E$ then it must be in F to be in $E \cup F$ and it will therefore be in $E^c \cap F$. Again both sides are subsets of the other and we have shown set equality.

Problem 6 (set expressions for various events)

Part (a): This would be given by the set $E \cap F^c \cap G^c$.

Part (b): This would be given by the set $E \cap G \cap F^c$.

Part (c): This would be given by the set $E \cup F \cup G$.

Part (d): This would be given by the set

$$((E \cap F) \cap G^c) \cup ((E \cap G) \cap F^c) \cup ((F \cap G) \cap E^c) \cup (E \cap F \cap G) .$$

This expresses the fact that satisfy this criterion by being inside two other events or by being inside three events.

Part (e): This would be given by the set $E \cap F \cap G$.

Part (f): This would be given by the set $(E \cup F \cup G)^c$.

Part (g): This would be given by the set

$$(E \cap F^c \cap G^c) \cup (E^c \cap F \cap G^c) \cup (E^c \cap F^c \cap G)$$

Part (h): At most two occur is the complement of all three taking place, so this would be given by the set $(E \cap F \cap G)^c$. Note that this includes the possibility that none of the events happen.

Part (i): This is a subset of the sets in Part (d) (i.e. without the set $E \cap F \cap G$) and is given by the set

$$((E \cap F) \cap G^c) \cup ((E \cap G) \cap F^c) \cup ((F \cap G) \cap E^c) .$$

Part (j): At most three of them occur must be the entire samples space since we only have three events total.

Problem 7 (set simplifications)

Part (a): We have that $(E \cup F) \cap (E \cup F^c) = E$.

Part (b): For the set

$$(E \cap F) \cap (E^c \cup F) \cap (E \cup F^c)$$

We begin with the set

$$\begin{aligned}(E \cap F) \cap (E^c \cup F) &= ((E \cap F) \cap E^c) \cup (E \cap F \cap F) \\ &= \emptyset \cup (E \cap F) \\ &= E \cap F .\end{aligned}$$

So the above becomes

$$\begin{aligned}(E \cap F) \cap (E \cup F^c) &= ((E \cap F) \cap E) \cup ((E \cap F) \cap F^c) \\ &= (E \cap F) \cup \emptyset \\ &= E \cap F .\end{aligned}$$

Part (c): We find that

$$\begin{aligned}(E \cup F) \cap (F \cup G) &= ((E \cup F) \cap F) \cup ((E \cup F) \cap G) \\ &= F \cup ((E \cap G) \cup (F \cap G)) \\ &= (F \cup (E \cap G)) \cup (F \cup (F \cap G)) \\ &= (F \cup (E \cap G)) \cup F \\ &= F \cup (E \cap G) .\end{aligned}$$

Problem 8 (counting partitions)

Part (a): As a simple example, we begin by considering all partition of the elements $\{1, 2, 3\}$. We have

$$\{\{1\}, \{2\}, \{3\}\}, \{\{1, 2, 3\}\}, \{\{1\}, \{2, 3\}\}, \{\{2\}, \{1, 3\}\}, \{\{3\}, \{1, 2\}\} ,$$

giving a count of five different partitions.

Part (b): Following the hint this result can be derived as follows. We select one of the $n+1$ items in our set of $n+1$ items to be denoted as special. With this item held out we partition the remaining n items into two sets a set of size k and its complement a set of size $n-k$ (we can take k values from $\{0,1,2,\ldots,n\}$). Each of these partitions has n or fewer elements. Specifically, the set of size k has T_k partitions. Lumping our special item with the set of size $n-k$ we obtain a set of size $n-k+1$. Grouped with the set of size k we have a partition of our original set of size $n+1$. Since the number of k subset elements can be chosen in $\left(\begin{array}{c} n \\ k \end{array}\right)$ ways we have

$$1+\sum_{k=1}^{n}\left(\begin{array}{c} n \\ k \end{array}\right)T_k\,,$$

possible partitions of the set $\{1,2,\ldots,n,n+1\}$. Note that the one in the above formulation represents the $k=0$ set and corresponds to the relatively trivial partition consisting of the entire set itself.

Problem 10

From the inclusion/exclusion principle we have

$$\begin{aligned} P(E\cup F\cup G) &= P(E)+P(F)+P(G)-P(E\cap F)-P(E\cap G)-P(F\cap G) \\ &+ P(E\cap F\cap G) \end{aligned}$$

Now consider the following decompositions of sets into mutually exclusive components

$$\begin{aligned} E\cap F &= (E\cap F\cap G^c)\cup(E\cap F\cap G) \\ E\cap G &= (E\cap G\cap F^c)\cup(E\cap G\cap F) \\ F\cap G &= (F\cap G\cap E^c)\cup(F\cap G\cap E)\,. \end{aligned}$$

Since each set above is mutually exclusive we have that

$$\begin{aligned} P(E\cap F) &= P(E\cap F\cap G^c)+P(E\cap F\cap G) \\ P(E\cap G) &= P(E\cap G\cap F^c)+P(E\cap G\cap F) \\ P(F\cap G) &= P(F\cap G\cap E^c)+P(F\cap G\cap E)\,. \end{aligned}$$

Adding these three sets we have that

$$P(E\cap F)+P(E\cap G)+P(F\cap G)=P(E\cap F\cap G^c)+P(E\cap F\cap F^c)+P(F\cap G\cap E^c)+3P(E\cap F\cap G)\,,$$

which when put into the inclusion/exclusion identity above gives the desired result.

Problem 11 (Bonferroni's inequality)

From the inclusion/exclusion identity for two sets we have

$$P(E\cup F)=P(E)+P(F)-P(EF)\,.$$

Since $P(E \cup F) \leq 1$, the above becomes

$$P(E) + P(F) - P(EF) \leq 1 .$$

or

$$P(EF) \geq P(E) + P(F) - 1 ,$$

which is known as Bonferroni's inequality. From the numbers given we find that

$$P(EF) \geq 0.9 + 0.8 - 1 = 0.7 .$$

Problem 12 (exactly one of E or F occurs)

Exactly one of the events E or F occurs is given by the probability of the set

$$(EF^c) \cup (E^c F) .$$

Since the two sets above are mutually exclusive the probability of this set is given by

$$P(EF^c) + P(E^c F) .$$

Since $E = (EF^c) \cup (EF)$, we then have that $P(E)$ can be expressed as

$$P(E) = P(EF^c) + P(EF) .$$

In the same way we have for $P(F)$ the following

$$P(F) = P(E^c F) + P(EF) .$$

so the above expression for our desired event (exactly one of E or F occurring) using these two expressions for $P(E)$ and $P(F)$ is given by

$$\begin{aligned} P(EF^c) + P(E^c F) &= P(E) - P(EF) + P(F) - P(EF) \\ &= P(E) + P(F) - 2P(EF) , \end{aligned}$$

as requested.

Problem 13 (E and not F)

Since $E = EF \cup EF^c$, and both sets on the right hand side of this equation are mutually exclusive we find that

$$P(E) = P(EF) + P(EF^c) ,$$

or solving for $P(EF^c)$ we find

$$P(EF^c) = P(E) - P(EF) ,$$

as expected.

Problem 15 (drawing k white balls from r total)

This is given by

$$P_k = \frac{\left(\begin{array}{c} M \\ k \end{array}\right)\left(\begin{array}{c} N \\ r-k \end{array}\right)}{\left(\begin{array}{c} M+N \\ r \end{array}\right)} \quad \text{for} \quad k \leq r\,.$$

Problem 16 (more Bonferroni)

From Bonferroni's inequality for two sets $P(EF) \geq P(E) + P(F) - 1$, when we apply this identity recursively we see that

$$\begin{aligned}
P(E_1E_2E_3\cdots E_n) &\geq P(E_1) + P(E_2E_3\cdots E_n) - 1 \\
&\geq P(E_1) + P(E_2) + P(E_3E_4\cdots E_n) - 2 \\
&\geq P(E_1) + P(E_2) + P(E_3) + P(E_4\cdots E_n) - 3 \\
&\geq \cdots \\
&\geq P(E_1) + P(E_2) + \cdots + P(E_n) - (n-1)\,.
\end{aligned}$$

That the final term is $n-1$ can be verified to be correct by evaluating this expression for $n = 2$ which yields the original Bonferroni inequality.

Problem 18 (the number of sequences with no consecutive heads)

If the first flip lands tails then we have f_{n-1} sequences that have n total flips and no consecutive heads (and that all start with a tail). If instead we get a head on the first flip then we cannot get a head on the second flip or we will have had two consecutive heads. In other words we must flip a tail for the second flip in order to count these sequences. Thus we have f_{n-2} additional sequences that we must count in this case. In total then we find

$$f_n = f_{n-1} + f_{n-2}\,.$$

Note that $f_1 = 2$ since we can toss either a head or a tail to not get two consecutive heads. We note that $f_2 = 3$ since we can through a HT, TT, or a TH and not get two consecutive heads. When we take $n = 2$ in the above we get

$$f_2 = f_1 + f_0 \quad \Rightarrow \quad 3 = 2 + f_0\,,$$

so $f_0 = 1$. The probability is given by $P_n = \frac{f_n}{2^n}$. Thus we need to compute f_{10} using the above recursion relationship.

Problem 19

k-balls will be with drawn if there are $r-1$ red balls in the first $k-1$ draws and the kth draw is the rth red ball. This happens with probability

$$\begin{aligned} P &= \frac{\binom{n}{r-1}\binom{m}{k-1-(r-1)}}{\binom{n+m}{k-1}} \cdot \frac{\binom{n-(r-1)}{1}}{\binom{n+m-(k-1)}{1}} \\ &= \frac{\binom{n}{r-1}\binom{m}{k-1-(r-1)}}{\binom{n+m}{k-1}} \cdot \left(\frac{n-(r-1)}{n+m-(k-1)}\right) . \end{aligned}$$

Here the first probability is that required to obtain $r-1$ red balls from n and $k-1-(r-1) = k-r$ blue balls from m. The next probability is the one requested to obtain the last kth red ball.

Problem 21 (counting total runs)

Following the example from 5o if we assume that we have an *even* number of total runs i.e. say $2k$, then we have two cases for the distribution of the win and loss runs. The wins and losses runs must be interleaved since we have the same number of each i.e. k, so we can start with a loosing block and end with a winning block or start with a winning block and end with a loosing block as in the following diagram

$$\begin{aligned} &L\,L\ldots L\,, W\,W\ldots W\,, L\ldots L\,, W\,W\ldots W \\ &W\,W\ldots W\,, L\,L\ldots L\,, W\ldots W\,, L\,L\ldots L\,. \end{aligned}$$

In either case, the number of wins including all winning streaks i must sum to the total number of wins n and the number of losses in all loosing streaks i must sum to the total number of losses. In equations, using x_i to denote the number of wins in the i-th winning streak and y_i to denote the number of losses in the i-th loosing streak we have that

$$\begin{aligned} x_1 + x_2 + \ldots + x_k &= n \\ y_1 + y_2 + \ldots + y_k &= m\,. \end{aligned}$$

Under the constraint that $x_i \geq 1$ and $y_i \geq 1$ since we are told that we have exactly k wins and losses (and therefore can't remove any of the unknowns. The number of solutions to the first and second equation above are given by

$$\binom{n-1}{k-1} \quad \text{and} \quad \binom{m-1}{k-1} .$$

Giving a total count on the number of possible situations where we have k winning streaks and k loosing streaks of

$$2 \cdot \binom{n-1}{k-1} \cdot \binom{m-1}{k-1}$$

Note that the "two" in the above formulation accounts for the two possibilities, i.e. we begin with a winning or loosing streak. Combined this give a probability of

$$\frac{2 \cdot \binom{n-1}{k-1} \cdot \binom{m-1}{k-1}}{\binom{n+m}{n}}.$$

If instead we are told that we have a total of $2k+1$ runs as an outcome we could have one more winning streak than loosing streak or corresponding one more loosing streak than winning streak. Assuming that we have one more winning streak than loosing our distribution of wins and looses looks schematically like the following

$$W\,W \ldots W\,,L\,L \ldots L\,,W\,W \ldots W\,,L \ldots L\,,W\,W \ldots W$$

Then counting the total number of wins and losses with our x_i and y_i variables we must have in this case

$$\begin{aligned} x_1 + x_2 + \ldots + x_k + x_{k+1} &= n \\ y_1 + y_2 + \ldots + y_k &= m\,. \end{aligned}$$

The first equation has $\binom{n-1}{k+1-1} = \binom{n-1}{k}$ solutions and the second has $\binom{m-1}{k-1}$. If instead we have one more loosing streak than winning our distribution of wins and looses looks schematically like the following

$$L\,L \ldots L\,,W\,W \ldots W\,,L\,L \ldots L\,,W \ldots W\,,L\,L \ldots L$$

Then counting the total number of wins and losses with our x_i and y_i variables we must have in this case

$$\begin{aligned} x_1 + x_2 + \ldots + x_k &= n \\ y_1 + y_2 + \ldots + y_k + y_{k+1} &= m\,. \end{aligned}$$

The first equation has $\binom{n-1}{k-1}$ solutions and the second has $\binom{m-1}{k+1-1} = \binom{m-1}{k}$.

Since either of these two mutually exclusive cases can occur the total number is given by

$$\binom{n-1}{k} \cdot \binom{m-1}{k-1} + \binom{n-1}{k-1} \cdot \binom{m-1}{k}.$$

Giving a probability of

$$\frac{\binom{n-1}{k} \cdot \binom{m-1}{k-1} + \binom{n-1}{k-1} \cdot \binom{m-1}{k}}{\binom{n+m}{n}}.$$

as expected.

Chapter 2: Self-Test Problems and Exercises

Problem 1 (a cafeteria sample space)

Part (a): We have two choices for the entree, three choices for the starch, and four choices for the dessert giving $2 \cdot 3 \cdot 4 = 24$ total outcomes in the sample space.

Part (b): Now we have two choices for the entrees, and three choices for the starch giving six total outcomes.

Part (c): Now we have three choices for the starch and four choices for the desert giving 12 total choices.

Part (d): The event $A \cap B$ means that we pick chicken for the entree and ice cream for the desert, so the three possible outcomes correspond to the three possible starches.

Part (e): We have two choices for an entree and four for a desert giving eight possible choices.

Part (f): This event is a dinner of chicken, rice, and ice cream.

Problem 2 (purchasing suits and ties)

Let S_u, S_h , and T be the events that a person purchases a s**u**it, a s**h**irt, and a **t**ie respectively. Then the problem gives the information that

$$\begin{array}{ccc} P(S_u) = 0.22 & P(S_h) = 0.3 & P(T) = 0.28 \\ P(S_u \cap S_h) = 0.11 & P(S_u \cap T) = 0.14 & P(S_h \cap T) = 0.1 \end{array}$$

and $P(S_u \cap S_h \cap T) = 0.06$.

Part (a): This is the event $P((S_u \cup S_h \cup T)^c)$, which we see is given by

$$\begin{aligned} P((S_u \cup S_h \cup T)^c) &= 1 - P(S_u \cup S_h \cup T) \\ &= 1 - P(S_u) - P(S_h) - P(T) + P(S_u \cap S_h) + P(S_u \cap T) \\ &+ P(S_h \cap T) - P(S_u \cap S_h \cap T) \\ &= 1 - 0.22 - 0.3 - 0.28 + 0.11 + 0.14 + 0.1 - 0.06 = 0.49\,. \end{aligned}$$

Part (b): Exactly one item means that we want to evaluate each of the following three mutually exclusive events

$$P(S_u \cap S_h^c \cap T^c) \quad \text{and} \quad P(S_u^c \cap S_h \cap T^c) \quad \text{and} \quad P(S_u^c \cap S_h^c \cap T)$$

and add the resulting probabilities up. We note that problem thirteen from this chapter was solved in this same way. To compute this probability we will begin by computing the probability that *two* or more items were purchased. This is the event

$$(S_u \cap S_h) \cup (S_u \cap T) \cup (S_h \cap T),$$

which we denote by E_2 for shorthand. Using the inclusion/exclusion identity we have that the probability of the event E_2 is given by

$$\begin{aligned}
P(E_2) &= P(S_u \cap S_h) + P(S_u \cap T) + P(S_h \cap T) \\
&- P(S_u \cap S_h \cap S_u \cap T) - P(S_u \cap S_h \cap S_h \cap T) - P(S_u \cap T \cap S_h \cap T) \\
&+ P(S_u \cap S_h \cap S_u \cap T \cap S_h \cap T) \\
&= P(S_u \cap S_h) + P(S_u \cap T) + P(S_h \cap T) \\
&- P(S_u \cap S_h \cap T) - P(S_u \cap S_h \cap T) - P(S_u \cap S_h \cap T) + P(S_u \cap S_h \cap T) \\
&= P(S_u \cap S_h) + P(S_u \cap T) + P(S_h \cap T) - 2P(S_u \cap S_h \cap T) \\
&= 0.11 + 0.14 + 0.1 - 2(0.06) = 0.23\,.
\end{aligned}$$

If we let E_0 and E_1 be the events that we purchase no items or one item, then the probability that we purchase exactly one item must satisfy

$$1 = P(E_0) + P(E_1) + P(E_2),$$

which we can solve for $P(E_1)$. We find that

$$P(E_1) = 1 - P(E_0) - P(E_2) = 1 - 0.49 - 0.23 = 0.28\,.$$

Problem 3 (the fourteenth card is an ace)

Since the probability that any one specific card is the fourteenth is 1/52 and we have four ways of getting an ace in the fourteenth spot we have a probability given by

$$\frac{4}{52} = \frac{1}{13}\,.$$

Another way to solve this problem is to recognized that we have 52! ways of ordering the 52 cards in the deck. Then the number of ways that the fourteenth card can be an ace is given by the fact that we have four choices for the ace in the fourteenth position and then the requirement that we need to place $52 - 1 = 51$ other cards in 51! ways so we have a probability of

$$\frac{4(51!)}{52!} = \frac{4}{52} = \frac{1}{13}\,.$$

To have the first ace occurs in the fourteenth spot we have to pick thirteen cards to place in the thirteen slots in front of this ace (from the $52 - 4 = 48$ "non" ace cards). This can be done in

$$48 \cdot 47 \cdot 46 \cdots (48 - 13 + 1) = 48 \cdot 47 \cdot 46 \cdots 36\,,$$

ways. Then we have four choices for the ace to pick in the fourteenth spot, then finally we have to place the remaining $52 - 14 = 38$ cards in 38! ways. Thus our probability is given by

$$\frac{(48 \cdot 47 \cdot 46 \cdots 36) \cdot 4 \cdot (38!)}{52!} = 0.03116\,.$$

Problem 4 (temperatures)

Let $A = \{t_{LA} = 70\}$ be the event that the temperature in LA is 70. Let $B = \{t_{NY} = 70\}$ be the event that the temperature in NY is 70. Let $C = \{\max(t_{LA}, t_{NY}) = 70\}$ be the event that the max of the two temperatures is 70. Let $D = \{\min(t_{LA}, t_{NY}) = 70\}$ be the event that the min of the two temperatures is 70. We note that $C \cap D = A \cap B$ and $C \cup D = A \cup B$. Then we want to compute $P(D)$. Since

$$P(C \cup D) = P(C) + P(D) - P(C \cap D) \, ,$$

by the inclusion/exclusion identity for two sets. We also have

$$\begin{aligned} P(C \cup D) &= P(A \cup B) = P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(C \cup D) \end{aligned}$$

By the relationship $C \cup D = A \cup B$ and the inclusion/exclusion identity for A and B. We can equate these two expressions to obtain

$$P(A) + P(B) - P(C \cap D) = P(C) + P(D) - P(C \cap D) \, ,$$

or

$$P(D) = P(A) + P(B) - P(C) = 0.3 + 0.4 - 0.2 = 0.5 \, .$$

Problem 5 (the top four cards)

Part (a): There are 52! arrangements of the cards. Then we have 52 choices for the first card, $52 - 4 = 48$ choices for the second card, $52 - 4 - 4 = 44$ choices for the third card etc. This gives a probability of

$$\frac{52 \cdot 48 \cdot 44 \cdot 40(52-4)!}{52!} = 0.67611 \, .$$

Part (b): For different suits we have 52! total arrangements and to impose that constraint that the top four all have different suits we have 52 choices for the first and then $52 - 13 = 39$ choices for the second card, $39 - 13 = 26$ choices for the third card etc. This gives a probability of

$$\frac{52 \cdot 39 \cdot 26 \cdot (52-4)!}{52!} = 0.1055 \, .$$

Problem 6 (balls of the same color)

We have this probability given by

$$\frac{\binom{3}{1}\binom{4}{1}}{\binom{6}{1}\binom{10}{1}} + \frac{\binom{3}{1}\binom{6}{1}}{\binom{6}{1}\binom{10}{1}} = \frac{1}{2} \, .$$

Where the first term is the probability that the first ball drawn is red and the second term is the probability that the second ball is drawn is black.

Problem 7 (the state lottery)

Part (a): We have

$$\frac{1}{\binom{40}{8}} = 1.3\,10^{-8}\,,$$

since there is only one way to get all eight numbers.

Part (b): We have

$$\frac{\binom{8}{7}\binom{40-8}{1}}{\binom{40}{8}} = \frac{\binom{8}{7}\binom{32}{1}}{\binom{40}{8}} = 3.3\,10^{-6}\,.$$

Part (c): To solve this part we now need the probability of selecting six numbers which is given by

$$\frac{\binom{8}{6}\binom{40-8}{2}}{\binom{40}{8}}\,,$$

which must be added to the probabilities in Part (a) and Part (b).

Problem 8 (committees)

Part (a): We have

$$\frac{\binom{3}{1}\binom{4}{1}\binom{4}{1}\binom{3}{1}}{\binom{3+4+4+3}{4}} = \frac{3\cdot 4\cdot 4\cdot 3}{\binom{14}{4}}\,.$$

Part (b): We have

$$\frac{\binom{4}{2}\binom{4}{2}}{\binom{14}{4}}\,.$$

Part (c): We can have no sophomores and four junior or one sophomore and three juniors or two sophomores and two juniors or three sophomores and one juniors or four sophomores and zero juniors. So our probability is given by

$$\frac{\binom{4}{0}\binom{4}{4}+\binom{4}{1}\binom{4}{3}+\binom{4}{2}\binom{4}{2}+\binom{4}{3}\binom{4}{1}+\binom{4}{4}\binom{4}{0}}{\binom{14}{4}}$$

From Problem 9 on Page 19 with $\binom{n}{k}=\binom{n}{n-k}$, the sum in the numerator is given by

$$\binom{2(4)}{4}=\binom{8}{4}.$$

Problem 9 (number of elements in various sets)

Both of these claims follow directly from the inclusion-exclusion identity if we assume that every element in our finite universal set S (with n elements) is equally likely and has probability $1/n$.

Problem 10 (horse experiments)

We have $N(A) = 3 \cdot 5! = 3 \cdot 120 = 360$. We have $N(B) = 5! = 120$, and

$$N(A \cap B) = 2 \cdot 4! = 2 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 2(12 \cdot 2) = 48.$$

The union gives

$$N(A \cup B) = N(A) + N(B) - N(A \cap B) = 720 + 120 - 48 = 432.$$

Problem 11

We have $\binom{52}{5}$ possible five card hands from our fifty-two cards. To have one card from each of the four suits we need to count the number of ways to select one club from the thirteen available, (this can be done in $\binom{13}{1}$ ways) one spade from the thirteen available, (this can be done in $\binom{13}{1}$ ways), one spade from the thirteen available in $\binom{13}{1}$ ways etc. The last card can be selected in

$$\binom{52-4}{1},$$

ways. Thus we have $\binom{13}{1}^4\binom{48}{1}$ possible hands containing one card from each suit, where the order of the choice made in the $\binom{48}{1}$ selections and the corresponding selection from the $\binom{13}{1}$ that has a suit that matches the $\binom{48}{1}$ selection mater. To better explain this say when picking clubs we get the three card. When we pick from the 48 remaining cards (after having selected a card of each suit) assume we select a four of clubs. This hand is equivalent to having picked the four of clubs *first* and then the three of clubs. So we must divide the above by a 2! giving a probability of

$$\frac{\frac{1}{2!}\binom{13}{1}^4\binom{48}{1}}{\binom{52}{5}} = 0.2637\,.$$

Problem 12 (basketball choices)

We have 10! ways of permutations of all the player (frontcourt and backcourt considered the same). Grouping the players list into pairs, we have five pairs and since the order within each pair does not matter we have $\frac{10!}{2^5}$ divisions of the ten players into a first roommate pair, a second roommate pair etc. Since the ordering of the roommate *pairs* does not matter we have $\frac{10!}{2^5 5!}$ pairs of roommates to choose from. Now there are

$$\binom{6}{2}\binom{4}{2}\,,$$

ways of selecting the two frontcourt and backcourt player and 2! ways of assigning them. We then have to create roommate pairs from only frontcourt and backcourt players. For the frontcourt we use the following logic to derive the number of pairs of total players

$$\frac{4!}{2^2(2!)} = 3\,.$$

For the backcourt players we have

$$\frac{2!}{2^1(1!)} = 1\,,$$

so we have a probability of

$$\frac{\binom{6}{2}\binom{4}{2}\cdot 2\cdot 3\cdot 1}{\left(\frac{10!}{2^5 5!}\right)} = 0.5714\,.$$

Problem 13 (random letter)

The same letter could be chosen if and only if it comes from one of R, E, or V. The probability of R is chosen from both words is

$$\left(\frac{2}{7}\right)\left(\frac{1}{8}\right) = \frac{2}{56}\,.$$

The probability of E is chosen from both words is

$$\left(\frac{3}{7}\right)\left(\frac{1}{8}\right) = \frac{3}{56}\,.$$

Finally the probability of V is chosen from both words is

$$\frac{1}{7}\left(\frac{1}{8}\right) = \frac{1}{56}\,.$$

So the total probability is the sum of all the above probabilities or

$$\frac{6}{56} = \frac{3}{28}\,.$$

Problem 14 (Boole's inequality)

We begin by decomposing the countable union of sets A_i

$$A_1 \cup A_2 \cup A_3 \ldots$$

into a countable union of disjoint sets C_j. Define these disjoint sets as

$$\begin{aligned}
C_1 &= A_1 \\
C_2 &= A_2 \backslash A_1 \\
C_3 &= A_3 \backslash (A_1 \cup A_2) \\
C_4 &= A_4 \backslash (A_1 \cup A_2 \cup A_3) \\
&\vdots \\
C_j &= A_j \backslash (A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_{j-1})
\end{aligned}$$

Then by construction

$$A_1 \cup A_2 \cup A_3 \cdots = C_1 \cup C_2 \cup C_3 \cdots\,,$$

and the C_j's are disjoint, so that we have

$$\Pr(A_1 \cup A_2 \cup A_3 \cup \cdots) = \Pr(C_1 \cup C_2 \cup C_3 \cup \cdots) = \sum_j \Pr(C_j)\,.$$

Since $\Pr(C_j) \le \Pr(A_j)$, for each j, this sum is bounded above by

$$\sum_j \Pr(A_j)\,,$$

Problem 15

From the fact that $\cap_i A_i$ is a *set*, its probability must be less than or equal to 1, that is

$$1 \geq P(\cap_i A_i) = P((\cup_i A_i^c)^c) = 1 - P(\cup_i A_i^c)\,.$$

By Boole's inequality we also have that

$$P(\cup_i A_i^c) \leq \sum_{i=1}^{\infty} P(A_i^c) = \sum_{i=1}^{\infty} (1 - P(A_i))\,.$$

But since $P(A_i) = 1$ each term in this sum is zeros and $P(\cup_i A_i^c) \leq 0$. Thus

$$1 \geq P(\cap_i A_i) \geq 1 - 0 = 1\,,$$

showing that $P(\cap_i A_i) = 1$.

Problem 16 (the number of non-empty partitions of size k)

Let $T_k(n)$ be the number of partitions of the set $\{1, 2, 3, \cdots, n\}$ into k nonempty subsets. Computing this number can be viewed as a counting the number of partitions with the singleton set $\{1\}$ in them, and counting the number of partitions without the singleton set $\{1\}$ in them. If $\{1\}$ is *in* a singleton set then we have used up one subset and are now looking at the number of partitions of a set of size $n-1$. Thus the number of partitions where $\{1\}$ is a singleton set must be $T_{k-1}(n-1)$. The number of partition where the element one is *in* a partition is given by $kT_k(n-1)$, since $T_k(n-1)$ gives the number of k partitions of a set of size $n-1$ and we can insert the element 1 into any of these k sets to derive a k partition of a set of n. Adding these two mutually exclusion results we obtain the following expression for $T_k(n)$

$$T_k(n) = T_{k-1}(n-1) + kT_k(n-1)\,.$$

Problem 17 (drawing balls from an urn)

Consider the complementary probability that no balls of a given color are chosen. For example let R be the event that no red balls are chosen, W the event that no white balls are chosen and B the event that no blue ball is chosen. The the desired probability is given by the complement of $P(R \cup W \cup B)$. By the inclusion/exclusion identity we have

$$P(R \cup W \cup B) = P(R) + P(W) + P(B) - P(R \cap W) - P(R \cap B) - P(W \cap B) + P(R \cap W \cap B)\,.$$

Now the individual probabilities are given by

$$
\begin{aligned}
P(R) &= \frac{\binom{13}{5}}{\binom{18}{5}}, \quad P(W) = \frac{\binom{5+7}{5}}{\binom{18}{5}} = \frac{\binom{12}{5}}{\binom{18}{5}} \\
P(B) &= \frac{\binom{5+6}{5}}{\binom{18}{5}} = \frac{\binom{11}{5}}{\binom{18}{5}} \\
P(R \cap W) &= \frac{\binom{7}{5}}{\binom{18}{5}}, \quad P(R \cap B) = \frac{\binom{6}{5}}{\binom{18}{5}} \\
P(W \cap B) &= \frac{\binom{5}{5}}{\binom{18}{5}} \quad P(R \cap W \cap B) = 0\,.
\end{aligned}
$$

Then adding these results together gives $P(R \cup W \cup B)$, and the desired result is $1 - P(R \cup W \cup B)$.

Chapter 3 (Conditional Probability and Independence)

Notes on the Text

The probability that event E happens before event F (page 93)

Let A be the event that E occurs before F (E and F are mutually exclusive). Here we are envisioning independent trials where E or F or $(E \cup F)^c$ are the only possible occurrences of each experiment. Then conditioning on each of these three events we have that

$$\begin{aligned} P(A) &= P(A|E)P(E) + P(A|F)P(F) + P(A|(E \cup F)^c)P((E \cup F)^c) \\ &= P(E) + (1 - P(E) - P(F))P(A) \,. \end{aligned}$$

Since $P(A|E) = 1$, $P(A|F) = 0$ and $P(A|(E \cup F)^c) = P(A)$. Solving for $P(A)$ gives

$$P(A) = \frac{P(E)}{P(E) + P(F)} \,. \tag{6}$$

From the symmetry of that equation we have that the probability that F happens before event E is then

$$P(A^c) = \frac{P(F)}{P(E) + P(F)} \,.$$

The duration of play problem (page 98)

I found this section of the text difficult to understand at first and wrote this simple explanation to help myself understand things better. In the ending arguments of example 4k, Ross applies the gamblers ruin problem to the duration of play problem of Huygens. In the duration of play problem if an eleven is thrown (with probability of $\frac{27}{216}$) the player B wins a point, if a fourteen is thrown (with probability of $\frac{15}{216}$) the player A wins a point, while anything else results in a continuation of the game. Since the outcome that A wins a point will only happen if a fourteen is thrown before an eleven we need to compute *that* probability to apply to the gamblers ruin problem. The probability that a fourteen is thrown before an eleven is given by example 4h and equals

$$\frac{P(E)}{P(E) + P(F)} = \frac{\frac{15}{216}}{\frac{27}{216} + \frac{15}{216}} = \frac{15}{42} \,,$$

the number given for p in the text.

Problem Solutions

Problem 1 (fair dice)

Let E be the event that at least one die is a six and F the event that the two die lands on different numbers. Then $P(E|F) = \frac{P(EF)}{P(F)}$. The event F can be any of the following pairs

$$(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,3),(2,4),(2,5),(2,6),(3,1),(3,2),(3,4),$$
$$(3,5),(3,6),(4,1),(4,2),(4,3),(4,5),(4,6),(5,1),(5,2),(5,3),(5,4),(5,6),(6,1),$$
$$(6,2),(6,3),(6,4), \quad \text{and} \quad (6,5)\,,$$

which has thirty elements giving a probability $P(F) = \frac{30}{36} = \frac{5}{6}$.

The event EF consist of the event where at least one die is a six and the other two die have different numbers. The elements of this set are given by

$$(1,6),(2,6),(3,6),(4,6),(5,6),(6,1),(6,2),(6,3),(6,4),(6,5)\,,$$

which has ten elements so $P(EF) = \frac{10}{36} = \frac{5}{18}$. With these two results we have

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{(5/18)}{(5/6)} = \frac{6}{18} = \frac{1}{3}\,.$$

Problem 2 (more fair dice)

Let E_i be the event that the sum of the two dice is i. Let F_6 denote the event that the first die is a six. Then we want to compute $P(F_6|E_i)$ for $i = 2, 3, \cdots, 12$. This expression is given by

$$\frac{P(F_6 \cap E_i)}{P(E_i)}\,.$$

From Problem 24 from Chapter 2 we know the values of $P(E_i)$ for $i = 2, 3, \cdots, 12$. Thus we only need to compute the events $F_6 \cap E_i$ for each i. We have (if ϕ is the empty set)

$$\begin{aligned}
F_6 \cap E_2 &= \phi, F_6 \cap E_4 = \phi, F_6 \cap E_6 = \phi\,, F_6 \cap E_3 = \phi, F_6 \cap E_5 = \phi, F_6 \cap E_7 = \{6,1\}\,, \\
F_6 \cap E_8 &= \{6,2\}, F_6 \cap E_9 = \{6,3\}, F_6 \cap E_{10} = \{6,4\}\,, F_6 \cap E_{11} = \{6,5\}, \quad \text{and} \\
F_6 \cap E_{12} &= \{6,6\}\,,
\end{aligned}$$

Thus if $F_6 \cap E_i = \phi$ then $P(F_6 \cap E_i) = 0$, while if $F_6 \cap E_i \neq \phi$, then $P(F_6 \cap E_i) = \frac{1}{36}$ So we get

$$P(F_6|E_2) = 0, P(F_6|E_3) = 0, P(F_6|E_4) = 0, P(F_6|E_5) = 0, P(F_6|E_6) = 0\,.$$

along with

$$
\begin{aligned}
P(F_6|E_7) &= \frac{1/36}{P(E_7)} = \frac{1/36}{6/36} = \frac{1}{6} \\
P(F_6|E_8) &= \frac{1/36}{P(E_8)} = \frac{1/36}{5/36} = \frac{1}{5} \\
P(F_6|E_9) &= \frac{1/36}{P(E_9)} = \frac{1/36}{4/36} = \frac{1}{4} \\
P(F_6|E_{10}) &= \frac{1/36}{P(E_{10})} = \frac{1/36}{3/36} = \frac{1}{3} \\
P(F_6|E_{11}) &= \frac{1/36}{P(E_{11})} = \frac{1/36}{2/36} = \frac{1}{2} \\
P(F_6|E_{12}) &= \frac{1/36}{P(E_{12})} = \frac{1/36}{1/36} = 1\,.
\end{aligned}
$$

Problem 3 (hands of bridge)

Equation 2.1 in the book is

$$
p(E|F) = \frac{p(EF)}{p(F)}\,.
$$

To use this equation for this problem, let E be the event that East has three spades and F be the event that the combined North-South pair has eight spades. Then

$$
P(F) = \frac{\binom{13}{8}\binom{39}{18}}{\binom{52}{26}}\,.
$$

This can be reasoned as follows. We have thirteen total spades from which we should pick eight to the give the North-South pair (the rest will go to the East-West pair). This gives the factor $\binom{13}{8}$. We then have $52 - 13 = 39$ other cards (non-spades) from which to pick the remaining $26 - 8 = 18$ cards to make the required total of 26 cards for the North-South pair. This gives the factor $\binom{39}{18}$. The product of these two expressions gives the total number of ways we can obtain the stated condition. This product is divided the number of ways to select 26 cards from the deck of 52 total cards. When we evaluate the above fraction we find $P(F) = 9102/56243 = 0.161833$.

Now the joint event EF means that East has three spades and North-South has eight spades so that West must have $13 - 3 - 8 = 2$ spades. Thus to evaluate $P(EF)$ the approach we take is to enumerate the required number and type of cards to East and then do the same for West. For each player we do this in two parts, first the number of spade cards and then the number of non-spade cards. Using this logic we find that

$$P(EF) = \frac{\binom{13}{3}\binom{39}{10}\binom{10}{2}\binom{52-13-10}{11}}{\binom{52}{13}\binom{39}{13}}.$$

This can be reasoned as follows. The first factor $\binom{13}{3}$ in the numerator is the number of ways we can select three required spades for East. The second factor $\binom{39}{10}$ is the number of ways we can select the remaining $13 - 3 = 10$ non-spade cards for East. The third factor $\binom{10}{2}$ is the number of ways we can select the required two spade cards for West. We then have $52 - 13 - 10$ remaining possible non-spade cards from which we need to draw 11 to complete the hand of West. This gives the factor $\binom{52-13-10}{11}$. The denominator is the number of ways we can draw East and West's hands without any restrictions. When we evaluate the above fraction we find $P(EF) = 2397/43675 = 0.054883$.

With these two results we see that $P(E|F) = 39/115 = 0.339130$, the same as in the back of the book. See the Matlab/Octave file `chap_3_prob_3.m` for the evaluation of the above fractions.

Problem 4 (at least one six)

This is solved in the same way as in problem number 2. In solving we will let E be the event that at least one of the pair of dice lands on a 6 and "$X = i$" be shorthand for the event the sum of the two dice is i. Then we desire to compute

$$p(E|X = i) = \frac{P(E, X = i)}{p(X = i)}.$$

We begin by computing $p(X = i)$ for $i = 2, 3, 4, \cdots, 12$. We find that

$$\begin{aligned}
p(X = 2) &= \frac{1}{36}, \quad p(X = 8) = \frac{5}{36} \\
p(X = 3) &= \frac{2}{36}, \quad p(X = 9) = \frac{4}{36} \\
p(X = 4) &= \frac{3}{36}, \quad p(X = 10) = \frac{3}{36} \\
p(X = 5) &= \frac{4}{36}, \quad p(X = 11) = \frac{2}{36} \\
p(X = 6) &= \frac{5}{36}, \quad p(X = 12) = \frac{1}{36} \\
p(X = 7) &= \frac{6}{36}.
\end{aligned}$$

We next compute $p(E, X = i)$ for $i = 2, 3, 4, \cdots, 12$ we find that

$$\begin{aligned}
p(E, X = 2) &= 0\,, \quad p(E, X = 8) = \frac{2}{36} \\
p(E, X = 3) &= 0\,, \quad p(E, X = 9) = \frac{2}{36} \\
p(E, X = 4) &= 0\,, \quad p(E, X = 10) = \frac{2}{36} \\
p(E, X = 5) &= 0\,, \quad p(E, X = 11) = \frac{2}{36} \\
p(E, X = 6) &= 0\,, \quad p(E, X = 12) = \frac{1}{36} \\
p(E, X = 7) &= \frac{2}{36}\,.
\end{aligned}$$

Finally computing our conditional probabilities we find that

$$P(E|X = 2) = p(E|X = 3) = p(E|X = 4) = p(E|X = 5) = p(E|X = 6) = 0\,.$$

and

$$\begin{aligned}
p(E|X = 7) &= \frac{1}{3}\,, \quad p(E|X = 10) = \frac{2}{3} \\
p(E|X = 8) &= \frac{2}{5}\,, \quad p(E|X = 11) = \frac{2}{2} = 1 \\
p(E|X = 9) &= \frac{1}{2}\,, \quad p(E|X = 12) = \frac{1}{1} = 1\,.
\end{aligned}$$

Problem 5 (the first two selected are white)

We have that

$$P = \frac{\binom{6}{2}\binom{9}{2}}{\binom{15}{4}}$$

is the probability of drawing two white balls and two black balls independently of the order of the draws. Since we are concerned with the probability of an *ordered* sequence of draws we should enumerate these. Let W by the event that the first two balls are white and B the event that the second two balls are black. Then we desire the probability $P(W \cap B) = P(W)\,P(B|W)$. Now

$$P(W) = \frac{\binom{6}{2}}{\binom{15}{2}} = \frac{15}{105} \approx 0.152$$

and

$$P(B|W) = \frac{\binom{9}{2}}{\binom{13}{2}} = \frac{36}{78} \approx 0.461$$

so that $P(W \cap B) = 0.0659 = \frac{6}{91}$.

Problem 6 (exactly three white balls)

Let F be the event that the first and third drawn balls are white and let E be the event that the sample contains exactly three white balls. Then we desire to compute $P(F|E) = \frac{P(F \cap E)}{P(E)}$. Working the without replacement we have that

$$P(E) = \frac{\left(\begin{array}{c} 8 \\ 3 \end{array}\right) \cdot \left(\begin{array}{c} 4 \\ 1 \end{array}\right)}{\left(\begin{array}{c} 12 \\ 4 \end{array}\right)} = \frac{224}{495}.$$

and $P(F \cap E)$ is the probability that our sample has three white balls and the first and third balls are white. To calculate this we can explicitly enumerate the possibilities in $F \cap E$ as $\{(W, W, W, B), (W, B, W, W)\}$, showing that

$$P(F \cap E) = \frac{2}{\left(\begin{array}{c} 12 \\ 4 \end{array}\right)}.$$

Given these two results we then have that

$$P(F|E) = \frac{2}{\left(\begin{array}{c} 8 \\ 3 \end{array}\right) \cdot \left(\begin{array}{c} 4 \\ 1 \end{array}\right)} = \frac{1}{112}.$$

To work the problem with replacement we have that

$$P(E) = \left(\begin{array}{c} 4 \\ 3 \end{array}\right) \left(\frac{2}{3}\right)^3 \left(\frac{1}{3}\right) = \frac{2^5}{3^4}.$$

As before we can enumerate the sample in $E \cap F$. This set is $\{(W, W, W, B), (W, B, W, W)\}$, and has probabilities given by

$$\left(\frac{2}{3}\right)^3 \frac{1}{3} + \left(\frac{2}{3}\right)^3 \frac{1}{3} = \frac{2^4}{3^4}.$$

so the probabilities we are after is

$$\frac{\frac{2^4}{3^4}}{\frac{2^5}{3^4}} = \frac{1}{2}.$$

Problem 7 (the king's sister)

The two possible children have a sample space given by

$$\{(M, M), (M, F), (F, M), (F, F)\},$$

each with probability 1/4. Then if we let E be the event that one child is a male and F be the event that one child is a female and one child is a male, the probability that we want to compute is given by

$$P(F|E) = \frac{P(FE)}{P(E)}.$$

Now

$$P(E) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

and FE consists of the set $\{(M, F), (F, M)\}$ so

$$P(FE) = \frac{1}{2},$$

so that

$$P(F|E) = \frac{1/2}{3/4} = \frac{2}{3}.$$

Problem 8 (two girls)

Let F be the event that both children are girls and E the event that the eldest child is a girl. Now $P(E) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ and the event EF has probability $\frac{1}{4}$. Then

$$P(F|E) = \frac{P(FE)}{P(E)} = \frac{1/4}{1/2} = \frac{1}{2}.$$

Problem 9 (a white ball from urn A)

Let F be the event that the ball chosen from urn A was white. Let E be the event that two while balls were chosen. Then the desired probability is $P(F|E) = \frac{P(FE)}{P(E)}$. Lets first calculate $P(E)$ or the probability that two white balls were chosen. This event can happen in the following mutually exclusive draws

$$(W, W, R), (W, R, W), (R, W, W).$$

We can calculate the probabilities of each of these events

- The first draw will happen with probability $\left(\frac{2}{6}\right)\left(\frac{8}{12}\right)\left(\frac{3}{4}\right) = \frac{1}{6}$
- The second draw will happen with probability $\left(\frac{1}{3}\right)\left(\frac{4}{12}\right)\left(\frac{1}{4}\right) = \frac{1}{36}$
- The third draw will happen with probability $\left(\frac{4}{6}\right)\left(\frac{8}{12}\right)\left(\frac{1}{4}\right) = \frac{1}{9}$

so that

$$P(E) = \frac{1}{6} + \frac{1}{36} + \frac{1}{9} = \frac{11}{36}.$$

Now FE consists of only the events $\{(W, W, R), (W, R, W)\}$ since now the first draw must be white. The event FE has probability given by $\frac{1}{6} + \frac{1}{36} = \frac{7}{36}$, so that we find

$$P(F|E) = \frac{7/36}{11/36} = \frac{7}{11} = 0.636\,.$$

Problem 10 (three spades given that we draw two others)

Let F be the event that the first card selected is a spade and E the event that the second and third cards are spades. Then we desire to compute $P(F|E) = \frac{P(FE)}{P(E)}$. Now $P(E)$ is the probability that the second and third cards are spades, which equals the union of two events. The first is event that the first, second, and third cards are spades and the second is the event that the first card is not a spade while the second and third cards are spades. Note that this first event is also FE above. Thus we have

$$P(FE) = \frac{13 \cdot 12 \cdot 11}{52 \cdot 51 \cdot 50}$$

Letting G be the event that the first card is not a spade while the second and third cards are spades, we have that

$$P(G) = \frac{(52 - 13) \cdot 13 \cdot 12}{52 \cdot 51 \cdot 50} = \frac{39 \cdot 13 \cdot 12}{52 \cdot 51 \cdot 50}\,,$$

so

$$P(E) = \frac{39 \cdot 13 \cdot 12}{52 \cdot 51 \cdot 50} + \frac{13 \cdot 12 \cdot 11}{52 \cdot 51 \cdot 50} = \frac{11}{39 + 11} = \frac{11}{50} = 0.22\,.$$

Problem 11 (probabilities on two cards)

We are told to let B be the event that both cards are aces, A_s the event that the ace of spades is chosen and A the event that at least one ace is chosen.

Part (a): We are asked to compute $P(B|A_s)$. Using the definition of conditional probabilities we have that

$$P(B|A_s) = \frac{P(BA_s)}{P(A_s)}\,.$$

The event BA_s is the event that both cards are aces and one is the ace of spades. This event can be represented by the sample space

$$\{(AD, AS), (AH, AS), (AC, AS)\}\,.$$

where D, S, H, and C stand for diamonds, spades, hearts, and clubs respectively and the order of these elements in the set above does not matter. So we see that

$$P(BA_s) = \frac{3}{\binom{52}{2}}\,.$$

The event A_s is given by the set $\{AS, *\}$ where $*$ is a wild-card denoting any of the possible fifty-one other cards besides the ace of spades. Thus we see that

$$P(A_s) = \frac{51}{\binom{52}{2}}.$$

These together give that

$$P(B|A_s) = \frac{3}{51} = \frac{1}{17}.$$

Part (b): We are asked to compute $P(B|A)$. Using the definition of conditional probabilities we have that

$$P(B|A) = \frac{P(BA)}{P(A)} = \frac{P(B)}{P(A)}.$$

The event B are the hand $\{(AD, AS), (AD, AH), (AD, \cdots)\}$ and has $\binom{4}{2}$ elements i.e. from the four total aces select two. So that

$$P(B) = \frac{\binom{4}{2}}{\binom{52}{2}}.$$

The set A is the event that at least one ace is chosen. This is the complement of the set that no ace is chosen. No ace can be chosen in $\binom{48}{2}$ ways so that

$$P(A) = 1 - \frac{\binom{48}{2}}{\binom{52}{2}} = \frac{\binom{52}{2} - \binom{48}{2}}{\binom{52}{2}}.$$

This gives for $P(B|A)$ the following

$$P(B|A) = \frac{\binom{4}{2}}{\binom{52}{2} - \binom{48}{2}} = \frac{6}{198} = \frac{1}{33}.$$

Problem 12 (passing the actuarial exams)

We let E_i be the event that the ith actuarial exam is passed. Then the given probabilities can be expressed as

$$P(E_1) = 0.9, \quad P(E_2|E_1) = 0.8, \quad P(E_3|E_1, E_2) = 0.7.$$

Part (a): The desired probability is given by $P(E_1E_2E_3)$ or conditioning we have

$$P(E_1E_2E_3) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2) = 0.9 \cdot 0.8 \cdot 0.7 = 0.504\,.$$

Part (b): The desired probability is given by $P(E_2^c|(E_1E_2E_3)^c)$ and can be expressed using the set identity

$$(E_1E_2E_3)^c = E_1 \cup (E_1E_2^c) \cup (E_1E_2E_3^c)\,,$$

are the only ways that one can not pass all three tests i.e. one must fail one of the first three tests. Note that these sets are mutually independent. Now

$$P(E_2^c|(E_1E_2E_3)^c) = \frac{P(E_2^c\,(E_1E_2E_3)^c)}{P((E_1E_2E_3)^c)}\,.$$

We know how to compute $P((E_1E_2E_3)^c)$ because it is equal to $1 - P(E_1E_2E_3)$ and we can compute $P(E_1E_2E_3)$. From the above set identity the event $E_2^c\,(E_1E_2E_3)^c$ is composed of only one set, namely $E_1E_2^c$, since if we don't pass the second test we don't take the third test. We now need to evaluate the probability of this event. We find

$$\begin{aligned}
P(E_1E_2^c) &= P(E_2^c|E_1)P(E_1) \\
&= (1 - P(E_2|E_1))P(E_1) \\
&= (1 - 0.8)(0.9) = 0.18\,.
\end{aligned}$$

With this the conditional probability sought is given by $\frac{0.18}{1-0.504} = 0.3629$

Problem 13

Define p by $p \equiv P(E_1E_2E_3E_4)$. Then by conditioning on the events E_1, E_1E_2, and $E_1E_2E_3$ we see that p is given by

$$\begin{aligned}
p &= P(E_1E_2E_3E_4) \\
&= P(E_1)P(E_2E_3E_4|E_1) \\
&= P(E_1)P(E_2|E_1)P(E_3E_4|E_1E_2) \\
&= P(E_1)P(E_2|E_1)P(E_3|E_1E_2)P(E_4|E_1E_2E_3)\,.
\end{aligned}$$

So we need to compute each probability in this product. We have

$$
\begin{aligned}
P(E_1) &= \frac{\binom{4}{1}\binom{48}{12}}{\binom{52}{13}} \\
P(E_2|E_1) &= \frac{\binom{3}{1}\binom{36}{12}}{\binom{39}{13}} \\
P(E_3|E_1E_2) &= \frac{\binom{2}{1}\binom{24}{12}}{\binom{26}{13}} \\
P(E_4|E_1E_2E_3) &= \frac{\binom{1}{1}\binom{12}{12}}{\binom{13}{13}} = 1\,.
\end{aligned}
$$

so this probability is then given by (when we multiply each of the above expressions)

$$p = 0.1055\,.$$

See the Matlab file `chap_3_prob_13.m` for these calculations.

Problem 14

Part (a): We will compute this as a conditional probability since the number of each colored balls depend on the results from the previous draws. Let B_i be the event that a black ball is selected on the ith draw and W_i the event that a white ball is selected on the ith draw. Then the probability we are looking for is given by

$$
\begin{aligned}
P(B_1B_2W_3W_4) &= P(B_1)P(B_2|B_1)P(W_3|B_1B_2)P(W_4|B_1B_2W_3) \\
&= \left(\frac{7}{5+7}\right)\left(\frac{9}{5+9}\right)\left(\frac{5}{5+11}\right)\left(\frac{7}{7+11}\right) = 0.0455\,.
\end{aligned}
$$

See the Matlab file `chap_3_prob_14.m` for these calculations.

Part (b): The set discussed is given by the $\binom{4}{2} = 6$ sets given by

$$
\begin{aligned}
&(B_1, B_2, W_3, W_4)\,, \quad (B_1, W_2, B_3, W_4)\,, \quad (B_1, W_2, W_3, B_4) \\
&(W_1, B_2, B_3, B_4)\,, \quad (W_1, B_2, W_3, B_4)\,, \quad (W_1, W_2, B_3, B_4)\,.
\end{aligned}
$$

The probabilities of each of these events can be computed as in Part (a) of this problem. The probability requested is then the sum of the probabilities of all these mutually exclusive events.

Problem 15 (ectopic pregnancy among smokers)

Let S be the event a woman is a smoker and E the event that a woman has an ectopic pregnancy. Then the information given in the problem statement is that $P(E|S) = 2P(E|S^c)$, $P(S) = 0.32$, $P(S^c) = 0.68$, and we want to calculate $P(S|E)$. We have using Bayes' rule that

$$\begin{aligned} P(S|E) &= \frac{P(E|S)P(S)}{P(E|S)P(S) + P(E|S^c)P(S^c)} \\ &= \frac{2P(E|S^c)(0.32)}{2P(E|S^c)(0.32) + P(E|S^c)(0.68)} \\ &= \frac{2(0.32)}{2(0.32) + 0.68} = 0.4848\,. \end{aligned}$$

Problem 16 (surviving a Cesarean birth)

Let C be the event of a Cesarean section birth, let S be the event that the baby survives. The facts given in the problem are that

$$P(S) = 0.98\,, \quad P(S^c) = 0.02\,, \quad P(C) = 0.15\,, \quad P(C^c) = 0.85\,, \quad P(S|C) = 0.96\,.$$

We want to calculate $P(S|C^c)$. We can compute $P(S)$ by C (the type of birth) as

$$P(S) = P(S|C)P(C) + P(S|C^c)P(C^c)\,.$$

Using the information given in the problem into the above we find that

$$0.98 = 0.96(0.15) + P(S|C^c)(0.85)\,,$$

or that $P(S|C^c) = 0.983$.

Problem 17 (owning pets)

Let D be the event a family owns a dog, and C the event that a family owns a cat. Then from the numbers given in the problem we have that $P(D) = 0.36$, $P(C) = 0.3$, and $P(C|D) = 0.22$.

Part (a): We are asked to compute $P(CD) = P(C|D)P(D) = 0.22 \cdot 0.36 = 0.0792$.

Part (b): We are asked to compute

$$P(D|C) = \frac{P(C|D)P(D)}{P(C)} = \frac{0.22 \cdot (0.36)}{0.3} = 0.264\,.$$

Problem 18 (types of voters)

Let I, L, and C be the event that a random person is an independent, liberal, or a conservative respectfully. Let V be the event that a person voted. Then from the problem we are given that

$$P(I) = 0.46\,,\quad P(L) = 0.3\,,\quad P(C) = 0.24\,,$$

and

$$P(V|I) = 0.35\,,\quad P(V|L) = 0.62\,,\quad P(V|C) = 0.58\,.$$

We want to compute $P(I|V)$, $P(L|V)$, and $P(C|V)$ which by Bayes' rule are given by (for $P(I|V)$ for example)

$$P(I|V) = \frac{P(V|I)P(I)}{P(V)} = \frac{P(V|I)P(I)}{P(V|I)P(I) + P(V|L)P(L) + P(V|C)P(C)}\,.$$

All desired probabilities will need to calculate $P(V)$ which we do (as above) by conditioning on the various types of voters. We find that it is given by

$$\begin{aligned} P(V) &= P(V|I)P(I) + P(V|L)P(L) + P(V|C)P(C) \\ &= 0.35(0.46) + 0.62(0.3) + 0.58(0.24) = 0.4862\,. \end{aligned}$$

Then the requested conditional probabilities are given by

$$\begin{aligned} P(I|V) &= \frac{0.35(0.46)}{0.48} = 0.3311 \\ P(L|V) &= \frac{P(V|L)P(L)}{P(V)} = \frac{0.62(0.3)}{0.4862} = 0.38256 \\ P(C|V) &= \frac{P(V|C)P(C)}{P(V)} = \frac{0.58(0.24)}{0.4862} = 0.2863\,. \end{aligned}$$

Part (d): This is $P(V)$ which from Part (c) we know to be equal to 0.48.

Problem 19 (attending a smoking success party)

Let M be the event a person who attends the party is male, W the event a person who attends the party is female, and E the event that a person was smoke free for a year. The problem gives

$$P(E|M) = 0.37\,,\quad P(M) = 0.62\,,\quad P(E|W) = 0.48\,,\quad P(W) = 1 - P(M) = 0.38\,.$$

Part (a): We are asked to compute $P(W|E)$ which by Bayes' rule is given by

$$\begin{aligned} P(W|E) &= \frac{P(E|W)P(W)}{P(E)} = \frac{P(E|W)P(W)}{P(E|W)P(W) + P(E|M)P(M)} \\ &= \frac{0.48(0.38)}{0.48(0.38) + 0.37(0.62)} = 0.442\,. \end{aligned}$$

Part (b): For this part we want to compute $P(E)$ which by conditioning on the sex of the person equals $P(E) = P(E|W)P(W) + P(E|M)P(M) = 0.4118$.

Problem 20 (majoring in computer science)

Let F be the event that a student is female. Let C be the event that a student is majoring in computer science. Then we are told that $P(F) = 0.52$, $P(C) = 0.05$, and $P(FC) = 0.02$.

Part (a): We are asked to compute $P(F|C) = \frac{P(FC)}{P(C)} = \frac{0.02}{0.05} = 0.4$.

Part (b): We are asked to compute $P(C|F) = \frac{P(FC)}{P(F)} = \frac{0.02}{0.52} = 0.3846$.

Problem 21 (salaries for married workers)

We are given the following joint probabilities

$$
\begin{aligned}
P(W_<, H_<) &= \frac{212}{500} = 0.424 \\
P(W_<, H_>) &= \frac{198}{500} = 0.396 \\
P(W_>, H_<) &= \frac{36}{500} = 0.072 \\
P(W_>, H_>) &= \frac{54}{500} = 0.108 \,.
\end{aligned}
$$

Where the notation $W_<$ is the event that the wife makes less than $25,000$, $W_>$ is the event that the wife makes more than $25,000$, $H_<$ and $H_>$ are the events that the husband makes less than or more than $25,000$ respectively.

Part (a): We desire to compute $P(H_<)$, which we can do by considering all possible situations involving the wife. We have

$$
P(H_<) = P(H_<, W_<) + P(H_<, W_>) = \frac{212}{500} + \frac{36}{500} = 0.496 \,.
$$

Part (b): We desire to compute $P(W_>|H_>)$ which we do by remembering the definition of conditional probability. We have $P(W_>|H_>) = \frac{P(W_>, H_>)}{P(H_>)}$. Since $P(H_>) = 1 - P(H_<) = 1 - 0.496 = 0.504$ using the above we find that $P(W_>|H_>) = 0.2142 = \frac{3}{14}$.

Part (c): We have

$$
P(W_>|H_<) = \frac{P(W_>, H_<)}{P(H_<)} = \frac{0.072}{0.496} = 0.145 = \frac{9}{62} \,.
$$

Problem 22 (ordering colored dice)

Part (a): The probability that no two dice land on the same number means that each die must land on a unique number. To count the number of such possible combinations we see that there are six choices for the red die, five choices for the blue die, and then four choices for the yellow die yielding a total of $6 \cdot 5 \cdot 4 = 120$ choices where each die has a different number. There are a total of 6^3 total combinations of all possible die through giving a probability of

$$\frac{120}{6^3} = \frac{5}{9}$$

Part (b): We are asked to compute $P(B < Y < R|E)$ where E is the event that no two dice land on the same number. From Part (a) above we know that the count of the number of rolls that satisfy event E is 120. Now the number of rolls that satisfy the event $B < Y < R$ can be counted in a manner like Problem 6 from Chapter 1. For example, if R shows a roll of three then the only possible valid rolls where $B < Y < R$ for B and Y are $B = 1$ and $Y = 2$. If R shows a four then we have $\left(\begin{array}{c} 3 \\ 2 \end{array}\right) = 3$ possible choices i.e. either

$$(B = 1, Y = 2)\,, \quad (B = 1, Y = 3)\,, \quad (B = 2, Y = 3)\,.$$

for the possible assignments to the two values for the B and Y die. If $R = 5$ we have $\left(\begin{array}{c} 4 \\ 2 \end{array}\right) = 6$ possible assignments to B and Y. Finally, if $R = 6$ we have $\left(\begin{array}{c} 5 \\ 2 \end{array}\right) = 10$ possible assignments to B and Y. Thus we find that

$$P(B < Y < R|E) = \frac{1 + 3 + 6 + 10}{120} = \frac{1}{6}$$

Part (c): We see that

$$P(B < Y < R) = P(B < Y < R|E)P(E) + P(B < Y < R|E^c)P(E^c)\,,$$

Since $P(B < Y < R|E^c) = 0$ from the above we have that

$$P(B < Y < R) = \left(\frac{1}{6}\right)\left(\frac{5}{9}\right) = \frac{5}{54}\,.$$

Problem 23 (some urns)

Part (a): Let W be the event that the ball chosen from urn II is white. Then we should solve this problem by conditioning on the color of the ball drawn from first urn. Specifically

$$P(W) = P(W|B_I = w)P(B_I = w) + P(W|B_I = r)P(B_I = r)\,.$$

Here $B_I = w$ is the event that the ball drawn from the first urn is white and $B_I = r$ is the event that the the drawn ball is red. We know that $P(B_I = w) = \frac{1}{3}$, $P(B_I = r) = \frac{2}{3}$, $P(W|B_I = w) = \frac{2}{3}$, and $P(W|B_I = r) = \frac{1}{3}$. We then have

$$P(W) = \frac{2}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3} = \frac{2+2}{9} = \frac{4}{9}$$

Part (b): Now we are looking for

$$P(B_I = w|W) = \frac{P(W|B_I = w)P(B_I = w)}{P(W)}.$$

Since everything is known in the above we can compute this as

$$P(B_I = w|W) = \frac{\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)}{\frac{4}{9}} = \frac{1}{2}.$$

Problem 24 (painted balls in an urn)

Part (a): Let E be the event that both balls are gold and F the event that at least one ball is gold. The probability we desire to compute is then $P(E|F)$. Using the definition of conditional probability we have that

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{P(\{G,G\})}{P(\{G,G\},\{G,B\},\{B,G\})} = \frac{1/4}{1/4+1/4+1/4} = \frac{1}{3}$$

Part (b): Since now the balls are mixed together in the urn, the difference between the pair $\{G,B\}$ and $\{B,G\}$ is no longer present. Thus we really have two cases to consider.

- Either both balls are gold or
- One ball is gold and the other is black.

Thus to have a second ball be gold will occur once out of these two choices and our probability is then 1/2.

Problem 25 (estimating the number of people over fifty)

Let F denote the event that a person is over fifty and denote this probability by p which is also the number we desire to estimate. Let α_1 denote the proportion of the time a person *under* fifty spends on the streets and α_2 the same proportion for people *over* fifty. Let S

denote the event that a person (of any age) is found in the streets. Then this event S can be decomposed into the sets where the person on the streets is less than or greater than fifty as

$$S = SF \cup SF^c .$$

Since the two sets on the right-hand-side of this expression are disjoint we have

$$P(S) = P(SF) + P(SF^c) .$$

These sets can be written in terms of S conditional on the persons age F as

$$\begin{aligned} P(SF) &= P(F)P(S|F) = pP(S|F) \\ P(SF^c) &= P(F^c)P(S|F^c) = (1-p)P(S|F) . \end{aligned}$$

Now by taking measurements of the number/proportion of people over fifty during the day as suggested by the initial part of this problem we are actually measuring the probability

$$P(F|S) ,$$

and not $P(F)$. The expression $P(F|S)$ is related to p and what we desire to measure by

$$P(F|S) = \frac{P(SF)}{P(S)} = \frac{pP(S|F)}{pP(S|F) + (1-p)P(S|F^c)} .$$

Since we are told that α_1 should be the proportion of time someone under the age of fifty spends in the streets we can express this variable in terms of the above expressions simply as $P(S|F^c)$. In the same way $P(S|F) = \alpha_2$. Using this notation we thus have

$$\begin{aligned} P(F|S) &= \frac{\alpha_2 p}{\alpha_2 p + \alpha_1(1-p)} \\ &= \frac{\alpha_2 p}{\alpha_1 + (\alpha_2 - \alpha_1)p} . \end{aligned}$$

From the above we see that if $\alpha_1 = \alpha_2$ we will have $P(F|S) = p$ and we will have actually measured what we intended to measure.

Problem 26 (colorblindness)

From the problem, assuming that CB represents the event that a person is colorblind, we are told that

$$P(CB|M) = 0.05 , \quad \text{and} \quad P(CB|W) = 0.0025 .$$

We are asked to compute $P(M|CB)$, which we will do by using the Bayes' rule. We find

$$P(M|CB) = \frac{P(CB|M)P(M)}{P(CB)} .$$

We will begin by computing $P(CB)$ by conditioning on the sex of the person. We have

$$\begin{aligned} P(CB) &= P(CB|M)P(M) + P(CB|F)P(F) \\ &= 0.05(0.5) + 0.0025(0.5) = 0.02625 . \end{aligned}$$

Then using Bayes' rule we find that

$$P(M|CB) = \frac{0.05(0.5)}{0.02625} = 0.9523 = \frac{20}{21}\,.$$

If the population consisted of twice as many males as females we would then have $P(M) = 2P(F)$ giving $P(M) = \frac{2}{3}$ and $P(F) = \frac{1}{3}$ and our calculation becomes

$$P(CB) = 0.05\left(\frac{2}{3}\right) + 0.0025\left(\frac{1}{3}\right) = 0.03416\,.$$

so that

$$P(M|CB) = \frac{0.05(2/3)}{0.03416} = 0.9756 = \frac{40}{41}\,.$$

Problem 27 (counting the number of people in each car)

Since we desire to estimate the number of people in a given car, if we choose the first method we will place too much emphasis on cars that carry a large number of people. For example if we imagine that a large bus of people arrives then on average we will select more people from this bus than from cars that only carry one person. This is the same effect as in the discussion in the book about the number of students counted on various numbers of buses and would not provide an unbiased estimate. The second method suggested would provide an unbiased estimate and would be the preferred method.

Another way to see this is to recognize that this problem is testing an understanding of the ideas of conditional probability. The question asks about the number of people in a car *given* that the car is in the company parking lot (the second method). If we start our sampling by looking at the person level (the first method) we will be counting people who may get to work by other means (like walk, ride a bicycle, etc.). As far as the number of people in each car in the parking lot is concerned we are not interested in these later people and they should not be polled.

Problem 28 (the 21st card)

Part (a): Let F be the event the 20th card is the first ace and let E be the event the 21st card is the ace of spades. For this part of the problem we want to compute $P(E|F)$. From the definition of conditional probability this can be written as

$$p(E|F) = \frac{p(EF)}{p(F)}\,.$$

Thus we can compute $P(E|F)$ if we can compute $P(F)$ and $P(EF)$. We begin by computing the value of $P(F)$. To compute this probability we will count the number of ways we can obtain the special card ordering denoted by event F and then divide this number by the number of ways we can have all 52 cards ordered with no restrictions on their ordering. This latter number is given by 52!. To compute the number of card ordering that given rise to event F consider that in selecting the first card we can select any card that is not an ace and thus have $52 - 4 = 48$ cards to select from. To select the second card we have one less or 47 cards to select from. Continuing this patter down to the 20th card we have

$$48 \cdot 47 \cdot 46 \cdots 32 \cdot 31 \cdot 30 \,.$$

ways to select the cards up to the 20th. For the 20th we have four choices (any one of the aces). After this card is selected we can select any card from the $52 - 20 = 32$ remaining cards for the 21st card. For the 22nd card we can select any of the 31 remaining cards. Thus the number of ways to select the remaining block of cards can be done in 32! ways. In total then we can compute $P(F)$ as

$$P(F) = \frac{(48 \cdot 47 \cdot 46 \cdots 32 \cdot 31 \cdot 30)4(32!)}{52!} = \frac{992}{54145} \,.$$

Next we need to compute $P(EF)$. Since the event EF is similar to the event F but with exception that the 20th card cannot be the ace of spaces (because the 21st card) the number of ways we can get the event EF is given by

$$(48 \cdot 47 \cdot 46 \cdots 32 \cdot 31 \cdot 30) \cdot 3 \cdot 1 \cdot (31!) \,.$$

Thus the probability $P(EF)$ is given by

$$P(EF) = \frac{(48 \cdot 47 \cdot 46 \cdots 32 \cdot 31 \cdot 30) \cdot 3 \cdot 1 \cdot (31!)}{52!} = \frac{93}{216580} \,.$$

Using these two results we compute

$$P(E|F) = \frac{992/54145}{93/216580} = \frac{3}{128} \,.$$

As an alternative method to compute these probabilities we can express the events E and F as boolean combinations of simpler component events A_i, where this component event describes whether the card at location i in the deck is an ace. The event F defined above represents the case where the the first 19 cards are not aces while the 20th card is and can be written in terms of these A_i events as

$$F = A_1^c \cdots A_{19}^c A_{20} \,.$$

With this product representation $P(F)$ can be computed by conditioning as

$$P(F) = P(A_1^c \cdots A_{19}^c A_{20}) = P(A_{20}|A_1^c \cdots A_{19}^c) P(A_1^c \cdots A_{19}^c) \,. \tag{7}$$

We can compute the probability the first 19 cards are not aces represented by the expression $P(A_1^c \cdots A_{19}^c)$ by further conditioning on earlier cards as

$$\begin{aligned}
P(A_1^c \cdots A_{19}^c) &= P(A_2^c A_3^c \cdots A_{19}^c | A_1^c) P(A_1^c) \\
&= P(A_3^c \cdots A_{19}^c | A_1^c A_2^c) P(A_2^c | A_1^c) P(A_1^c) \\
&= P(A_{19}^c | A_1^c A_2^c A_3^c \cdots A_{18}^c) \cdots P(A_3^c | A_1^c A_2^c) P(A_2^c | A_1^c) P(A_1^c) \,. \qquad (8)
\end{aligned}$$

We can now more easily evaluate these probabilities since

$$P(A_1^c) = \frac{48}{52}, \quad P(A_2^c|A_1^c) = \frac{47}{51} \quad \text{etc.}$$

Thus changing the order of the product in Equation 8 we find

$$\begin{aligned} P(A_1^c \cdots A_{19}^c) &= P(A_1^c)P(A_2^c|A_1^c)P(A_3^c|A_1^cA_2^c)\cdots P(A_{19}^c|A_1^cA_2^cA_3^c\cdots A_{18}^c) \\ &= \frac{48}{52}\cdot\frac{47}{51}\cdot\frac{46}{50}\cdots\frac{30}{34} = \frac{8184}{54145}\,. \end{aligned}$$

In the same way we have $P(A_{20}|A_1^c \cdots A_{19}^c) = \frac{4}{33}$ so that using Equation 7 we find

$$P(F) = \frac{8184}{54145}\cdot\frac{4}{33} = \frac{992}{54145}\,,$$

the same result we found earlier.

Next to compute $P(EF)$ we first introduce the event S to denote what type of ace the 20th is. To do that let S be the event that the 20th ace is the ace of spades. Since using S we have $A_{20} = S \cup S^c$ we can write the event EF as

$$EF = A_1^c \cdots A_{19}^c A_{20} E = A_1^c \cdots A_{19}^c SE \cup A_1^c \cdots A_{19}^c S^c E\,,$$

and have

$$P(EF) = P(A_1^c \cdots A_{19}^c SE) + P(A_1^c \cdots A_{19}^c S^c E)\,. \tag{9}$$

To evaluate each of these expressions we can condition like Equation 7 to get

$$\begin{aligned} P(A_1^c \cdots A_{19}^c SE) &= P(E|A_1^c \cdots A_{19}^c S)P(A_1^c \cdots A_{19}^c S) \quad \text{and} \\ P(A_1^c \cdots A_{19}^c S^c E) &= P(E|A_1^c \cdots A_{19}^c S^c)P(A_1^c \cdots A_{19}^c S^c)\,. \end{aligned}$$

Since S and E cannot both happen $P(E|A_1^c \cdots A_{19}^c S) = 0$ and in Equation 9 we are left with

$$\begin{aligned} P(EF) &= P(A_1^c \cdots A_{19}^c S^c E) = P(E|A_1^c \cdots A_{19}^c S^c)P(A_1^c \cdots A_{19}^c S^c) \\ &= P(E|A_1^c \cdots A_{19}^c S^c)P(S^c|A_1^c \cdots A_{19}^c)P(A_1^c \cdots A_{19}^c) \\ &= \frac{1}{32}\cdot\frac{3}{33}\cdot\frac{8184}{54145} = \frac{93}{216580}\,, \end{aligned}$$

the same result as earlier.

Part (b): As in the first method in Part (a) above for this part let F again be the event the 20th card is the first ace, but now let E be the event the 21st card is the 2 of clubs. As before we will solve this problem using definition of conditional probability or

$$p(E|F) = \frac{p(EF)}{p(F)}\,.$$

It remains to compute $p(EF)$ in this case since $P(F)$ is the same as previously. Since the event EF is similar to the event F but with exception that we now know the identity of the 21st card the number of ways we can get the event EF is given by

$$(47 \cdot 46 \cdot 45 \cdots 31 \cdot 30 \cdot 29) \cdot 4 \cdot 1 \cdot (31!)\,.$$

Thus the probability $P(EF)$ is given by

$$P(EF) = \frac{(47 \cdot 46 \cdot 45 \cdots 31 \cdot 30 \cdot 29) \cdot 4 \cdot 1 \cdot (31!)}{52!} = \frac{18}{52037}.$$

Using these two results we compute

$$P(E|F) = \frac{18/52037}{98/5349} = \frac{29}{1536}.$$

See the Matlab/Octave file `chap_3_prob_28.m` for the fractional simplifications needed in this problem.

Problem 29 (used tennis balls)

Let E_0, E_1, E_2, E_3 be the event that we select 0, 1, 2, or 3 used tennis balls during our first draw consisting of three balls. Then let A be the event that when we draw three balls the second time *none* of the selected balls have been used. The problem asks us to compute $P(A)$, which we can compute $P(A)$ by conditioning on the mutually exclusive events E_i for $i = 0, 1, 2, 3$ as

$$P(A) = \sum_{i=0}^{3} P(A|E_i)P(E_i).$$

Now we can compute the prior probabilities $P(E_i)$ as follows

$$\begin{aligned} P(E_0) &= \frac{\binom{6}{0}\binom{9}{3}}{\binom{15}{3}}, \quad P(E_1) = \frac{\binom{6}{1}\binom{9}{2}}{\binom{15}{3}} \\ P(E_2) &= \frac{\binom{6}{2}\binom{9}{1}}{\binom{15}{3}}, \quad P(E_3) = \frac{\binom{6}{3}\binom{9}{0}}{\binom{15}{3}}. \end{aligned}$$

Where the random variable representing the number of selected used tennis balls is a hypergeometric random variable and we have explicitly enumerated these probabilities above. We can now compute $P(A|E_i)$ for each i. Beginning with $P(A|E_0)$ which we recognize as the probability of event A under the situation where in the first draw of three balls we draw no used balls initially i.e. we draw all new balls. Since event E_0 is assumed to happen with certainty when we go to draw the second of three balls we have 6 new balls and 9 used balls. This gives the probability of event A as

$$P(A|E_0) = \frac{\binom{9}{0}\binom{6}{3}}{\binom{15}{3}}.$$

In the same way we can compute the other probabilities. We find that

$$P(A|E_1) = \frac{\binom{8}{0}\binom{7}{3}}{\binom{15}{3}}, \quad P(A|E_2) = \frac{\binom{7}{0}\binom{8}{3}}{\binom{15}{3}}, \quad P(A|E_3) = \frac{\binom{6}{0}\binom{9}{3}}{\binom{15}{3}}.$$

With these results we can calculate $P(A)$. This is done in the Matlab file `chap_3_prob_29.m` where we find that $P(A) \approx 0.0893$.

Problem 30 (boxes with marbles)

Let B be the event that the drawn ball is black and let X_1 (X_2) be the event that we select the first (second) box. Then to calculate $P(B)$ we will condition on the box drawn from as

$$P(B) = P(B|X_1)P(X_1) + P(B|X_2)P(X_2).$$

Now $P(B|X_1) = 1/2$, $P(B|X_2) = 2/3$, $P(X_1) = P(X_2) = 1/2$ so

$$P(B) = \frac{1}{2}\left(\frac{1}{2}\right) + \frac{1}{2}\left(\frac{2}{3}\right) = \frac{7}{12}.$$

If we see that the ball is white (i.e. it is not black i.e event B^c has happened) we now want to compute that it was drawn from the first box i.e.

$$P(X_1|B^c) = \frac{P(B^c|X_1)P(X_1)}{P(B^c|X_1)P(X_1) + P(B^c|X_2)P(X_2)} = \frac{3}{5}.$$

Problem 31 (Ms. Aquina's holiday)

After Ms. Aquina's tests are completed and the doctor has the results he will flip a coin. If it lands *heads* and the results of the tests are *good* he will call with the good news. If the results of the test are *bad* he will not call. If the coin flip lands *tails* he will not call regardless of the tests outcome. Lets let B denote the event that Ms. Aquina has cancer and the and the doctor has bad news. Let G be the event that Ms. Aquina does not have cancer and the results of the test are good. Finally let C be the event that the doctor calls the house during the holiday.

Part (a): Now the event that the doctor does not call (i.e. C^c) will add support to the hypothesis that Ms. Aquina has cancer (or event B) if and only if it is more likely that the doctor will not call given that she does have cancer. This is the event C^c will cause $\beta \equiv P(B|C^c)$ to be greater than $\alpha \equiv P(B)$ if and only if

$$P(C^c|B) \geq P(C^c|B^c) = P(C^c|G).$$

From a consideration of all possible outcomes we have that

$$P(C^c|B) = 1\,,$$

since if the results of the tests come back negative (and Ms. Aquina has cancer), the doctor will not call regardless of the coin flip. We also have that

$$P(C^c|G) = \frac{1}{2}\,,$$

since if the results of the test are good, the doctor will only call if the coin flip lands heads and not call otherwise. Thus the fact that the doctor does not call adds evidence to the belief that Ms. Aquina has cancer. Logic similar to this is discussed in the book after the example of the bridge championship controversy.

Part (b): We want to explicitly find $\beta = P(B|C^c)$ using Bayes' rule. We find that

$$\beta = \frac{P(C^c|B)P(B)}{P(C^c)} = \frac{1(\alpha)}{(3/4)} = \frac{4}{3}\alpha > \alpha\,.$$

Which explicitly verifies the intuition obtained in Part (a).

Problem 32 (the number of children)

Let C_1, C_2, C_3, C_4 be the events that the family has $1, 2, 3, 4$ children respectively. Let E be the evidence that the chosen child is the eldest in the family.

Part (a): We want to compute

$$P(C_1|E) = \frac{P(E|C_1)P(C_1)}{P(E)}\,.$$

We will begin by computing $P(E)$. We find that

$$P(E) = \sum_{i=1}^{4} P(E|C_i)P(C_i) = 1(0.1) + \frac{1}{2}(0.25) + \frac{1}{3}(0.35) + \frac{1}{4}(0.3) = 0.4167\,,$$

so that $P(C_1|E) = 1(0.1)/0.4167 = 0.24$.

Part (b): We want to compute

$$P(C_4|E) = \frac{P(E|C_4)P(C_4)}{P(E)} = \frac{(0.25)(0.3)}{0.4167} = 0.18\,.$$

These calculations are done in the file `chap_3_prob_32.m`.

Problem 33 (English vs. American)

Let E (A) be the event that this man is English (American). Also let L be the evidence found on the letter. Then we want to compute $P(E|L)$ which we will do with Bayes' rule. We find (counting the number of vowels in each word) that

$$\begin{aligned} P(E|L) &= \frac{P(L|E)P(E)}{P(L|E)P(E) + P(L|E^c)P(E^c)} \\ &= \frac{(3/6)(0.4)}{(3/6)(0.4) + (2/5)(0.6)} = \frac{5}{11} \,. \end{aligned}$$

Problem 34 (some new interpretation of the evidence)

From Example 3f in the book we had that

$$P(G|C) = \frac{P(GC)}{P(C)} = \frac{P(C|G)P(G)}{P(C|G)P(G) + P(C|G^c)P(G^c)} \,.$$

But now we are told $P(C|G) = 0.9$, since we are assuming that if we are guilty we will have the given characteristic with 90% certainty. Thus we now would compute for $P(G|C)$ the following

$$P(G|C) = \frac{0.9(0.6)}{0.9(0.6) + 0.2(0.4)} = \frac{27}{31} \,.$$

Problem 35 (which class is superior)

In this problem the superior class is the one that has the larger concentration of good students. An expert examines a student selected from class A and a student from class B. To formulate this problem in terms of probabilities lets introduce three events E, F, and R as follows. Let E be the event class A is the superior class, F be the event the expert finds the student from class A to be Fair, and R be the event the expert finds the student from class B to be Poor (P might have been a more intuitive notation to use for this last event but the letter P conflicts with the notation for probability). Using this notation for this problem we want to evaluate $P(E|FR)$. Using the definition of conditional probability we have

$$P(E|FR) = \frac{P(FR|E)P(E)}{P(FR)} = \frac{P(FR|E)P(E)}{P(FR|E)P(E) + P(FR|E^c)P(E^c)} \,.$$

To evaluate the above, first assume the events E and E^c are equally likely, that is $P(E) = P(E^c) = \frac{1}{2}$. This is reasonable since the labeling of A and B was done randomly and so the event that the label A was assigned to the superior class would happen with a probability of $\frac{1}{2}$. Next given E (that is A is the superior class) the two events F and R are conditionally independent. That is

$$P(FR|E) = P(F|E)P(R|E) \,,$$

and a similar expression when the event FR is conditioned on E^c. This states that given A is the superior class, a student selected from one class is Good, Fair, or Poor independent of a student selected from the other class being Good, Fair or Poor.

To evaluate these probabilities we reason as follows. If we are given the event E then A is the superior class and thus has 10 Fair students, so $P(F|E) = \frac{10}{30}$, while B is not the superior class and has 15 Poor students giving $P(R|E) = \frac{15}{30}$. If we are given E^c then A is not the superior class so $P(F|E^c) = \frac{5}{30}$ and $P(R|E^c) = \frac{10}{30}$. Using all of these results we have

$$\begin{aligned} P(E|FR) &= \frac{P(F|E)P(R|E)P(E)}{P(F|E)P(R|E)P(E) + P(F|E^c)P(R|E^c)P(E^c)} \\ &= \frac{P(F|E)P(R|E)}{P(F|E)P(R|E) + P(F|E^c)P(R|E^c)} \\ &= \frac{(10/30)(15/30)}{(10/30)(15/30) + (10/30)(5/30)} = \frac{3}{4} \,. \end{aligned}$$

Problem 36 (resignations from store C)

To solve this problem lets begin by introducing several events. Let A be the event a person works for company A, B be the event a person works for company B, and C be the event a person works for company C. Finally let W be the event a person is female (a woman). We desire to find $P(C|W)$. Using the definition of conditional probability we have

$$P(C|W) = \frac{P(CW)}{P(W)} = \frac{P(W|C)P(C)}{P(W|A)P(A) + P(W|B)P(B) + P(W|C)P(C)} \,. \tag{10}$$

Since 50, 75, and 100 people work for companies A, B, and C, respectively the total number of workers is $50 + 75 + 100 = 225$ and the individual probabilities of A, B, or C is given by

$$P(A) = \frac{50}{225} = \frac{2}{9} \,, \quad P(B) = \frac{75}{225} = \frac{1}{3} \,, \quad \text{and} \quad P(C) = \frac{100}{225} = \frac{4}{9} \,.$$

We are also told that .5, .6, and .7 are the percentages of the female employees of the companies A, B, C, respectively. Thus

$$P(W|A) = 0.5 \,, \quad P(W|B) = 0.6 \,, \quad \text{and} \quad P(W|C) = 0.7 \,.$$

Using these results in Equation 10 we get

$$P(C|W) = \frac{(0.7)(4/9)}{(0.5)(2/9) + (0.6)(1/3) + (0.7)(4/9)} = \frac{1}{2} \,.$$

See the Matlab/Octave file `chap_3_prob_36.m` for the fractional simplifications needed in this problem.

Problem 37 (gambling with a fair coin)

Let F denote the event that the gambler is observing results from a fair coin. Also let O_1, O_2, and O_3 denote the three observations made during our experiment. We will assume that before any observations are made the probability that we have selected the fair coin is 1/2.

Part (a): We desire to compute $P(F|O_1)$ or the probability we are looking at a fair coin given the first observation. This can be computed using Bayes' theorem. We have

$$\begin{aligned} P(F|O_1) &= \frac{P(O_1|F)P(F)}{P(O_1|F)P(F) + P(O_1|F^c)P(F^c)} \\ &= \frac{\frac{1}{2}\left(\frac{1}{2}\right)}{\frac{1}{2}\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right)} = \frac{1}{3} . \end{aligned}$$

Part (b): With the second observation and using the "posteriori's become priors" during a recursive update we now have

$$\begin{aligned} P(F|O_2, O_1) &= \frac{P(O_2|F, O_1)P(F|O_1)}{P(O_2|F, O_1)P(F|O_1) + P(O_2|F^c, O_1)P(F^c|O_1)} \\ &= \frac{\frac{1}{2}\left(\frac{1}{3}\right)}{\frac{1}{2}\left(\frac{1}{3}\right) + 1\left(\frac{2}{3}\right)} = \frac{1}{5} . \end{aligned}$$

Part (c): In this case because the two-headed coin cannot land tails we can immediately conclude that we have selected the fair coin. This result can also be obtained using Bayes' theorem as we have in the other two parts of this problem. Specifically we have

$$\begin{aligned} P(F|O_3, O_2, O_1) &= \frac{P(O_3|F, O_2, O_1)P(F|O_2, O_1)}{P(O_3|F, O_2, O_1)P(F|O_2, O_1) + P(O_3|F^c, O_2, O_1)P(F^c|O_2, O_1)} \\ &= \frac{\frac{1}{2}\left(\frac{1}{5}\right)}{\frac{1}{2}\left(\frac{1}{5}\right) + 0} = 1 . \end{aligned}$$

Verifying what we know must be true.

Problem 38 (drawing white balls)

Let W and B represent the events of drawing a white ball or a black respectively, and let H and T denote the event of obtaining a head or a tail when we flip the coin. As stated in the problem when the outcome of the coin flip is heads (event H) a ball is selected from urn A. This urn has 5 white and 7 black balls. Thus $P(W|H) = \frac{5}{12}$. Similarly, when the coin flip results in tails a ball is selected from urn B, which has 3 white and 12 black balls. Thus $P(W|T) = \frac{3}{15}$. We would like to compute $P(T|W)$. Using Bayes' formula we have

$$\begin{aligned} P(T|W) &= \frac{P(W|T)P(T)}{P(W)} = \frac{P(W|T)P(T)}{P(W|T)P(T) + P(W|H)P(H)} \\ &= \frac{\frac{3}{15}\left(\frac{1}{2}\right)}{\frac{3}{15}\left(\frac{1}{2}\right) + \frac{5}{12}\left(\frac{1}{2}\right)} = \frac{12}{37} . \end{aligned}$$

Problem 39 (having accidents)

From example 3a in the book where A_1 is the event that a person has an accident during the first year we recall that $P(A_1) = 0.26$. In this problem we are asked to find $P(A_2|A_1^c)$. We can find this probability by conditioning on whether or not the person is accident prone (event A). We have

$$P(A_2|A_1^c) = \frac{P(A_2A_1^c)}{P(A_1^c)} = \frac{P(A_2A_1^c|A)P(A) + P(A_2A_1^c|A^c)P(A^c)}{P(A_1^c)} .$$

We assume that A_2 and A_1 are conditionally independent given A and thus have

$$P(A_2A_1^c|A) = P(A_2|A)P(A_1^c|A) \quad \text{and} \quad P(A_2A_1^c|A^c) = P(A_2|A^c)P(A_1^c|A^c) . \qquad (11)$$

With these simplifications and using the numbers from example 3a we can evaluate $P(A_2|A_1^c)$. We thus find

$$\begin{aligned} P(A_2|A_1^c) &= \frac{P(A_2|A)P(A_1^c|A)P(A) + P(A_2|A^c)P(A_1^c|A^c)P(A^c)}{P(A_1^c)} \\ &= \frac{0.4(1-0.4)(0.3) + 0.2(1-0.2)(0.7)}{1-0.26} = \frac{46}{185} . \end{aligned}$$

Note that instead of assuming conditional independence to simplify probabilities such as $P(A_2A_1^c|A)$ appearing in Equations 11 we could also simply condition on *earlier* events by writing this expression as $P(A_2|A_1^c, A)P(A_1^c|A)$. The numerical values used to evaluate this expression would be the same as presented above.

Problem 40 (selecting k white balls)

For this problem we draw balls from an urn that starts with 5 white and 7 red balls and on each draw we replace each drawn ball with one of the same color as the one drawn. Then to solve the requested problem let W_k denote the event that a white ball was selected during the kth draw and R_k denote the even that a red ball was selected on the kth draw for $k = 1, 2, 3$. We then can decompose each of the higher level events (the number of white balls) in terms of the component events W_k and R_k as follows

Part (a): To get 0 white balls requires the event $R_1R_2R_3$. To compute this probability we use conditioning to find

$$\begin{aligned} P(R_1R_2R_3) &= P(R_1)P(R_2R_3|R_1) = P(R_1)P(R_2|R_1)P(R_3|R_1R_2) \\ &= \frac{7}{12} \cdot \frac{8}{13} \cdot \frac{9}{14} = \frac{3}{13} . \end{aligned}$$

Part (b): We can represent drawing only 1 white ball by the following event

$$W_1R_2R_3 \cup R_1W_2R_3 \cup R_1R_2W_3 .$$

As in Part (a) by conditioning we have that the probability of the above event is given by

$$\begin{aligned} P(W_1R_2R_3 \cup R_1W_2R_3 \cup R_1R_2W_3) &= P(W_1R_2R_3) + P(R_1W_2R_3) + P(R_1R_2W_3) \\ &= P(W_1)P(R_2|W_1)P(R_3|W_1R_2) \\ &+ P(R_1)P(W_2|R_1)P(R_3|R_1W_2) \\ &+ P(R_1)P(R_2|R_1)P(W_3|R_1R_2) \\ &= \frac{5}{12} \cdot \frac{7}{13} \cdot \frac{8}{14} + \frac{7}{12} \cdot \frac{5}{13} \cdot \frac{8}{14} + \frac{7}{12} \cdot \frac{8}{13} \cdot \frac{5}{14} \\ &= \frac{5}{13}. \end{aligned}$$

Part (c): We can draw 3 white balls in only one way $W_1W_2W_3$. Using the above logic as in Part (a) we have that the probability of this event given by

$$\begin{aligned} P(W_1W_2W_3) &= P(W_1)P(W_2|W_1)P(W_3|W_1W_2) \\ &= \frac{5}{12}\frac{6}{13}\frac{7}{14} = \frac{5}{52}. \end{aligned}$$

Part (d): We can draw two white balls in the following way

$$R_1W_2W_3 \cup W_1R_2W_3 \cup W_1W_2R_3.$$

Again, using the same logic as in Part (a) we have that the probability of the above event given by

$$\begin{aligned} P(R_1W_2W_3 \cup W_1R_2W_3 \cup W_1W_2R_3) &= P(R_1W_2W_3) + P(W_1R_2W_3) + P(W_1W_2R_3) \\ &= P(R_1)P(W_2|R_1)P(W_3|R_1W_2) \\ &+ P(W_1)P(R_2|W_1)P(W_3|W_1R_2) \\ &+ P(W_1)P(W_2|W_1)P(R_3|W_1W_2) \\ &= \frac{7}{12} \cdot \frac{5}{13} \cdot \frac{6}{14} + \frac{5}{12} \cdot \frac{7}{13} \cdot \frac{6}{14} + \frac{5}{12} \cdot \frac{6}{13} \cdot \frac{7}{14} \\ &= \frac{15}{52}. \end{aligned}$$

Problem 41 (drawing the same ace)

We want to compute if the second card drawn is an ace. Denoting this event by E and following the hint lets compute $P(E)$ by conditioning on whether we select the original ace drawn from the first deck. Let this event by A_0. Then we have

$$P(E) = P(E|A_0)P(A_0) + P(E|A_0^c)P(A_0^c).$$

Now $P(A_0) = \frac{1}{27}$ since it is one of the twenty seven cards in this second stack and $P(A_0^c) = \frac{26}{27}$ and $P(E|A_0) = 1$. Using these values we get

$$P(E) = 1 \cdot \frac{1}{27} + \frac{26}{27} \cdot P(E|A_0^c).$$

Thus it remains to compute the expression $P(E|A_0^c)$. Since under the event A_0^c we know that we do not draw the original ace, this probability is related to how the original deck of cards was split. In that case the current half deck could have $3, 2, 1$ or 0 aces in it. For each of these cases we have probabilities given by $\frac{3}{26}$, $\frac{2}{26}$, $\frac{1}{26}$, and $\frac{0}{26}$ of drawing an ace if we have three aces, two aces, one ace, and no aces respectively in our second half deck. Conditioning on the number of aces in this half deck we have (using D_3, D_2 and D_1 as notation for the events that this half deck has 3, 2 or 1 aces in it) we obtain

$$P(E|A_0^c) = P(E|D_3, A_0^c)P(D_3|A_0^c) + P(E|D_2, A_0^c)P(D_2|A_0^c) + P(E|D_1, A_0^c)P(D_1|A_0^c)\,.$$

Since one of the aces was found to be in the first pile, the second pile contains $k = 1, 2, 3$ aces with probability

$$P(D_k) = \frac{\binom{3}{k}\binom{48}{26-k}}{\binom{51}{26}} \quad \text{for} \quad k = 1, 2, 3\,,$$

Evaluating the above expression these numbers become

$$\begin{aligned}
P(E|D_3, A_0^c) &= \frac{\binom{3}{3}\binom{52-4}{26-3}}{\binom{51}{26}} = \frac{104}{833} \\
P(E|D_2, A_0^c) &= \frac{\binom{3}{2}\binom{52-4}{26-2}}{\binom{51}{26}} = \frac{325}{833} \\
P(E|D_1, A_0^c) &= \frac{\binom{3}{1}\binom{52-4}{26-1}}{\binom{51}{26}} = \frac{314}{833}\,.
\end{aligned}$$

We can now evaluate $P(E|A_0^c)$ the probability of selecting an ace, given that it must be one of the original 26 cards in the second pile, as

$$\begin{aligned}
P(E|A_0^c) &= \sum_{k=1}^{3} P(ED_k|A_0^c) = \sum_{k=1}^{3} P(D_k|A_0^c)P(E|D_k, A_0^c) \\
&= \sum_{k=1}^{3} \frac{k}{26}\left(\frac{\binom{3}{k}\binom{48}{26-k}}{\binom{51}{26}}\right) = \frac{1}{17}\,,
\end{aligned}$$

when we perform the required summation. We can also reason that $P(E|A_0^c) = \frac{1}{17}$ in another way. This probability is equivalent to the case where we have simply removed one ace from the deck and recognized that the second card drawn could be any one of the remaining 51 cards (three of which remain as aces). Thinking like this would give $P(E|A_0^c) = \frac{3}{51} = \frac{1}{17}$ the same result as argued above. The fact that cards are in separate piles is irrelevant. Any one of the 51 cards could be in any position in either pile.

We can now finally evaluate $P(E)$ we have

$$P(E) = 1 \cdot \frac{1}{27} + \frac{26}{27} \cdot \frac{1}{17} = \frac{43}{459},$$

the same as in the back of the book. See the Matlab file `chap_3_prob_41.m` for these calculations.

Problem 42 (special cakes)

Let R be the event that the special cake will rise correctly. Then from the problem statement we are told that $P(R|A) = 0.98$, $P(R|B) = 0.97$, and $P(R|C) = 0.95$, with the prior information of $P(A) = 0.5$, $P(B) = 0.3$, and $P(C) = 0.2$. Then this problem asks for $P(A|R^c)$. Using Bayes' rule we have

$$P(A|R^c) = \frac{P(R^c|A)P(A)}{P(R^c)},$$

where $P(R^c)$ is given by conditioning on A, B, or C as

$$\begin{aligned} P(R^c) &= P(R^c|A)P(A) + P(R^c|B)P(B) + P(R^c|C)P(C) \\ &= 0.02(0.5) + 0.03(0.3) + 0.05(0.2) = 0.029, \end{aligned}$$

so that $P(A|R^c)$ is given by

$$P(A|R^c) = \frac{0.02(0.5)}{0.029} = 0.344.$$

Problem 43 (three coins in a box)

Let C_1, C_2, C_3 be the event that the first, second, and third coin is chosen and flipped. Then let H be the event that the flipped coin showed heads. Then we would like to evaluate $P(C_1|H)$. Using Bayes' rule we have

$$P(C_1|H) = \frac{P(H|C_1)P(C_1)}{P(H)}.$$

We compute $P(H)$ first. We find conditioning on the the coin selected that

$$\begin{aligned} P(H) &= \sum_{i=1}^{3} P(H|C_i)P(C_i) = \frac{1}{3}\sum_{i=1}^{3} P(H|C_i) \\ &= \frac{1}{3}\left(1 + \frac{1}{2} + \frac{3}{4}\right) = \frac{3}{4}. \end{aligned}$$

Then $P(C_1|H)$ is given by

$$P(C_1|H) = \frac{1(1/3)}{(3/4)} = \frac{4}{9}.$$

Problem 44 (a prisoners' dilemma)

I will argue that this problem is similar to the so called "Monty Hall Problem" and because of this connection the probability of execution of the prisoner *stays* at 1/3 instead of 1/2. See [1] for a nice discussion of the Monty Hall Problem. The probabilities that do change, however, are the probabilities of the *other* two prisoners. The probability of execution of the prisoner to be set free falls to 0 while the probability of the other prisoner increases to 2/3.

To show the similarity of this problem to the Monty Hall Problem, think of the three prisoners as surrogates for Monty Hall's three doors. Think of execution as the "prize" that is hidden "behind" one of the prisoners. Finally, the prisoner that the guard admits to freeing is equivalent to Monty Hall opening up a door in which he knows does not contain this "prize". In the Monty Hall Problem the initial probabilities associated with each door are 1/3 but once a non-selected door has been opened the probability of having selected the correct door *does not* increase from 1/3 to 1/2. The opening of the other door is irrelevant to your odds of winning if you keep your selection. The remaining door has a probability of 2/3 of containing the prize.

Following the analogy the jailer revealing a prisoner that can go free is Monty opening a door known not to contain the prize. By symmetry to Monty Hall, A's probability of being executed must remain at 1/3 and not increase to 1/2.

A common error in logic is to argue as follows. Before asking his question the probability of event A (A is to be executed) is $P(A) = 1/3$. If prisoner A is told that B (or C) is to be set free then we need to compute $P(A|B^c)$. Where A, B, and C are the events that prisoner A, B, or C is to be executed respectively. Now from Bayes' rule

$$P(A|B^c) = \frac{P(B^c|A)P(A)}{P(B^c)}.$$

We have that $P(B^c)$ is given by

$$P(B^c) = P(B^c|A)P(A) + P(B^c|B)P(B) + P(B^c|C)P(C) = \frac{1}{3} + 0 + \frac{1}{3} = \frac{2}{3}.$$

So the above probability then becomes

$$P(A|B^c) = \frac{1(1/3)}{2/3} = \frac{1}{2} > \frac{1}{3}.$$

Thus the probability that prisoner A will be executed has increased as claimed by the jailer.

While there is nothing wrong with above logic, the problem with it is that it is not answering the real question that we want the answer to. This question is: what is the probability that A will be executed given the statement that the jailer makes. Now the jailer has only two choices for what he can say; either B or C will be set free. Lets compute $P(A|J_B)$ where J_B is the event that the jailer says that prisoner B will be set free. We expect by symmetry that $P(A|J_B) = P(A|J_C)$. We then have using Bayes' rule

$$\begin{aligned}
P(A|J_B) &= \frac{P(J_B|A)P(A)}{P(J_B)} \\
&= \frac{P(J_B|A)P(A)}{P(J_B|A)P(A) + P(J_B|B)P(B) + P(J_B|C)P(C)} \\
&= \frac{(1/2)(1/3)}{(1/2)(1/3) + 0 + 1(1/3)} \\
&= \frac{1}{3}.
\end{aligned}$$

In performing this computation we have used the facts that

$$P(J_B|A) = \frac{1}{2},$$

since the jailer has two choices he can say if A is to be executed,

$$P(J_B|B) = 0,$$

since the jailer cannot say that B will be set free if B is to be executed, and

$$P(J_B|C) = 1,$$

since in this case prisoner C is to be executed so the jailer cannot say C and being unable to say that prisoner A is to be set free he must say that prisoner B is to be set free. Thus we see that regardless to what the jailer says the probability that A is to executed *stays the same.*

Problem 45 (is it the fifth coin?)

Let C_i be the event that the ith coin was selected to be flipped. Since any coin is equally likely we have $P(C_i) = \frac{1}{10}$ for all i. Let H be the event that the flipped coin shows heads, then we want to compute $P(C_5|H)$. From Bayes' rule we have

$$P(C_5|H) = \frac{P(H|C_5)P(C_5)}{P(H)}.$$

We compute $P(H)$ by conditioning on the selected coin C_i we have

$$\begin{aligned}
P(H) &= \sum_{i=1}^{10} P(H|C_i)P(C_i) \\
&= \sum_{i=1}^{10} \frac{i}{10}\left(\frac{1}{10}\right) = \frac{1}{100}\sum_{i=1}^{10} i \\
&= \frac{1}{100}\left(\frac{10(10+1)}{2}\right) = \frac{11}{20}.
\end{aligned}$$

So that

$$P(C_5|H) = \frac{(5/10)(1/10)}{(11/20)} = \frac{1}{11}.$$

Problem 46 (one accident means its more likely that you will have another)

Consider the expression $P(A_2|A_1)$. By the definition of conditional probability this can be expressed as

$$P(A_2|A_1) = \frac{P(A_1, A_2)}{P(A_1)},$$

so the desired expression to show is then equivalent to the following

$$\frac{P(A_1, A_2)}{P(A_1)} > P(A_1),$$

or $P(A_1, A_2) > P(A_1)^2$. Considering first the expression $P(A_1)$ by conditioning on the sex of the policy holder we have

$$P(A_1) = P(A_1|M)P(M) + P(A_1|W)P(W) = p_m\alpha + p_f(1-\alpha).$$

where M is the event the policy holder is male and W is the event that the policy holder is female. In the same way we have for the joint probability $P(A_1, A_2)$ that

$$P(A_1, A_2) = P(A_1, A_2|M)P(M) + P(A_1, A_2|W)P(W).$$

Assuming that A_1 and A_2 are *independent* given the specification of the policy holders sex we have that

$$P(A_1, A_2|M) = P(A_1|M)P(A_2|M),$$

the same expression holds for the event W. Using this in the expression for $P(A_1, A_2)$ above we obtain

$$\begin{aligned} P(A_1, A_2) &= P(A_1|M)P(A_2|M)P(M) + P(A_1|W)P(A_2|W)P(W) \\ &= p_m^2\alpha + p_f^2(1-\alpha). \end{aligned}$$

We now look to see if $P(A_1, A_2) > P(A_1)^2$. Computing the expression $P(A_1, A_2) - P(A_1)^2$, (which we hope to be able to show is always positive) we have that

$$\begin{aligned} P(A_1, A_2) - P(A_1)^2 &= p_m^2\alpha + p_f^2(1-\alpha) - (p_m\alpha + p_f(1-\alpha))^2 \\ &= p_m^2\alpha + p_f^2(1-\alpha) - p_m^2\alpha^2 - 2p_mp_f\alpha(1-\alpha) - p_f^2(1-\alpha)^2 \\ &= p_m^2\alpha(1-\alpha) + p_f^2(1-\alpha)\alpha - 2p_mp_f\alpha(1-\alpha) \\ &= \alpha(1-\alpha)(p_m^2 + p_f^2 - 2p_mp_f) \\ &= \alpha(1-\alpha)(p_m - p_f)^2. \end{aligned}$$

Note that this is always positive. Thus we have shown that $P(A_1|A_2) > P(A_1)$. In words, this means that given that we have an accident in the first year this information will increase the probability that we will have an accident in the second year to a value greater than we would have without the knowledge of the accident during year one (A_1).

Problem 47 (the probability on which die was rolled)

Let X be the the random variable that specifies the number on the die roll i.e. the integer $1, 2, 3, \cdots, 6$. Let W be the event that all the balls drawn are white. Then we want to evaluate $P(W)$, which can be computed by conditioning on the value of X. Thus we have

$$P(W) = \sum_{i=1}^{6} P\{W|X = i\}P(X = i)$$

Since $P\{X = i\} = 1/6$ for every i, we need only to compute $P\{W|X = i\}$. We have that

$$\begin{aligned}
P\{W|X = 1\} &= \frac{5}{15} \approx 0.33 \\
P\{W|X = 2\} &= \left(\frac{5}{15}\right)\left(\frac{4}{14}\right) \approx 0.095 \\
P\{W|X = 3\} &= \left(\frac{5}{15}\right)\left(\frac{4}{14}\right)\left(\frac{3}{13}\right) \approx 0.022 \\
P\{W|X = 4\} &= \left(\frac{5}{15}\right)\left(\frac{4}{14}\right)\left(\frac{3}{13}\right)\left(\frac{2}{12}\right) \approx 0.0036 \\
P\{W|X = 5\} &= \left(\frac{5}{15}\right)\left(\frac{4}{14}\right)\left(\frac{3}{13}\right)\left(\frac{2}{12}\right)\left(\frac{1}{11}\right) \approx 0.0003 \\
P\{W|X = 6\} &= 0
\end{aligned}$$

Then we have

$$P(W) = \frac{1}{6}(0.33 + 0.95 + 0.022 + 0.0036 + 0.0003) = 0.0756\,.$$

If all the balls selected are white then the probability our die showed a three was

$$P\{X = 3|W\} = \frac{P\{W|X = 3\}P(X = 3)}{P(W)} = 0.048\,.$$

Problem 48 (which cabinet did we select)

This question is the same as asking what is the probability we select cabinet A given that a silver coin is seen on our draw. Then we want to compute $P(A|S) = \frac{P(S|A)P(A)}{P(S)}$. Now

$$P(S) = P(S|A)P(A) + P(S|B)P(B) = 1\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{3}{4}$$

Thus

$$P(A|S) = \frac{1(1/2)}{(3/4)} = \frac{2}{3}\,.$$

Problem 49 (prostate cancer)

Let C be the event that man has cancer and A (for antigen) the event of taking an elevated PSA measurement. Then in the problem we are given

$$\begin{aligned} P(A|C^c) &= 0.135 \\ P(A|C) &= 0.268\,, \end{aligned}$$

and in addition we have $P(C) = 0.7$.

Part (a): We want to evaluate $P(C|A)$ or

$$\begin{aligned} P(C|A) &= \frac{P(A|C)P(C)}{P(A)} \\ &= \frac{P(A|C)P(C)}{P(A|C)P(C) + P(A|C^c)P(C^c)} \\ &= \frac{(0.268)(0.7)}{(0.268)(0.7) + (0.135)(0.3)} = 0.822\,. \end{aligned}$$

Part (b): We want to evaluate $P(C|A^c)$ or

$$\begin{aligned} P(C|A^c) &= \frac{P(A^c|C)P(C)}{P(A^c)} \\ &= \frac{(1 - 0.268)(0.7)}{1 - 0.228} = 0.6637\,. \end{aligned}$$

If the prior probability of cancer changes (i.e. $P(C) = 0.3$) then the above formulas yield

$$\begin{aligned} P(C|A) &= 0.459 \\ P(C|A^c) &= 0.266\,. \end{aligned}$$

Problem 50 (assigning probabilities of risk)

Let G, A, B be the events that a person is of good risk, an average risk, or a bad risk respectively. Then in the problem we are told that (if E denotes the event that an accident occurs)

$$\begin{aligned} P(E|G) &= 0.05 \\ P(E|A) &= 0.15 \\ P(E|B) &= 0.3 \end{aligned}$$

In addition the a priori assumptions on the proportion of people that are good, average and bad risks are given by $P(G) = 0.2$, $P(A) = 0.5$, and $P(B) = 0.3$. Then in this problem we are asked to compute $P(E)$ or the probability that an accident will happen. This can be computed by conditioning on the probability of a person having an accident from among the three types, i.e.

$$\begin{aligned} P(E) &= P(E|G)P(G) + P(E|A)P(A) + P(E|B)P(B) \\ &= 0.05(0.2) + (0.15)(0.5) + (0.3)(0.3) = 0.175\,. \end{aligned}$$

If a person had no accident in a given year we want to compute $P(G|E^c)$ or

$$\begin{aligned} P(G|E^c) &= \frac{P(E^c|G)P(G)}{P(E^c)} = \frac{(1-P(E|G))P(G)}{1-P(E)} \\ &= \frac{(1-0.05)(0.2)}{1-0.175} = \frac{38}{165} \end{aligned}$$

also to compute $P(A|E^c)$ we have

$$\begin{aligned} P(A|E^c) &= \frac{P(E^c|A)P(A)}{P(E^c)} = \frac{(1-P(E|A))P(A)}{1-P(E)} \\ &= \frac{(1-0.15)(0.5)}{1-0.175} = \frac{17}{33} \end{aligned}$$

Problem 51 (letters of recommendation)

Let R_s, R_m, and R_w be the event that our worker receives a strong, moderate, or weak recommendation respectively. Let J be the event that our applicant gets the job. Then the problem specifies

$$\begin{aligned} P(J|R_s) &= 0.8 \\ P(J|R_m) &= 0.4 \\ P(J|R_w) &= 0.1\,, \end{aligned}$$

with priors on the type of recommendation given by

$$\begin{aligned} P(R_s) &= 0.7 \\ P(R_m) &= 0.2 \\ P(R_w) &= 0.1\,, \end{aligned}$$

Part (a): We are asked to compute $P(J)$ which by conditioning on the type of recommendation received is

$$\begin{aligned} P(J) &= P(J|R_s)P(R_s) + P(J|R_m)P(R_m) + P(J|R_w)P(R_w) \\ &= 0.8(0.7) + (0.4)(0.2) + (0.1)(0.1) = 0.65 = \frac{13}{20}\,. \end{aligned}$$

Part (b): Given the event J is held true then we are asked to compute the following

$$\begin{aligned} P(R_s|J) &= \frac{P(J|R_s)P(R_s)}{P(J)} = \frac{(0.8)(0.7)}{(0.65)} = \frac{56}{65} \\ P(R_m|J) &= \frac{P(J|R_m)P(R_m)}{P(J)} = \frac{(0.4)(0.2)}{(0.65)} = \frac{8}{65} \\ P(R_w|J) &= \frac{P(J|R_w)P(R_w)}{P(J)} = \frac{(0.1)(0.1)}{(0.65)} = \frac{1}{65} \end{aligned}$$

Note that this last probability can also be calculated as $P(R_w|J) = 1 - P(R_w|J) - P(R_w|J)$.

Part (c): For this we are asked to compute

$$\begin{aligned} P(R_s|J^c) &= \frac{P(J^c|R_s)P(R_s)}{P(J^c)} = \frac{(1-0.8)(0.7)}{(0.35)} = \frac{2}{5} \\ P(R_m|J^c) &= \frac{P(J^c|R_m)P(R_m)}{P(J^c)} = \frac{(1-0.4)(0.2)}{(0.35)} = \frac{12}{35} \\ P(R_w|J^c) &= \frac{P(J^c|R_w)P(R_w)}{P(J^c)} = \frac{(1-0.1)(0.1)}{(0.35)} = \frac{9}{35}\,. \end{aligned}$$

Problem 52 (college acceptance)

Let M, T, W, R, F, and S correspond to the events that mail comes on Monday, Tuesday, Wednesday, Thursday, Friday, or Saturday (or later) respectively. Let A be the event that our student is accepted.

Part (a): To compute $P(M)$ we can condition on whether or not the student is accepted as

$$P(M) = P(M|A)P(A) + P(M|A^c)P(A^c) = 0.15(0.6) + 0.05(0.4) = 0.11\,.$$

Part (b): We desire to compute $P(T|M^c)$. Using the definition of conditional probability we find that (again conditioning $P(T)$ on whether she is accepted or not)

$$\begin{aligned} P(T|M^c) &= \frac{P(T,M^c)}{P(M^c)} = \frac{P(T)}{1-P(M)} \\ &= \frac{P(T|A)P(A) + P(T|A^c)P(A^c)}{1-P(M)} \\ &= \frac{0.2(0.6)+0.1(0.4)}{1-0.11} = \frac{16}{89}\,. \end{aligned}$$

Part (c): We want to calculate $P(A|M^c, T^c, W^c)$. Again using the definition of conditional probability (twice) we have that

$$P(A|M^c,T^c,W^c) = \frac{P(A,M^c,T^c,W^c)}{P(M^c,T^c,W^c)} = \frac{P(M^c,T^c,W^c|A)P(A)}{P(M^c,T^c,W^c)}\,.$$

To evaluate terms like $P(M^c, T^c, W^c|A)$, and $P(M^c, T^c, W^c|A^c)$, lets compute the probability that mail will come on Saturday or later given that she is accepted or not. Using the fact that $P(\cdot|A)$ and $P(\cdot|A^c)$ are both probability densities and must sum to one over their first argument we calculate that

$$\begin{aligned} P(S|A) &= 1 - 0.15 - 0.2 - 0.25 - 0.15 - 0.1 = 0.15 \\ P(S|A^c) &= 1 - 0.05 - 0.1 - 0.1 - 0.15 - 0.2 = 0.4\,. \end{aligned}$$

With this result we can calculate that

$$\begin{aligned} P(M^c, T^c, W^c|A) &= P(R|A) + P(F|A) + P(S|A) = 0.15 + 0.1 + 0.15 = 0.4 \\ P(M^c, T^c, W^c|A^c) &= P(R|A^c) + P(F|A^c) + P(S|A^c) = 0.15 + 0.2 + 0.4 = 0.75\,. \end{aligned}$$

Also we can compute $P(M^c, T^c, W^c)$ by conditioning on whether she is accepted or not. We find

$$\begin{aligned} P(M^c, T^c, W^c) &= P(M^c, T^c, W^c|A)P(A) + P(M^c, T^c, W^c|A^c)P(A^c) \\ &= 0.4(0.6) + 0.75(0.4) = 0.54\,. \end{aligned}$$

Now we finally have all of the components we need to compute what we were asked to. We find that

$$P(A|M^c, T^c, W^c) = \frac{P(M^c, T^c, W^c|A)P(A)}{P(M^c, T^c, W^c)} = \frac{0.4(0.6)}{0.54} = \frac{4}{9}\,.$$

Part (d): We are asked to compute $P(A|R)$ which using Bayes' rule gives

$$P(A|R) = \frac{P(R|A)P(A)}{P(R)}\,.$$

To compute this lets begin by computing $P(R)$ again obtained by conditioning on whether our student is accepted or not. We find

$$P(R) = P(R|A)P(A) + P(R|A^c)P(A^c) = 0.15(0.6) + 0.15(0.4) = 0.15\,.$$

So that our desired probability is given by

$$P(A|R) = \frac{0.15(0.6)}{0.15} = \frac{3}{5}\,.$$

Part (e): We want to calculate $P(A|S)$. Using Bayes' rule gives

$$P(A|S) = \frac{P(S|A)P(A)}{P(S)}\,.$$

To compute this, lets begin by computing $P(S)$ again obtained by conditioning on whether our student is accepted or not. We find

$$P(S) = P(S|A)P(A) + P(S|A^c)P(A^c) = 0.15(0.6) + 0.4(0.4) = 0.25\,.$$

So that our desired probability is given by

$$P(A|S) = \frac{0.15(0.6)}{0.25} = \frac{9}{25}\,.$$

Problem 53 (the functioning of a parallel system)

With n components a parallel system will be working if at least one component is working. Let H_i be the event that the component i for $i = 1, 2, 3, \cdots, n$ is working. Let F be the event that the entire system is functioning. We want to compute $P(H_1|F)$. We have

$$P(H_1|F) = \frac{P(F|H_1)P(H_1)}{P(F)}.$$

Now $P(F|H_1) = 1$ since if the first component is working the system is functioning. In addition, $P(F) = 1 - \left(\frac{1}{2}\right)^n$ since to be *not* functioning all components must not be working. Finally $P(H_1) = 1/2$. Thus our probability is

$$P(H_1|F) = \frac{1/2}{1 - (1/2)^n}.$$

Problem 54 (independence of E and F)

Part (a): These two events would be independent. The fact that one person has blue eyes and another unrelated person has blue eyes are in no way related.

Part (b): These two events seem unrelated to each other and would be modeled as independent.

Part (c): As height and weigh are related, I would think that these two events are not independent.

Part (d): Since the United States is in the western hemisphere these two two events are related and they are not independent.

Part (e): Since rain one day would change the probability of rain on other days I would say that these events are related and therefore not independent.

Problem 55 (independence in class)

Let S be a random variable denoting the sex of the randomly selected person. The S can take on the values m for male and f for female. Let C be a random variable representing denoting the class of the chosen student. The C can take on the values f for freshman and s for sophomore. We want to select the number of sophomore girls such that the random variables S and C are independent. Let n denote the number of sophomore girls. Then counting up the number of students that satisfy each requirement we have

$$\begin{aligned} P(S=m) &= \frac{10}{16+n} \\ P(S=f) &= \frac{6+n}{16+n} \\ P(C=f) &= \frac{10}{16+n} \\ P(C=s) &= \frac{6+n}{16+n}. \end{aligned}$$

The joint density can also be computed and are given by

$$\begin{aligned} P(S=m, C=f) &= \frac{4}{16+n} \\ P(S=m, C=s) &= \frac{6}{16+n} \\ P(S=f, C=f) &= \frac{6}{16+n} \\ P(S=f, C=s) &= \frac{n}{16+n}. \end{aligned}$$

Then to be independent we must have $P(C,S) = P(S)P(C)$ for all possible C and S values. Considering the point case where $(S = m, C = f)$ we have that n must satisfy

$$\begin{aligned} P(S=m, C=f) &= P(S=m)P(C=f) \\ \frac{4}{16+n} &= \left(\frac{10}{16+n}\right)\left(\frac{10}{16+n}\right) \end{aligned}$$

which when we solve for n gives $n = 9$. Now one should check that this value of n works for all other equalities that must be true, for example one needs to check that when $n = 9$ the following are true

$$\begin{aligned} P(S=m, C=s) &= P(S=m)P(C=s) \\ P(S=f, C=f) &= P(S=f)P(C=f) \\ P(S=f, C=s) &= P(S=f)P(C=s). \end{aligned}$$

As these can be shown to be true, $n = 9$ is the correct answer.

Problem 56 (is the nth coupon new?)

Let C_i be the identity of the n coupon and A_i the event that after collecting $n-1$ coupons at least one coupon of type i exists. Then the event that the nth coupon is new if we obtain one of type i is the event $A_i^c \cap C_i$, which is the the event that the ith coupon is not in the first $n-1$ coupons i.e. A_i^c and that the nth coupon is not the ith one. Then if E_i is the event that the ith coupon is new we have

$$P(E_i) = P(A_i^c \cap C_i) = P(A_i^c | C_i)P(C_i) = P(A_i^c)p_i = (1 - P(A_i))p_i\,,$$

where from Example 4 i from the book we have that $P(A_i) = 1-(1-p_i)^{n-1}$ so the probability of a new coupon being of type i is $(1-(1-p_i)^{n-1})p_i$, so the probability of a new coupon (at all) is given by

$$P(E) = \sum_{i=1}^{m} P(E_i) = \sum_{i=1}^{m} (1-(1-p_i)^{n-1})p_i \,.$$

Problem 57 (the price path of a stock)

For this problem it helps to draw a diagram of the stocks path v.s. time for the various situations.

Part (a): To be the same price in two days the stock can go up and then back down or down and then back up. Giving a total probability of $2p(1-p)$.

Part (b): To go up only one unit in three steps we must go up twice and down once. We can have the single down day happen on any of the three days. Thus, the three possible paths are (with $+1$ denoting a day where the stock goes up and -1 denoting a day where the stock goes down) given by

$$(+1,+1,-1)\,, \quad (+1,-1,+1)\,, \quad (-1,+1,+1)\,,$$

each with probability $p^2(1-p)$. Thus since each path is mutually exclusive we have a total probability of $3p^2(1-p)$.

Part (c): When we count the number of paths where we go up on the first day (two) and divide by the total number of paths (three) we get the probability $\frac{2}{3}$.

Problem 58 (generating fair flips with a biased coin)

Part (a): Consider pairs of flips. Let E be the event that a pair of flips returns (H,T) and let F be the event that the pair of flips returns (T,H). From the discussion on Page 93 Example 4h the event E will occur first with probability

$$\frac{P(E)}{P(E)+P(F)}\,.$$

Now $P(E) = p(1-p)$ and $P(F) = (1-p)p$, so the probability of obtaining event E and declaring *tails* before the event F (from which we would declare *heads*) would be

$$\frac{p(1-p)}{2p(1-p)} = \frac{1}{2}\,.$$

In the same way we will have the event F occur before the event E with probability $\frac{1}{2}$. Thus we have an equally likely chance of obtaining heads or tails. Note: its important to note that the procedure described is effectively working with ordered pairs of flips, we flip two coins and only make a decision after looking at both coins and the order in which they come out.

Part (b): Lets compute the probability of declaring heads under this procedure. Assume we are considering a sequence of coin flips. Let H be the event that we declare a head. Then conditioning on the outcome of the previous two flips we have with P_f and C_f random variables denoting the previous and the current flip respectively that

$$\begin{aligned} P(H) &= P(H|P_f = T, C_f = T)P\{P_f = T, C_f = T\} \\ &+ P(H|P_f = T, C_f = H)P\{P_f = T, C_f = H\} \\ &+ P(H|P_f = H, C_f = T)P\{P_f = H, C_f = T\} \\ &+ P(H|P_f = H, C_f = H)P\{P_f = H, C_f = H\}\,. \end{aligned}$$

Now since

$$\begin{aligned} P\{P_f = T, C_f = T\} &= 0\,, \quad P\{P_f = T, C_f = H\} = 1 \\ P\{P_f = H, C_f = T\} &= 0\,, \quad P\{P_f = H, C_f = H\} = 0\,. \end{aligned}$$

we see that $P(H) = P(P_f = T, C_f = H) = (1-p)p \neq \frac{1}{2}$. In the same way $P(T) = p(1-p)$. Thus this procedure would produce a head or a tail with equal probability but this probability would not be 1/2.

Problem 59 (the first four outcomes)

Part (a): This probability would be p^4.

Part (b): This probability would be $(1-p)p^3$.

Part (c): Given two mutually exclusive events E and F the probability that E occurs before F is given by

$$\frac{P(E)}{P(E)+P(F)}\,.$$

Denoting E by the event that we obtain a T, H, H, H pattern and F the event that we obtain a H, H, H, H pattern the above becomes

$$\frac{p^3(1-p)}{p^4 + p^3(1-p)} = \frac{1-p}{p+(1-p)} = 1-p\,.$$

Problem 60 (the color of your eyes)

Since Smith's sister has blue eyes and this is a recessive trait, both of Smith's parents must have the gene for blue eyes. Let R denote the gene for brown eyes and L denote the gene for blue eyes (these are the second letters in the words brown and blue respectively). Then

Smith will have a gene makeup possibly given by $(R,R),(R,L),(L,R)$, where the left gene is the one received from his mother and the right gene is the one received from his father.

Part (a): With the gene makeup given above we see that in two cases from three total Smith will have a blue gene. Thus this probability is 2/3.

Part (b): Since Smith's wife has blue eyes, Smith's child will receive a L gene from his mother. The probability Smith's first child will have blue eyes is then dependent on what gene they receive from Smith. Letting B be the event that Smith's first child has blue eyes (and conditioning on the possible genes Smith could give his child) we have

$$P(B) = 0\left(\frac{1}{3}\right) + \frac{1}{2}\left(\frac{1}{3}\right) + \frac{1}{2}\left(\frac{1}{3}\right) = \frac{1}{3}.$$

As stated above, this result is obtained by conditioning on the possible gene makeups of Smith. For example let (X,Y) be the notation for the "event" that Smith has a gene make up given by (X,Y) then the above can be written symbolically (in terms of events) as

$$P(B) = P(B|(R,R))P(R,R) + P(B|(R,L))P(R,L) + P(B|(L,R))P(L,R).$$

Evaluating each of the above probabilities gives the result already stated.

Part (c): The fact that the first child has brown eyes makes it more likely that Smith has a genotype, given by (R,R). We compute the probability of this genotype given the event E (the event that the first child has brown eyes using Bayes' rule as)

$$\begin{aligned}
P((R,R)|E) &= \frac{P(E|(R,R))P(R,R)}{P(E|(R,R))P(R,R) + P(E|(R,L))P(R,L) + P(E|(L,R))P(L,R)} \\
&= \frac{1\left(\frac{1}{3}\right)}{1\left(\frac{1}{3}\right) + \frac{1}{2}\left(\frac{1}{3}\right) + \frac{1}{2}\left(\frac{1}{3}\right)} \\
&= \frac{1}{2}.
\end{aligned}$$

In the same way we have for the other possible genotypes that

$$P((R,L)|E) = \frac{\frac{1}{2}\left(\frac{1}{3}\right)}{\frac{2}{3}} = \frac{1}{4} = P((L,R)|E).$$

Thus the same calculation as in Part (b), but now conditioning on the fact that the first child has brown eyes (event E) gives for a probability of the event B_2 (that the second child we have *blue* eyes)

$$\begin{aligned}
P(B_2|E) &= P(B_2|(R,R),E)P((R,R)|E) + P(B_2|(R,L),E)P((R,L)|E) + P(B_2|(L,R),E)P((L,R)|E) \\
&= 0\left(\frac{1}{2}\right) + \frac{1}{2}\left(\frac{1}{4}\right) + \frac{1}{2}\left(\frac{1}{4}\right) = \frac{1}{4}.
\end{aligned}$$

This means that the probability that the second child has *brown* eyes is then

$$1 - P(B_2|E) = \frac{3}{4}.$$

Problem 61 (more recessive traits)

From the information that the two parents are normal but that they produced an albino child we know that both parents must be carriers of albinism. Their non-albino child can have any of three possible genotypes each with probability 1/3 given by $(A, A), (A, a), (a, A)$. Lets denote this parent by P_1 and the event that this parent is a carrier for albinism as C_1. Note that $P(C_1) = 2/3$ and $P(C_1^c) = 1/3$. We are told that the spouse of this person (denoted P_2) is a carrier for albinism.

Part (a): The probability their first offspring is an albino depends on how likely our first parent is a carrier of albinism. We have (with E_1 the event that their first child is an albino) that

$$P(E_1) = P(E_1|C_1)P(C_1) + P(E_1|C_1^c)P(C_1^c)\,.$$

Now $P(E_1|C_1) = \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4}$, since both parents must contribute their albino gene, and $P(E_1|C_1^c) = 0$ so we have that

$$P(E_1) = \frac{1}{4}\left(\frac{2}{3}\right) = \frac{1}{6}\,.$$

Part (b): The fact that the first newborn is not an albino changes the probability that the first parent is a carrier or the value of $P(C_1)$. To calculate this we will use Bayes' rule

$$\begin{aligned} P(C_1|E_1^c) &= \frac{P(E_1^c|C_1)P(C_1)}{P(E_1^c|C_1)P(C_1) + P(E_1^c|C_1^c)P(C_1^c)} \\ &= \frac{\frac{3}{4}\left(\frac{2}{3}\right)}{\frac{3}{4}\left(\frac{2}{3}\right) + 1\left(\frac{1}{3}\right)} \\ &= \frac{3}{5}\,. \end{aligned}$$

so we have that $P(C_1^c|E_1^c) = \frac{2}{5}$, and following the steps in Part (a) we have (with E_2 the event that the couples second child is an albino)

$$\begin{aligned} P(E_2|E_1^c) &= P(E_2|E_1^c, C_1)P(C_1|E_1^c) + P(E_2|E_1^c, C_1^c)P(C_1^c|E_1^c) \\ &= \frac{1}{4}\left(\frac{3}{5}\right) = \frac{3}{20}\,. \end{aligned}$$

Problem 62 (target shooting with Barbara and Dianne)

Let H be the event that the duck is "hit", by either Barbra or Dianne's shot. Let B and D be the events that Barbra (respectively Dianne) hit the target. Then the outcome of the experiment where both Dianne and Barbra fire at the target (assuming that their shots work independently is)

$$\begin{aligned} P(B^c, D^c) &= (1 - p_1)(1 - p_2) \\ P(B^c, D) &= (1 - p_1)p_2 \\ P(B, D^c) &= p_1(1 - p_2) \\ P(B, D) &= p_1 p_2\,. \end{aligned}$$

Part (a): We desire to compute $P(B, D|H)$ which equals

$$P(B, D|H) = \frac{P(B, D, H)}{P(H)} = \frac{P(B, D)}{P(H)}$$

Now $P(H) = (1 - p_1)p_2 + p_1(1 - p_2) + p_1 p_2$ so the above probability becomes

$$\frac{p_1 p_2}{(1 - p_1)p_2 + p_1(1 - p_2) + p_1 p_2} = \frac{p_1 p_2}{p_1 + p_2 - p_1 p_2}.$$

Part (b): We desire to compute $P(B|H)$ which equals

$$P(B|H) = P(B, D|H) + P(B, D^c|H).$$

Since the first term $P(B, D|H)$ has already been computed we only need to compute $P(B, D^c|H)$. As before we find it to be

$$P(B, D^c|H) = \frac{p_1(1 - p_2)}{(1 - p_1)p_2 + p_1(1 - p_2) + p_1 p_2}.$$

So the total result becomes

$$P(B|H) = \frac{p_1 p_2 + p_1(1 - p_2)}{(1 - p_1)p_2 + p_1(1 - p_2) + p_1 p_2} = \frac{p_1}{p_1 + p_2 - p_1 p_2}.$$

Problem 63 (dueling)

For a given trial while dueling we have the following possible outcomes (events) and their associated probabilities

- Event I: A is hit and B is not hit. This happens with probability $p_B(1 - p_A)$.
- Event II: A is not hit and B is hit. This happens with probability $p_A(1 - p_B)$.
- Event III: A is hit and B is hit. This happens with probability $p_A p_B$.
- Event IV: A is not hit and B is not hit. This happens with probability $(1-p_A)(1-p_B)$.

With these definitions we can compute the probabilities of various other events.

Part (a): To solve this we recognize that A is hit if events I and III happen and the dueling continues if event IV happens. We can compute $p(A)$ (the probability that A is hit) by conditioning on the outcome of the first duel. We have

$$p(A) = p(A|I)p(I) + p(A|II)p(II) + p(A|III)p(III) + p(A|IV)p(IV).$$

Now in the case of event IV the dual continues afresh and we see that $p(A|IV) = p(A)$. Using this fact and the definitions of events I-IV we have that the above becomes

$$p(A) = 1 \cdot p_B(1 - p_A) + 0 \cdot p_A(1 - p_B) + 1 \cdot p_A p_B + p(A) \cdot (1 - p_A)(1 - p_B).$$

Now solving for $p(A)$ in the above we find that

$$p(A) = \frac{p_B}{(1 - (1 - p_A)(1 - p_B))}.$$

Part (b): Let D be the event that both duelists are hit. Then to compute this, we can condition on the outcome of the first dual. Using the same arguments as above we find

$$\begin{aligned} p(D) &= p(D|I)p(I) + p(D|II)p(II) + p(D|III)p(III) + p(D|IV)p(IV) \\ &= 0 + 0 + 1 \cdot p_A p_B + p(D) \cdot (1 - p_A)(1 - p_B). \end{aligned}$$

On solving for $P(D)$ we have

$$p(D) = \frac{p_A p_B}{1 - (1 - p_A)(1 - p_B)}.$$

Part (c): Lets begin by computing the probability that the dual ends after one dual. Let G_1 be the event that the game ends with *more than* (or after) one dual. We have, conditioning on the events *I*-*IV* that

$$p(G_1) = 0 + 0 + 0 + 1 \cdot (1 - p_A)(1 - p_B) = (1 - p_A)(1 - p_B).$$

Now let G_2 be the event that the game ends with *more than* (or after) two duals. Then

$$p(G_2) = (1 - p_A)(1 - p_B)p(G_1) = (1 - p_A)^2(1 - p_B)^2.$$

Generalizing this result we have for the probability that the games ends after n duels is

$$p(G_n) = (1 - p_A)^n(1 - p_B)^n.$$

Part (d): Let G_1 be the event that the game ends with more than one dual and let A be the event that A is hit. Then to compute $p(G_1|A^c)$ by conditioning on the first experiment we have

$$\begin{aligned} p(G_1|A^c) &= p(G_1, I|A^c)p(I) + p(G_1, II|A^c)p(II) \\ &+ p(G_1, III|A^c)p(III) + p(G_1, IV|A^c)p(IV) \\ &= 0 + 0 + 0 + p(G_1, IV|A^c)(1 - p_A)(1 - p_B). \end{aligned}$$

So now we need to evaluate $p(G_1, IV|A^c)$, which we do using the definition of conditional probability. We find

$$p(G_1, IV|A^c) = \frac{p(G_1, IV, A^c)}{p(A^c)} = \frac{1}{p(A^c)}.$$

Where $p(A^c)$ is the probability that A is not hit *on the first experiment.* This can be computed as

$$\begin{aligned} p(A) &= p_B(1 - p_A) + p_A p_B = p_B \quad \text{so} \\ p(A^c) &= 1 - p_B, \end{aligned}$$

	Woman answers correctly	Woman answers incorrectly
Man answers correctly	p^2	$p(1-p)$
Man answers incorrectly	$(1-p)p$	$(1-p)^2$

Table 7: The possible probabilities of agreement for the couple in Problem 64, Chapter 3. When asked a question four possible outcomes can occur, corresponding to the correctness of the mans (woman's) answer. The first row corresponds to the times when the husband answers the question correctly, the second row to the times when the husband answers the question incorrectly. In the same way, the first column corresponds to the times when the wife is correct and second column to the times when the wife is incorrect.

and the above is then given by

$$p(G_1|A^c) = \frac{(1-p_A)(1-p_B)}{1-p_B} = 1-p_A\,.$$

In the same way as before this would generalize to the following (for the event G_n)

$$p(G_n) = (1-p_A)^n(1-p_B)^{n-1}$$

Part (e): Let AB be the event that both duelists are hit. Then in the same way as Part (d) above we see that

$$p(G_1, IV|AB) = \frac{p(G_1, IV, AB)}{p(AB)} = \frac{1}{p(AB)}\,.$$

Here $p(AB)$ is the probability that A and B are hit on any given experiment so $p(AB) = p_A p_B$, and

$$p(G_1|AB) = \frac{(1-p_A)(1-p_B)}{p_A p_B}$$

and in general

$$p(G_n|AB) = \frac{(1-p_A)^n(1-p_B)^n}{p_A p_B}\,.$$

Problem 64 (game show strategies)

Part (a): Since each person has probability p of getting the correct answer, either one selected to represent the couple will answer correctly with probability p.

Part (b): To compute the probability that the couple answers correctly under this strategy we will condition our probability on the "agreement" matrix in Table 7, i.e. the possible combinations of outcomes the couple may encounter when asked a question that they both answer. Lets define E be the event that the couple answers correctly, and let C_m (C_w) be the events that the man (women) answers the question correctly. We find that

$$\begin{aligned} P(E) &= P(E|C_m, C_w)P(C_m, C_w) + P(E|C_m, C_w^c)P(C_m, C_w^c) \\ &+ P(E|C_m^c, C_w)P(C_m^c, C_w) + P(E|C_m^c, C_w^c)P(C_m^c, C_w^c)\,. \end{aligned}$$

Now $P(E|C_m^c, C_w^c) = 0$ since both the man and the woman agree but they both answer the question incorrectly. In that case the couple would return the incorrect answer to the question. In the same way we have that $P(E|C_m, C_w) = 1$. Following the strategy of flipping a coin when the couple answers disagree we note that $P(E|C_m, C_w^c) = P(E|C_m^c, C_w) = 1/2$, so that the above probability when using this strategy becomes

$$P(E) = 1 \cdot p^2 + \frac{1}{2}p(1-p) + \frac{1}{2}(1-p)p = p\,,$$

where in computing this result we have used the joint probabilities found in Table 7 to evaluate terms like $P(C_m, C_w^c)$. Note that this result is the same as in Part (a) of this problem showing that there is no benefit to using this strategy.

Problem 65 (how accurate are we when we agree/disagree)

Part (a): We want to compute (using the notation from the previous problem)

$$P(E|(C_m, C_w) \cup (C_m^c, C_w^c))\,.$$

Defining the event A to be equal to $(C_m, C_w) \cup (C_m^c, C_w^c)$. We see that this is equal to

$$P(E|(C_m, C_w) \cup (C_m^c, C_w^c)) = \frac{P(E, A)}{P(A)} = \frac{p^2}{p^2 + (1-p)^2} = \frac{0.36}{0.36 + 0.16} = \frac{9}{13}\,.$$

Part (b): We want to compute $P(E|(C_m^c, C_w) \cup (C_m, C_w^c))$, but in the second strategy above if the couple disagrees they flip a fair coin to decide. Thus this probability is equal to $1/2$.

Problem 66 (relay circuits)

Part (a): Let E be the event that current flows from A to B. Then

$$\begin{aligned} P(E) &= P(E|5\,\text{Closed})p_5 \\ &= p(1 \text{ and } 2 \text{ closed } \textit{or } 3 \textit{ and } 4 \text{ closed}|5\,\text{closed})p_5 \\ &= (p_1 p_2 + p_3 p_4)p_5\,. \end{aligned}$$

Part (b): Conditioning on relay 3. Let C_i be the event the ith relay is closed. Then

$$\begin{aligned} P(E) &= P(E|C_3)P(C_3) + P(E|C_3^c)P(C_3^c) \\ &= (p_1 p_4 + p_1 p_5 + p_2 p_5 + p_2 p_4)p_3 + (p_1 p_4 + p_2 p_5)(1 - p_3)\,. \end{aligned}$$

Both of these can be checked by considering the entire joint distribution and eliminating combinations that don't allow current to flow. For example for Part (a) we have (conditioned on switch five being closed) that

$$\begin{aligned}
1 &= p_1p_2p_3p_4 + (1-p_1)p_2p_3p_4 + p_1(1-p_2)p_3p_4 + p_1p_2(1-p_3)p_4 \\
&+ p_1p_2p_3(1-p_4) + (1-p_1)(1-p_2)p_3p_4 + (1-p_1)p_2(1-p_3)p_4 \\
&+ (1-p_1)p_2p_3(1-p_4) + p_1(1-p_2)(1-p_3)p_4 + p_1(1-p_2)p_3(1-p_4) \\
&+ p_1p_2(1-p_3)(1-p_4) + (1-p_1)(1-p_2)(1-p_3)p_4 + (1-p_1)(1-p_2)p_3(1-p_4) \\
&+ (1-p_1)p_2(1-p_3)(1-p_4) + p_1(1-p_2)(1-p_3)(1-p_4) \\
&+ (1-p_1)(1-p_2)(1-p_3)(1-p_4)\,.
\end{aligned}$$

This explicit enumeration is possible is possible because these are Bernoulli random variables (they can be *on* or *off*) and thus there are $2^{|c|}$ total elements in the joint distribution. From the above enumeration we find (eliminating the non-functioning combinations) that

$$\begin{aligned}
P(E|C_5) &= p_1p_2p_3p_4 + (1-p_1)p_2p_3p_4 + p_1(1-p_2)p_3p_4 \\
&+ p_1p_2(1-p_3)p_4 + (1-p_1)(1-p_2)p_3p_4 + p_1p_2(1-p_3)(1-p_4) \\
&= p_2p_3p_4 + (1-p_2)p_3p_4 + p_1p_2(1-p_3) \\
&= p_3p_4 + p_1p_2(1-p_3)\,.
\end{aligned}$$

Problem 67 (k-out-of-n systems)

Part (a): We must have two or more of the four components functioning so we can have (with E the event that we have a functioning system) that

$$\begin{aligned}
P(E) &= p_1p_2p_3p_4 + (1-p_1)p_2p_3p_4 + p_1(1-p_2)p_3p_4 + p_1p_2(1-p_3)p_4 \\
&+ p_1p_2p_3(1-p_4) + (1-p_1)(1-p_2)p_3p_4 + (1-p_1)p_2(1-p_3)p_4 \\
&+ (1-p_1)p_2p_3(1-p_4) + p_1(1-p_2)(1-p_3)p_4 + p_1(1-p_2)p_3(1-p_4) \\
&+ p_1p_2(1-p_3)(1-p_4)\,.
\end{aligned}$$

Problem 68 (is the relay open?)

For this problem let C_i be the event the ith relay is *open* and let E be the event current flows from A to B. We can write the event E in terms of the events C_i as

$$E = (C_1C_2 \cup C_3C_4)C_5\,.$$

The probability we want to evaluate is $P((C_1C_2)^c|E) = 1 - P(C_1C_2|E)$. Now from the definition of conditional probability we have

$$P(C_1C_2|E) = \frac{P(C_1C_2E)}{P(E)}\,.$$

Consider C_1C_2E as an "set" using the expression for E above it can be written as

$$C_1C_2E = C_1C_2C_5 \cup C_1C_2C_3C_4C_5 = C_1C_2C_3C_4C_5\,,$$

P_1	P_2	$C = P_1$	$C = P_2$
a,a	a,a	1	1
a,a	a,A	1/2	1/2
a,a	A,A	0	0
a,A	a,a	1/2	1/2
a,A	a,A	1/2	1/2
a,A	A,A	1/2	1/2
A,A	a,a	0	0
A,A	a,A	1/2	1/2
A,A	A,A	1	1

Table 8: The probability of various matching *genotypes*, for the child denoted by C and the two parents P_1 and P_2. The notations $C = P_1$ and $C = P_2$ means the the child's genotype matches that of the first and second parent respectively.

since $C_1C_2C_5$ is a larger set than $C_1C_2C_3C_4C_5$ in other words $C_1C_2C_5 \supset C_1C_2C_3C_4C_5$. Thus

$$P(C_1C_2E) = P(C_1C_2C_3C_4C_5) = p_1p_2p_3p_4p_5 \,.$$

Next using the relationship between the probability of the union of events we have

$$\begin{aligned} P(E) &= P((C_1C_2 \cup C_3C_4)C_5) = p_5 P(C_1C_2 \cup C_3C_4) \\ &= p_5(P(C_1C_2) + P(C_3C_4) - P(C_1C_2C_3C_4)) \\ &= p_5(p_1p_2 + p_3p_4 - p_1p_2p_3p_4) \,. \end{aligned}$$

Using these two expressions we have

$$P(C_1C_2|E) = \frac{p_1p_2p_3p_4}{p_1p_2 + p_3p_4 - p_1p_2p_3p_4} \,,$$

so that the desired probability is given by

$$P((C_1C_2)^c|E) = 1 - P(C_1C_2|E) = \frac{p_1p_2 + p_3p_4 - 2p_1p_2p_3p_4}{p_1p_2 + p_3p_4 - p_1p_2p_3p_4} \,.$$

Problem 69 (genotypes and phenotypes)

For this problem lets first consider the simplified case where we compute the probability that a child receives various genotypes/phenotypes when crossing one single gene pair from each parent. In Table 8 we list the probability of a child having *genotypes* that match the first parent and second parent. These probabilities can be computed by considering the possible genes that a given parent can give to his/her offspring. An example of this type of subcalculation that goes into producing the entries in the second row of Table 8 is given in Table 9. In this table we see the first parent has the gene pair aa and the second parent has the gene pair Aa (or aA since they are equivalent). See the table caption for more details. In the same way in Table 10 we list the probability of a child having *phenotypes* that match (or not) the two parents. As each pair of genes is independent of the others by using the two Tables 8 and 10 we can now answer the questions for this problem.

	A	a
a	aA	aa
a	aA	aa

Table 9: Example of the potential genotypes (and phenotypes) of the offspring produced from mating the first parent with an aa genotype and a second parent with an Aa genotype. We see that in 2 of 4 possible cases we get the gene pair aA and in 2 of the 4 possible cases we get the gene pair aa. This gives the probability of 1/2 for either genotype aa or aA, and probabilities of 1/2 for the recessive phenotype a and the dominant phenotype A.

Part (a): Using Table 8 and independence we see that to get a child's genotype that matches the first parent will happen with probability of

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{2^5}=\frac{1}{32}\,.$$

Using Table 10 and independence we see that to get a child's phenotype that matches the first parent will happen with probability of

$$\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{1}{2}=\frac{9}{2^7}=\frac{9}{128}\,.$$

Part (b): Using the Tables 8 and 10 we have the probability that a child's genotype and phenotype matches the second parent will happen with

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{2^5}=\frac{1}{32}\,,$$

and

$$\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{1}{2}=\frac{9}{2^7}=\frac{9}{128}\,.$$

Part (c): To match genotypes with either parent we must exactly match the first or the second parent. Thus using the results from Part (a) and (b) above we get the probability of matching either parents genotype is given by

$$\frac{1}{32}+\frac{1}{32}=\frac{1}{16}\,.$$

In the same way to match phenotypes means that our phenotype must match the first or the second child. Thus again using the results from Part (a) and (b) above we have the probability of matching either parents phenotype is given by

$$\frac{9}{128}+\frac{9}{128}=\frac{9}{64}\,.$$

Part (d): If we desire the probability we *don't* match either parent, this is the complement of the probability we do match one of the parents. Thus using the result from Part (c) above we have that the probability of matching neither parents genotype is given by

$$1-\frac{1}{16}=\frac{15}{16}\,,$$

P_1	P_2	$C = P_1$	$C = P_2$
a,a	a,a	1	1
a,a	a,A	1/2	1/2
a,a	A,A	0	1
a,A	a,a	1/2	1/2
a,A	a,A	3/4	3/4
a,A	A,A	1	1
A,A	a,a	1	0
A,A	a,A	1	1
A,A	A,A	1	1

Table 10: The probability of various matching *phenotypes*, for the child denoted by C and the two parents P_1 and P_2. The notations here match the same ones given in Table 8 but is for phenotypes rather than genotypes.

and phenotype is given by

$$1 - \frac{9}{64} = \frac{55}{64}.$$

Problem 70 (hemophilia and the queen)

Let C be the event that the queen is a carrier of the gene for hemophilia. We are told that $P(C) = 0.5$. Let H_i be the event that the i-th prince has hemophilia. The we observe the event $H_1^c H_2^c H_3^c$ and we want to compute $P(C|H_1^c H_2^c H_3^c)$. Using Bayes' rule we have that

$$P(C|H_1^c H_2^c H_3^c) = \frac{P(H_1^c H_2^c H_3^c|C)P(C)}{P(H_1^c H_2^c H_3^c|C)P(C) + P(H_1^c H_2^c H_3^c|C^c)P(C^c)}.$$

Now

$$P(H_1^c H_2^c H_3^c|C) = P(H_1^c|C)P(H_2^c|C)P(H_3^c|C).$$

By the independence of the birth of the princes. Now $P(H_i^c|C) = 0.5$ so that the above is given by

$$P(H_1^c H_2^c H_3^c|C) = (0.5)^3 = \frac{1}{8}.$$

Also $P(H_1^c H_2^c H_3^c|C^c) = 1$ so the above probability becomes

$$P(C|H_1^c H_2^c H_3^c) = \frac{(0.5)^3(0.5)}{(0.5)^3(0.5) + 1(0.5)} = \frac{1}{9}.$$

In the next part of this problem (below) we will need the complement of this probability or

$$P(C^c|H_1^c H_2^c H_3^c) = 1 - P(C|H_1^c H_2^c H_3^c) = \frac{8}{9}.$$

If the queen has a fourth prince, then we want to compute $P(H_4|H_1^c H_2^c H_3^c)$. Let A be the event $H_1^c H_2^c H_3^c$ (so that we don't have to keep writing this) then conditioning on whether

Probability	Win	Loss	Total Wins	Total Losses
$\left(\frac{1}{2}\right)^3 = \frac{1}{8}$	0	3	87	75
$3\left(\frac{1}{2}\right)^3 = \frac{3}{8}$	1	2	88	74
$3\left(\frac{1}{2}\right)^3 = \frac{3}{8}$	2	1	89	73
$\left(\frac{1}{2}\right)^3 = \frac{1}{8}$	3	0	90	72

Table 11: The win/loss record for the Atlanta Braves each of the four total possible outcomes when they play the San Diego Padres.

S.F.G. Total Wins	S.F.G. Total Losses	L.A.D. Total Wins	L.A.D. Total Losses
86	76	89	73
87	75	88	74
88	74	87	75
89	73	86	76

Table 12: The total win/loss record for both the San Francisco Giants (S.F.G) and the Los Angeles Dodgers (L.A.D.). The first row corresponds to the San Francisco Giants winning *no* games while the Los Angeles Dodgers win *three* games. The number of wins going to the San Francisco Giants increases as we move down the rows of the table, until we reach the third row where the Giants have won three games and the Dodgers none.

the queen is a carrier, we see that the probability we seek is given by

$$\begin{aligned} P(H_4|A) &= P(H_4|C,A)P(C|A) + P(H_4|C^c,A)P(C^c|A) \\ &= P(H_4|C)P(C|A) + P(H_4|C^c)P(C^c|A) \\ &= \frac{1}{2}\left(\frac{1}{9}\right) = \frac{1}{18}\,. \end{aligned}$$

Problem 71 (winning the western division)

We are asked to compute the probabilities that each of the given team wins the western division. We will assume that the team with the largest total number of wins will be the division winner. We are also told that each team is equally likely to win each game it plays. We can take this information to mean that each team wins each game it plays with probability 1/2. We begin to solve this problem, by considering the three games that the Atlanta Braves play against the San Diego Padres. In Table 11 we enumerate all of the possible outcomes, i.e. the total number of wins or losses that can occur to the Atlanta Braves during these three games, along with the probability that each occurs.

We can construct the same type of a table for the San Francisco Giants when they play the Los Angeles Dodgers. In Table 12 we list all of the possible total win/loss records for both the San Francisco Giants and the Los Angeles Dodgers. Since the probabilities are the same as listed in Table 11 the table does not explicitly enumerate these probabilities.

From these results (and assuming that the the team with the most wins will win the division)

	1/8	3/8	3/8	1/8
1/8	D	D	G	G
3/8	D	B/D	B/G	G
3/8	B/D	B	B	B/G
1/8	B	B	B	B

Table 13: The possible division winners depending on the outcome of the three games that each team must play. The rows (from top to bottom) correspond to the Atlanta Braves winning more and more games (from the three that they play). The columns (from left to right) correspond to the San Francisco Giants winning more and more games (from the three they play). Note that as the Giants win more games the Dodgers must loose more games. Ties are determined by the presence of two symbols at a given location.

we can construct a table which represents for each of the possible wins/losses combination above, which team will be the division winner. Define the events B, G, and D to be the events that the Braves, Giants, and Los Angles Dodgers win the western division. Then in Table 13 we summarize the results of the two tables above where for the first row assumes that the Atlanta Braves win *none* of their games and the last row assumes that the Atlanta Braves win *all* of their games. In the same way the first column corresponds to the case when the San Francisco Giants win *none* of their games and the last column corresponds to the case when they win *all* of their games.

In anytime that two teams tie each team has a 1/2 of a chance of winning the tie-breaking game that they play next. Using this result and the probabilities derived above we can evaluate the individual probabilities that each team wins. We find that

$$
\begin{aligned}
P(D) &= \frac{1}{8}\left(\frac{1}{8}+\frac{3}{8}\right)+\frac{3}{8}\left(\frac{1}{8}+\frac{1}{2}\cdot\frac{3}{8}\right)+\frac{3}{8}\left(\frac{1}{2}\cdot\frac{1}{8}\right)=\frac{13}{64} \\
P(G) &= \frac{1}{8}\left(\frac{3}{8}+\frac{1}{8}\right)+\frac{3}{8}\left(\frac{1}{2}\cdot\frac{3}{8}+\frac{1}{8}\right)+\frac{3}{8}\left(\frac{1}{2}\cdot\frac{1}{8}\right)=\frac{13}{64} \\
P(B) &= \frac{3}{8}\left(\frac{1}{2}\cdot\frac{3}{8}+\frac{1}{2}\cdot\frac{3}{8}\right)+\frac{3}{8}\left(\frac{1}{2}\cdot\frac{1}{8}+\frac{3}{8}+\frac{3}{8}+\frac{1}{2}\cdot\frac{1}{8}\right)+\frac{1}{8}(1)=\frac{19}{32}
\end{aligned}
$$

Note that these probabilities to add to one as they should. The calculations for this problems are `chap_3_prob_71.m`.

Problem 72 (town council vote)

We can solve this first part of the problem in two ways. In the first, we select a member of the steering committee (SC) to consider. Without loss of generality let this person be the first member of the steering committee or $i = 1$ for $i = 1, 2, 3$. Let V_i be the event that the ith person on the steering committee votes for the given piece of legislation. Then the event V_i^c is the event the i member votes against. Now all possible voting for a given piece of legislation are

$$
V_1V_2V_3\,,\, V_1^cV_2V_3\,,\, V_1V_2^cV_3\,,\, V_1V_2V_3^c\,,\, V_1V_2^cV_3^2\,,\, V_1^cV_2V_3^c\,,\, V_1^cV_2^cV_3\,,\, V_1^cV_2^cV_3^c\,.
$$

Since each event V_i is independent with probability p the probability of each of the events above can be easily calculated. From the above, in only the events

$$V_1 V_2^c V_3 \,,\, V_1 V_2 V_3^c \,,\, V_1^c V_2 V_3^c \,,\, V_1^c V_2^c V_3 \,,$$

will changing the vote of the $i = 1$ member change the total outcome. Summing the probability of these four events we find the probability we seek given by

$$p^2(1-p)+p^2(1-p)+p(1-p)^2+p(1-p)^2 = 2p^2(1-p)+2p(1-p)^2 = 2p(1-p)[p+1-p] = 2p(1-p)\,.$$

As a second way to work this problem let E be the event that the total vote outcome from the steering committee will be different if the selected member changes his vote. Lets compute $P(E)$ by conditioning on whether V, our member voted for the legislation, or V^c he did not. We have

$$P(E) = P(E|V)P(V) + P(E|V^c)P(V^c) = P(E|V)p + P(E|V^c)(1-p)\,.$$

Now to determine $P(E|V)$ the event that changing from a "yes" vote to a "no" vote will change the outcome of the total decision we note that in order for that to be true we need to have one "yes" vote and one "no" vote from the other members. In that case if we change from "yes" to "no" the legislation will be rejected. Having one "yes" and one "no" vote happens with probability

$$\binom{2}{1} p(1-p)\,.$$

Now to determine $P(E|V^c)$ we reason the same way. This is the event that changing from a "no" vote to a "yes" vote will change the outcome of the total decision. In order for that to be true, we need to have one "no" vote and one "yes" vote from the other members. In that case if we change from "no" to "yes" the legislation will be accepted. Having one "no" and one "yes" vote happens (again) with probability

$$\binom{2}{1} (1-p)p\,.$$

Thus, summing the two results above we find

$$P(E) = 2p^2(1-p) + 2(1-p)^2 p = 2p(1-p)\,,$$

the same as before.

When we move to the case with seven councilmen we assume that there is no guarantee that the members of the original steering committee will vote the same way as earlier. Then again to evaluate $P(E)$ we condition on V and we have

$$P(E) = P(E|V)p + P(E|V^c)(1-p)\,.$$

Evaluating $P(E|V)$, since there are 7 total people we need to have 4 total "yes" and 3 total "no" thus removing the fact that V has occurred we need 6 "yes" and 3 "no" for a probability of

$$P(E|V) = \binom{6}{3} p^3(1-p)^3\,.$$

The calculation of $P(E|V^c)$ is the same. We need 4 total "no" and 3 total "yes" thus removing the fact that V^c has occurred (we voted "no") we need 6 "no" and 3 "yes" for a probability of

$$P(E|V^c) = \binom{6}{3}(1-p)^3 p^3 .$$

Thus we find since $\binom{6}{3} = 20$ that

$$P(E) = 20p^4(1-p)^3 + 20p^3(1-p)^4 = 20p^3(1-p)^3 .$$

Problem 73 (5 children)

Part (a): To have all of the same type of children means that they are all girls or all boys and will happen with a probability

$$\left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 = \frac{1}{32} + \frac{1}{32} = \frac{1}{16} .$$

Part (b): To first have 3 boys and then 2 girls will happen with probability

$$\left(\frac{1}{2}\right)^5 = \frac{1}{32} .$$

Part (c): To have exactly 3 boys (independent of their ordering) will happen with probability

$$\binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^5 = \frac{10}{32} .$$

Part (d): To have the first 2 children be girls (independent of what the other children are) will happen with probability

$$\left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

Part (e): To have at least one girl (not all boys) is the complement of the even of having no girls, or of having all boys. Thus

$$1 - \left(\frac{1}{2}\right)^5 = 1 - \frac{1}{32} = \frac{31}{32} .$$

Problem 74 (our dice sum to 9 or 6)

From Equation 2 we see that the probability player A rolls a 9 is given by $p_A = \frac{1}{9}$ and the probability player B rolls a 6 is given by $p_B = \frac{5}{36}$. We can compute the probability the games stops after n total rolls by reasoning as follows

- Since A starts the probability we stop with only one roll is p_A.
- The probability we stop at the second roll is $(1 - p_A)p_B$.
- The probability we stop at the third roll is $(1 - p_A)(1 - p_B)p_A$.
- The probability we stop at the fourth roll is

$$(1 - p_A)(1 - p_B)(1 - p_A)p_B = (1 - p_A)^2(1 - p_B)p_B \,.$$

- The probability we stop at the fifth roll is $(1 - p_A)^2(1 - p_B)^2 p_A$.
- The probability we stop at the sixth roll is $(1 - p_A)^3(1 - p_B)^2 p_B$.

From the above special cases, it looks like the probability we stop after an odd number say, $2n - 1$ rolls for $n \geq 1$ is given by

$$(1 - p_A)^{n-1}(1 - p_B)^{n-1} p_A \,,$$

and we stop after an even number of rolls say $2n$ for $n \geq 1$ is

$$(1 - p_A)^n (1 - p_B)^{n-1} p_B \,.$$

The final roll is made by A with probability we roll 1, 3, 5, or an odd number of times. We can evaluate the probability that A wins then as the sum of the elemental probabilities above. We find

$$\begin{aligned} P\{A \text{ wins}\} &= p_A + (1 - p_A)(1 - p_B)p_A + \cdots + (1 - p_A)^{n-1}(1 - p_B)^{n-1} p_A + \cdots \\ &= \sum_{n=1}^{\infty} p_A (1 - p_A)^{n-1}(1 - p_B)^{n-1} = p_A \sum_{n=0}^{\infty} (1 - p_A)^n (1 - p_B)^n \\ &= p_A \left(\frac{1}{1 - (1 - p_A)(1 - p_B)} \right) . \end{aligned}$$

When we use the values of p_A and p_B stated at the beginning of this problem we find $P\{A \text{ wins}\} = \frac{9}{19}$.

As another way to work this problem, let E be the event the last roll is made by player A and let A_i be the event that A wins on round i for $i = 1, 3, 5, \cdots$ and let B_i be the event B wins on round i for $i = 2, 4, 6, \cdots$. Then the event E (A wins) is the union of disjoint events (much as above) as

$$E = A_1 \cup A_1^c B_2^c A_3 \cup A_1^c B_2^c A_3^c B_4^c A_5 \cup A_1^c B_2^c A_3^c B_4^c A_5^c B_6^c A_7 \cup \ldots$$

Note that we can write the above in a way that emphasizes whether A wins on the first trial (or not) in the following way

$$E = A_1 \cup A_1^c B_2^c \left[A_3 \cup A_3^c B_4^c A_5 \cup A_3^c B_4^c A_5^c B_6^c A_7 \ldots \right] . \tag{12}$$

Notice that the term in brackets or

$$A_3 \cup A_3^c B_4^c A_5 \cup A_3^c B_4^c A_5^c B_6^c A_7 \ldots \,,$$

is the event that A wins given that he did not win on the first roll and that B did not win on the second roll. The probability of this even is the same as that of E since when A and B do not win we are starting the game anew. Thus we can evaluate this as

$$\begin{aligned} P(A_1^c B_2^c [A_3 \cup A_3^c B_4^c A_5 \cup A_3^c B_4^c A_5^c B_6^c A_7 \ldots]) &= P(A_3 \cup A_3^c B_4^c A_5 \cup A_3^c B_4^c A_5^c B_6^c A_7 \ldots | A_1^c B_2^c) P(A_1^c B_2^c) \\ &= P(E) P(A_1^c B_2^c) \,. \end{aligned}$$

Using Equation 12 we thus have shown

$$P(E) = P(A_1) + P(A_1^c B_2^c) P(E) \,,$$

or solving for $P(E)$ in the above we get

$$P(E) = \frac{P(A_1)}{1 - P(A_1^c B_2^c)} = \frac{P(A_1)}{1 - P(A_1^c) P(B_2^c)} = \frac{p_A}{1 - (1 - p_A)(1 - p_B)} \,,$$

the same expression as before.

Problem 75 (the eldest son)

Part (a): Let assume we have N families each of which can have four possible birth orderings of their two children

$$BB \,, BG \,, GB \,, GG \,.$$

Here the notation B stands for boy child and G stands for girl child, thus the combined event BG means that the family has a boy first followed by a girl and the same for the others. From the N total families that have two children there are $\frac{N}{4}$ of them have two sons, $\frac{N}{2}$ of them have one son, and $\frac{N}{4}$ of them have no sons. The total number of boys is then

$$2\left(\frac{N}{4}\right) + 1\left(\frac{N}{2}\right) = N \,.$$

Since in the birth ordering BB there is only one eldest sons, while in BG and GB there is also one eldest son the total number of eldest sons is

$$\frac{N}{4} + \frac{N}{2} = \frac{3N}{4} \,.$$

Thus the fraction of all sons that are an eldest son is

$$\frac{\frac{3}{4}N}{N} = \frac{3}{4} \,.$$

Part (b): As in the first part, each family can have three children from the following possible birth order

$$BBB \,, BBG \,, BGB \,, GBB \,, BGG \,, GBG \,, GGB \,, GGG \,.$$

From this we see that

- $\frac{N}{8}$ of the families have 3 sons of which only 1 is the eldest.
- $\frac{3N}{8}$ of the families have 2 sons of which only 1 is the eldest.
- $\frac{3N}{8}$ of the families have 1 son (who is also the oldest).
- $\frac{N}{8}$ of the families have 0 sons.

The total number of sons then is

$$3\left(\frac{N}{8}\right) + 2\left(\frac{3N}{8}\right) + \left(\frac{3N}{8}\right) = \frac{3N}{2} .$$

The total number of eldest sons is

$$\frac{N}{8} + \frac{3N}{8} + \frac{3N}{8} = \frac{7N}{8} .$$

The fraction of all sons that are eldest is

$$\frac{\frac{7N}{8}}{\frac{3N}{2}} = \frac{7}{12} .$$

Problem 76 (mutually exclusive events)

If E and F are mutually exclusive events in an experiment, then $P(E \cup F) = P(E) + P(F)$. We desire to compute the probability that E occurs before F, which we will denote by p. To compute p we condition on the three mutually exclusive events E, F, or $(E \cup F)^c$. This last event are all the outcomes not in E or F. Letting the event A be the event that E occurs before F, we have that

$$p = P(A|E)P(E) + P(A|F)P(F) + P(A|(E \cup F)^c)P((E \cup F)^c) .$$

Now

$$\begin{aligned} P(A|E) &= 1 \\ P(A|F) &= 0 \\ P(A|(E \cup F)^c) &= p \, , \end{aligned}$$

since if neither E or F happen the next experiment will have E before F (and thus event A with probability p). Thus we have that

$$\begin{aligned} p &= P(E) + pP((E \cup F)^c) \\ &= P(E) + p(1 - P(E \cup F)) \\ &= P(E) + p(1 - P(E) - P(F)) . \end{aligned}$$

Solving for p gives

$$p = \frac{P(E)}{P(E) + P(F)} ,$$

as we were to show.

Problem 77 (running independent trials)

Part (a): Since the trials are independent, knowledge of the third experiment does not give us any information about the outcome of the first experiment. Thus it could be any of the 3 choices (equally likely) so the probability we obtained a value of 1 is thus 1/3.

Part (b): Again by independence, we have a 1/3 chance in the first trial of getting a 1 and a 1/3 chance of getting a 1 on the second trial. Thus we get a 1/9 chance of the first two trials outputting 1's.

Problem 78 (play till someone wins)

Part (a): According to the problem description whoever wins the last game of these four is the winner and in the previous three games must have won once more than the other player. Let E be the event that the game ends after only four games. Let E_A be the event that the winner of the last game is A. The event that B wins the last game is then E_A^c. Then conditioning on who wins the last game we get

$$\begin{aligned} P(E) &= P(E|E_A)P(E_A) + P(E|E_A^c)P(E_A^c) \\ &= P(E|E_A)p + P(E|E_A^c)(1-p)\,. \end{aligned}$$

To calculate $P(E|E_A)$ we note that in the three games before the last one is played because the rules state we declare a winner if A or B gets two games ahead when we consider all the possible ways we can assign winners to these three games:

$$AAA\,, AAB\,, ABA\,, BAA\,, ABB\,, BAB\,, BBA\,, BBB\,,$$

we see that several of the orderings are not relevant for us. For example with the win pattern AAA, AAB, BBA, BBB we would have stopped playing before the third game. For the win patterns ABB and BAB we would keep playing after the fourth game won by A. Thus only *two* orderings ABA and BAA will result in A winning after he wins the fourth game. Thus

$$p(E|E_A) = 2p^2(1-p)\,.$$

By symmetry of A and B and p and $1-p$ we thus have

$$P(E|E_A^c) = 2(1-p)^2 p\,,$$

and the probability we want is given by

$$P(E) = 2p^3(1-p) + 2(1-p)^3 p = 2p(1-p)(p^2 + (1-p)^2)\,.$$

Part (b): Consider the first two games. If we stop and declare a total winner then they must be A_1A_2 or B_1B_2. If we don't stop they must be A_1B_2 or B_1A_2. If we don't stop, then since A and B both have won one game they are "tied" and its like the game is started all over again. Thus if we let E be the event that A wins the match we can evaluate $P(E)$ by conditioning on the result of the first two games. We find

$$\begin{aligned} P(E) &= P(E|A_1A_2)P(A_1A_2) + P(E|B_1B_2)P(B_1B_2) \\ &+ P(E|A_1B_2)P(A_1B_2) + P(E|B_1A_2)P(B_1A_2) \,. \end{aligned}$$

Note that

$$\begin{aligned} P(E|A_1A_2) &= 1 \\ P(E|B_1B_2) &= 0 \\ P(E|A_1B_2) &= P(E|B_1A_2) = P(E) \,, \end{aligned}$$

as in the last two cases the match "starts over". Using independence we then have

$$P(E) = p^2 + 2p(1-p)P(E) \,.$$

Solving for $P(E)$ in the above we get

$$P(E) = \frac{p^2}{1 - 2p(1-p)} = \frac{p^2}{(1-p+p)^2 - 2p(1-p)} = \frac{p^2}{(1-p)^2 + p^2} \,.$$

Problem 79 (2 7's before 6 even numbers)

In the problem statement we take the statement that we get 2 7's before 6 even numbers to not require 2 *consecutive* 7's or 6 *consecutive* even numbers. What is required is that the second occurrence of a 7 is before the sixth occurrence of an even number in the total sequence of trials. We will evaluate this probability by considering what happens on the first trial. Let S_1 be the event a **s**even is rolled on the first trial, let E_1 be the event an **e**ven number is rolled on the first trial and let X_1 be the event that neither a 7 or an even number is rolled on the first trial. Finally, let $R_{i,j}$ be the event we roll i 7's before j even numbers for $i \geq 0$ and $j \geq 0$. We will compute the probability of the event $R_{i,j}$ by conditioning on what happens on the first trial as

$$P(R_{i,j}) = P(R_{i,j}|S_1)P(S_1) + P(R_{i,j}|E_1)P(E_1) + P(R_{i,j}|X_1)P(X_1) \,. \tag{13}$$

Using Equation 2 the probabilities of the "one step" events S_1, E_1, and X_1 are given by

$$P(S_1) = \frac{1}{6} \equiv s \,, \quad P(E_1) = \frac{1}{2} \equiv e \,, \quad P(X_1) = 1 - \frac{1}{6} - \frac{1}{2} = \frac{1}{3} \equiv x \,,$$

where these probabilities are constant on each subsequent trial. The probabilities of the conditional events when $i > 1$ and $j > 1$ are given by

$$\begin{aligned} P(R_{i,j}|S_1) &= P(R_{i-1,j}) \\ P(R_{i,j}|E_1) &= P(R_{i,j-1}) \\ P(R_{i,j}|X_1) &= P(R_{i,j}) \,. \end{aligned}$$

With all of this Equation 13 above becomes

$$P(R_{i,j}) = sP(R_{i-1,j}) + eP(R_{i,j-1}) + xP(R_{i,j}) \quad \text{for} \quad i > 1 \,, \quad j > 1 \,. \tag{14}$$

For this problem we want to evaluate $P(R_{2,6})$ which using Equation 14 with $i = 2$ gives

$$P(R_{2,j}) = sP(R_{1,j}) + eP(R_{2,j-1}) + xP(R_{2,j}) \,,$$

or solving for $P(R_{2,j})$ we have

$$P(R_{2,j}) = \frac{s}{1-x} P(R_{1,j}) + \frac{e}{1-x} P(R_{2,j-1}) \,. \tag{15}$$

The first term on the right-hand-side shows that we need to be able to evaluate is $P(R_{1,j})$ for $j > 1$. We can compute this expression by taking $i = 1$ in Equation 13, where we get

$$\begin{aligned} P(R_{1,j}) &= sP(R_{1,j}|S_1) + eP(R_{1,j}|E_1) + xP(R_{1,j}|X_1) \\ &= s + eP(R_{1,j-1}) + xP(R_{1,j}) \,, \end{aligned}$$

where we have simplified by using $P(R_{1,j}|S_1) = 1$ and $P(R_{1,j}|E_1) = P(R_{1,j-1})$. Solving for $P(R_{1,j})$ we find

$$P(R_{1,j}) = \frac{s}{1-x} + \frac{e}{1-x} P(R_{1,j-1}) \,.$$

Based on the terms in the above expression lets define $\eta = \frac{s}{1-x}$ and $\xi = \frac{e}{1-x}$ where the above becomes

$$P(R_{1,j}) = \eta + \xi P(R_{1,j-1}) \,.$$

This can be iterated for $j = 2, 3, \cdots$ to give

$$\begin{aligned} P(R_{1,2}) &= \eta + \xi P(R_{1,1}) \\ P(R_{1,3}) &= \eta + \xi\eta + \xi^2 P(R_{1,1}) \\ P(R_{1,4}) &= \eta + \xi\eta + \xi^2\eta + \xi^3 P(R_{1,1}) \\ &\vdots \\ P(R_{1,j}) &= \eta \sum_{k=0}^{j-2} \xi^k + \xi^{j-1} P(R_{1,1}) \quad \text{for} \quad j \geq 2 \,. \end{aligned}$$

We can find the value of $P(R_{1,1})$ by taking $i = 1$ and $j = 1$ in Equation 13. We find

$$\begin{aligned} P(R_{1,1}) &= P(R_{1,1}|S_1)P(S_1) + P(R_{1,1}|E_1)P(E_1) + P(R_{1,1}|X_1)P(X_1) \\ &= s + xP(R_{1,1}) \,. \end{aligned}$$

When we solve the above for $P(R_{1,1})$ we get $P(R_{1,1}) = \frac{s}{1-x} = \eta$. Thus, using this we see that the expression for $P(R_{1,j})$ then becomes

$$P(R_{1,j}) = \eta \sum_{k=0}^{j-1} \xi^k \quad \text{for} \quad j \geq 1 \,. \tag{16}$$

Note that for any given value of $j \geq 1$ we can evaluate the above sum to compute $P(R_{1,j})$. Thus we can consider this a *known* function of j. Using this, and the expression for $P(R_{2,j})$ in Equation 15 we have

$$P(R_{2,j}) = \eta P(R_{1,j}) + \xi P(R_{2,j-1}) \,.$$

This we can iterate by letting $j = 2, 3, \cdots$ and observing the resulting pattern. We find

$$\begin{aligned}
P(R_{2,2}) &= \eta P(R_{1,2}) + \xi P(R_{2,1}) \\
P(R_{2,3}) &= \eta P(R_{1,3}) + \xi P(R_{2,2}) = \eta P(R_{1,3}) + \xi \eta P(R_{1,2}) + \xi^2 P(R_{2,1}) \\
&= \eta (P(R_{1,3}) + \xi P(R_{1,2})) + \xi^2 P(R_{2,1}) \\
P(R_{2,4}) &= \eta P(R_{1,4}) + \xi \eta P(R_{1,3}) + \eta \xi^2 P(R_{1,2}) + \xi^3 P(R_{2,1}) \\
&= \eta (P(R_{1,4}) + \xi P(R_{1,3}) + \xi^2 P(R_{1,2})) + \xi^3 P(R_{2,1}) \\
&\vdots \\
P(R_{2,j}) &= \eta \sum_{k=0}^{j-2} \xi^k P(R_{1,j-k}) + \xi^{j-1} P(R_{2,1}) \quad \text{for} \quad j \geq 2\,.
\end{aligned}$$

To use this we need to evaluate $P(R)$. Again using Equation 13 we have

$$P(R_{2,1}) = sP(R_{1,1}) + eP(R_{2,1}|E_1) + xP(R_{2,1}) = sP(R_{1,1}) + xP(R_{2,1})\,,$$

Solving for $P(R_{2,1})$ we get

$$P(R_{2,1}) = \frac{s}{1-x} P(R_{1,1}) = \left(\frac{s}{1-x}\right)^2 = \eta^2\,.$$

Thus we finally have for $P(R_{2,j})$ the following expression

$$P(R_{2,j}) = \eta \sum_{k=0}^{j-2} \xi^k P(R_{1,j-k}) + \eta^2 \xi^{j-1} \quad \text{for} \quad j \geq 2\,.$$

From the numbers given for s, e, and x we find

$$\eta = \frac{s}{1-x} = \frac{s}{s+e} = \frac{\frac{1}{6}}{\frac{1}{6}+\frac{1}{2}} = \frac{1}{4}, \quad \xi = \frac{e}{1-x} = \frac{e}{s+e} = \frac{\frac{1}{2}}{\frac{1}{6}+\frac{1}{2}} = \frac{3}{4}\,.$$

Thus we can evaluate $P(R_{2,6})$ using the above sum. In the python code `chap_3_prob_79.py` we implement the above computation. When we evaluate that code we get the probability 0.555053710938.

Problem 80 (contest playoff)

Part (a): We have that A will play in the first contest with probability one. He will play in the second contest only if he wins the first contest which happens with probability $\frac{1}{2}$. He will play in the third contest only if he wins the first two contests which happens with probability $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$. In general we have

$$P(A_i) = \left(\frac{1}{2}\right)^i \quad \text{for} \quad 1 \leq i \leq n\,.$$

Part (d): Consider various values for n. Then by considering the game trees for these values we see that, if $n = 1$ then 2^1 people play 1 game. If $n = 2$ then 2^3 play 3 games. If $n = 3$

then 7 games are played. Note in all of these cases the number of games played is $2^n - 1$. Let G_n be the number of games played when we have 2^n players. We can derive a recursive relationship for this as follows. Since on the first round we pair the 2^n players in the first game we will play $\frac{1}{2}2^n$ games. After the first round is played, half of the players have lost and no longer need to be considered. We thus have to consider 2^{n-1} players who will play G_{n-1} games. The total number of games played is the sum of these two numbers. This gives that

$$G_n = \frac{1}{2}2^n + G_{n-1} = 2^{n-1} + G_{n-1}\,.$$

Using this recursion relationship and the first few value of G_n computed earlier we can prove by induction that $G_n = 2^n - 1$.

Problem 81 (the stock market)

We are told that the value of the stock goes up or down 1 point successively and we want the probability that the value of the stock goes up to 40 before it goes down to 10. Another way to state this is that since the initial value of the stock is 25, we want the probability the stock goes up 15 points before it goes down 15 points. This problem is like the gambler's ruin problem, discussed in this chapter of the book in Example 4k. In the terminology of that example we assume that the stock owner and "the stock market" are playing a game that the stock owner wins with probability $p = 0.55$ and each player starts with 15 points. We then want the probability the gambler's fortune goes up to 30 units before it goes down to 0 units starting with an initial fortune of 15 units. Let E be the event gambler A (the stock owner) ends up with all the money when he starts with 15 units and gambler B (the market) starts with $(30 - 15) = 15$ units. From Example 4k we have

$$P(E) = \frac{1 - (q/p)^{15}}{1 - (q/p)^{30}} = \frac{1 - (0.45/0.55)^{15}}{1 - (0.45/0.55)^{30}} = 0.95302\,.$$

Problem 82 (flipping coins until a tail occurs)

Because Parts (a) and (c) are similar while Parts (b) and (d) are similar, I've solved them grouped in that order.

Part (a): For this part we want the probability that A gets 2 heads in a row before B does. To find this, let A be the event that A gets 2 heads in a row before B in a game where A *goes first*. In the same way let B be the event A gets 2 heads in a row before B in a game where now B *goes first*. Note in each event A or B the player A "wins". These events will come up naturally when we try to evaluate $P(A)$ by conditioning on the outcome of the first two flips. To do that we introduce the events that denote what happens on the first two flips. Let $H_1^A, H_2^A, T_1^A, T_2^A$ be the events A's coin lands heads up or tails up on the first or second flips. Let $H_1^B, H_2^B, T_1^B, T_2^B$ be the events B's coin lands heads up or tails up on the first or second flips. Then conditioning on what happens in the first few flips we have

$$
\begin{aligned}
P(A) &= P(A|H_1^A H_2^A)P(H_1^A H_2^A) + P(A|H_1^A T_2^A)P(H_1^A T_2^A) + P(A|T_1^A)P(T_1^A) \\
&= (1)P(H_1^A H_2^A) + P(B)P(H_1^A T_2^A) + P(B)P(T_1^A) \\
&= P_1^2 + [P_1(1-P_1) + (1-P_1)]P(B) \\
&= P_1^2 + (1+P_1)(1-P_1)P(B)\,.
\end{aligned}
$$

Where we have used expressions like $P(A|H_1^A T_2^A) = P(B)$ since in the case where A first flips $H_1^A T_2^A$ (or T_1^A) the dice go to B and all memory of the number of heads flipped is forgotten. In the above expression we need to evaluate $P(B)$ we can do this by again conditioning on the first few flips. We find

$$
\begin{aligned}
P(B) &= P(B|H_1^B H_2^B)P(H_1^B H_2^B) + P(B|H_1^B T_2^B)P(H_1^B T_2^B) + P(B|T_1^B)P(T_1^B) \\
&= (0)P(H_1^B H_2^B) + P(A)P(H_1^B T_2^B) + P(A)P(T_1^B) \\
&= (P_2(1-P_2) + (1-P_2))P(A) \\
&= (1+P_2)(1-P_2)P(A)\,.
\end{aligned}
$$

Putting this expression into the previous expression derived for $P(A)$ and we get

$$
P(A) = P_1^2 + (1+P_1)(1-P_1)(1+P_2)(1-P_2)P(A)\,.
$$

Solving for $P(A)$ we get

$$
P(A) = \frac{P_1^2}{1-(1+P_1)(1-P_1)(1+P_2)(1-P_2)} = \frac{P_1^2}{1-(1-P_1^2)(1-P_2^2)}\,.
$$

Part (c): For this part we want the probability that A gets 3 heads in a row before B does. In the same way as Part (a), let A be the event that A gets 3 heads in a row before B in a game where A *goes first*. In the same way let B be the event A gets 3 heads in a row before B in a game where now B *goes first*. Again in each event the player A "wins". Again let $H_1^A, H_2^A, H_3^A, T_1^A, T_2^A, T_3^A$ be the events the A's coin lands heads up or tails up on the first, second or third flips. Let $H_1^B, H_2^B, H_3^B, T_1^B, T_2^B, T_3^B$ be the events the B's coin lands heads up or tails up on the first, second or third flips. By conditioning on the first few flips we have

$$
\begin{aligned}
P(A) &= P(A|H_1^A H_2^A H_3^A)P(H_1^A H_2^A H_3^A) \\
&+ P(A|H_1^A H_2^A T_3^A)P(H_1^A H_2^A T_3^A) + P(A|H_1^A T_2^A)P(H_1^A T_2^A) + P(A|T_1^A)P(T_1^A) \\
&= (1)P(H_1^A H_2^A H_3^A) + P(H_1^A H_2^A T_3^A)P(B) + P(H_1^A T_2^A)P(B) + P(T_1^A)P(B) \\
&= P_1^3 + [P_1^2(1-P_1) + P_1(1-P_1) + (1-P_1)]P(B) \\
&= P_1^3 + (1-P_1)(P_1^2 + P_1 + 1)P(B)\,.
\end{aligned}
$$

As this involves $P(B)$ in the same way as earlier we compute that now

$$
\begin{aligned}
P(B) &= P(B|H_1^B H_2^B H_3^B)P(H_1^B H_2^B H_3^B) \\
&+ P(B|H_1^B H_2^B T_3^B)P(H_1^B H_2^B T_3^B) + P(B|H_1^B T_2^B)P(H_1^B T_2^B) + P(B|T_1^B)P(T_1^B) \\
&= (0)P(H_1^B H_2^B H_3^B) + P(A)P(H_1^B H_2^B T_3^B) + P(A)P(H_1^B T_2^B) + P(A)P(T_1^B) \\
&= (P_2^2(1-P_2) + P_2(1-P_2) + (1-P_2))P(A) \\
&= (1-P_2)(P_2^2 + P_2 + 1)P(A)\,.
\end{aligned}
$$

Using this in the expression derived for $P(A)$ earlier we have

$$P(A) = P_1^3 + (1 - P_1)(P_1^2 + P_1 + 1)(1 - P_2)(P_2^2 + P_2 + 1)P(A)\,.$$

Solving this for $P(A)$ we find

$$\begin{aligned} P(A) &= \frac{P_1^3}{1 - (P_1^2 + P_1 + 1)(1 - P_1)(P_2^2 + P_2 + 1)(1 - P_2)} \\ &= \frac{P_1^3}{1 - (1 - P_1^3)(1 - P_2^3)}\,. \end{aligned}$$

Part (b): For this part we want the probability that A gets a total of 2 heads before B does. Unlike the previous parts of the problem we now have to *remember* the number of heads that a given player has already made. Since that number makes it more or less likely for him to win the game. The motivation for this solution is that if A starts and gets a head on the first flip then A continues flipping. We can view the rest of the game from this point on as a new game where A starts and A wins if A gets a total of $2 - 1 = 1$ head before B gets a total of 2 heads. In the case when A gets tails on the first flip then B starts flipping. We can view the rest of the game as a new game where B now starts and A wins if A gets a total of 2 heads before B gets a total of 2 heads. Motivated by this, let $A(i, j)$ be the event A gets $2 - i$ heads before B gets $2 - j$ heads in a game where A starts and in the same way we let $B(i, j)$ be the event A gets $2 - i$ heads before B gets $2 - j$ heads in a game where B starts. In all cases we have $0 \leq i \leq 2$. In both of these events A "wins". Thus the indices i and j count the number of heads that each player has already received. The problem asks us to then find $P(A(0, 0))$. To do this it will also be easier to compute $P(B(0, 0))$ at the same time. Then conditioning on the result of the first flip we can compute $P(A(i, j))$ and $P(B(i, j))$ as

$$\begin{aligned} P(A(i, j)) &= P(A(i + 1, j))P_1 + P((B(i, j))(1 - P_1) \\ P(B(i, j)) &= P(B(i, j + 1))P_2 + P((A(i, j))(1 - P_2)\,. \end{aligned} \tag{17}$$

This is a recursive system where we need to evaluate $P(A(i, j))$ and $P(B(i, j))$ at various value of i and j to get $P(A(0, 0))$ and $P(B(0, 0))$. Note that to evaluate each of $P(A(i, j))$ and $P(B(i, j))$ using the above we need to know $P(A(i + 1, j))$ and $P(B(i, j + 1))$. That is we need probability values at the neighboring grid points $(i + 1, j)$ and $(i, j + 1)$. This motivates us to start with the boundary conditions

$$\begin{aligned} P(A(2, j)) &= 1 \quad \text{for} \quad 0 \leq j \leq 1 \\ P(B(2, j)) &= 1 \quad \text{for} \quad 0 \leq j \leq 1 \\ P(A(i, 2)) &= 0 \quad \text{for} \quad 0 \leq i \leq 1 \\ P(B(i, 2)) &= 0 \quad \text{for} \quad 0 \leq i \leq 1\,. \end{aligned}$$

and work *backwards*. We do this by using the above equations to solve for $P(A(i, j))$ and $P(B(i, j))$ at $(i, j) = (1, 1)$, then $(i, j) = (0, 1)$, then $(i, j) = (1, 0)$ and finally $(i, j) = (0, 0)$. To begin, let $i = j = 1$ in Equation 17 to get

$$\begin{aligned} P(A(1, 1)) &= P_1 P(A(2, 1)) + (1 - P_1)P(B(1, 1)) = P_1 + (1 - P_1)P(B(1, 1)) \\ P(B(1, 1)) &= P_2 P(B(1, 2)) + (1 - P_2)P(A(1, 1)) = (1 - P_2)P(A(1, 1))\,. \end{aligned}$$

When we solve this for $P(A(1,1))$ and $P(B(1,1))$ we get

$$\begin{aligned} P(A(1,1)) &= \frac{P_1}{1-(1-P_1)(1-P_2)} = \frac{P_1}{P_1+P_2-P_1P_2} \\ P(B(1,1)) &= \frac{P_1(1-P_2)}{1-(1-P_1)(1-P_2)} = \frac{P_1(1-P_2)}{P_1+P_2-P_1P_2} \,. \end{aligned} \tag{18}$$

Now let $i=1$ and $j=0$ in Equation 17 and we get

$$\begin{aligned} P(A(1,0)) &= P_1P(A(2,0)) + (1-P_1)P(B(1,0)) = P_1 + (1-P_1)P(B(1,0)) \\ P(B(1,0)) &= P_2P(B(1,1)) + (1-P_2)P(A(1,0)) \,. \end{aligned}$$

Since we know the value of $P(B(1,1))$ via Equation 18 we can solve the above for $P(A(1,0))$ and $P(B(1,0))$. Using Mathematica we get

$$\begin{aligned} P(A(1,0)) &= \frac{P_1(P_1(1-P_2)^2 + (2-P_2)P_2)}{(P_1+P_2-P_1P_2)^2} \\ P(B(1,0)) &= \frac{P_1(1-P_2)(P_1+2P_2-P_1P_2)}{(P_1+P_2-P_1P_2)^2} \,. \end{aligned}$$

Next let $i=0$ and $j=1$ in Equation 17 to get

$$\begin{aligned} P(A(0,1)) &= P_1P(A(1,1)) + (1-P_1)P(B(0,1)) \\ P(B(0,1)) &= P_2P(B(0,2)) + (1-P_2)P(A(0,1)) = (1-P_2)P(A(0,1)) \,. \end{aligned}$$

Since we have already evaluated $P(A(1,1))$ we can solve the above for $P(A(0,1))$ and $P(B(0,1))$. Using Mathematica when we do that we get

$$\begin{aligned} P(A(0,1)) &= \frac{P_1^2}{(P_1+P_2-P_1P_2)^2} \\ P(B(0,1)) &= \frac{P_1^2(1-P_2)}{(P_1+P_2-P_1P_2)^2} \,. \end{aligned}$$

Finally let $i=0$ and $j=0$ in Equation 17 to get

$$\begin{aligned} P(A(0,0)) &= P_1P(A(1,0)) + (1-P_1)P(B(0,0)) \\ P(B(0,0)) &= P_2P(B(0,1)) + (1-P_2)P(A(0,0)) \,. \end{aligned}$$

Since we know expressions for $P(A(1,0))$ and $P(B(0,1))$ we can solve the above for $P(A(0,0))$ and $P(B(0,0))$. Where we find

$$\begin{aligned} P(A(0,0)) &= \frac{P_1^2((3-2P_2)P_2 + P_1(1-3P_2+2P_2^2))}{(P_1(1-P_2)+P_2)^3} \\ P(B(0,0)) &= \frac{P_1^2(1-P_2)(P_1(1-P_2)^2 + (3-P_2)P_2)}{(P_1(1-P_2)+P_2)^3} \,. \end{aligned}$$

Part (d): Just as in Part (b) above we now have to keep track of the number of heads that each player has received as they play the game. We will use the same notation as above. We now have the boundary conditions

$$\begin{aligned} P(A(3,j)) &= 1 \quad \text{for} \quad 0 \le j \le 2 \\ P(B(3,j)) &= 1 \quad \text{for} \quad 0 \le j \le 2 \\ P(A(i,3)) &= 0 \quad \text{for} \quad 0 \le i \le 2 \\ P(B(i,3)) &= 0 \quad \text{for} \quad 0 \le i \le 2\,. \end{aligned}$$

with the same recursion relationship given by Equation 17 and work backwards to derive an expression for $P(A(0,0))$ and $P(B(0,0))$. We will outline the calculation that we will perform before we will write out recurrence relationships, as before, such that we will always be able to solve for $P(A(\cdot,\cdot))$ and $P(B(\cdot,\cdot))$ at each step. We write out the recurrence relationship for the grid point (i,j) in the following orders: $(2,2)$, $(1,2)$, $(2,1)$, $(0,2)$, $(1,1)$, $(2,1)$, $(0,1)$, $(1,0)$, and finally $(0,0)$ which will be the desired result. The ordering of these points is starting from the upper right corner of the (i,j) plane and working to the lower left corner towards $(0,0)$ diagonally. In this manner we have the correct boundary equations and previous solutions in place to solve for the next values of $P(A(\cdot,\cdot))$ and $P(B(\cdot,\cdot))$. This procedure can be implemented in Mathematica where we obtain

$$\frac{P_1^3(P_2^2(10-12P_2+3P_2^2)+P_1^2(-1+P_2)^2(1-3P_2+3P_2^2)+P_1P_2(5-20P_2+21P_2^2-6P_2^3))}{(P_1+P_2-P_1P_2)^5}\,,$$

for $P(A(0,0))$ and

$$\frac{P_1^3(1-P_2)(P_1^2(-1+P_2)^4+P_2^2(10-8P_2+P_2^2)+P_1P_2(5-15P_2+12P_2^2-2P_2^3))}{(P_1+P_2-P_1P_2)^5}\,,$$

for $P(B(0,0))$. The algebra for this problem is worked in the Mathematica file `chap_3_prob_82.nb`.

Problem 83 (red and white dice)

Let H be the event the coin lands heads up and we select the die A. From this H^c is then the event that the coin lands tails up and we select die B. Let R_n be the event a red face is showing on the nth roll of the die.

Part (a): We can compute $P(R_n)$ by conditioning on the result of the coin flip. We have

$$\begin{aligned} P(R_n) &= P(H)P(R_n|H) + P(H^c)P(R_n|H^c) \\ &= \left(\frac{1}{2}\right)\left(\frac{4}{6}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{6}\right) = \frac{1}{2}\,, \end{aligned}$$

when we simplify.

Part (b): We want to compute $P(R_3|R_1,R_2)$. We will do that using the definition of conditional probability or

$$P(R_3|R_1R_2) = \frac{P(R_1R_2R_3)}{P(R_1R_2)}\,.$$

We will evaluate $P(R_1R_2)$ and $P(R_1R_2R_3)$ by conditioning on the result of the first coin flip. We have

$$\begin{aligned}P(R_1R_2) &= P(R_1R_2|H)P(H) + P(R_1R_2|H^c)P(H^c)\\ &= P(R_1|H)P(R_2|H)P(H) + P(R_1|H^c)P(R_2|H^c)P(H^c)\,.\end{aligned}$$

and the same for $P(R_1R_2R_3)$ or

$$\begin{aligned}P(R_1R_2R_3) &= P(R_1R_2R_3|H)P(H) + P(R_1R_2R_3|H^c)P(H^c)\\ &= P(H)P(R_1|H)P(R_2|H)P(R_3|H) + P(H^c)P(R_1|H)P(R_2|H)P(R_3|H^c)\,.\end{aligned}$$

Thus we get for the probability we want

$$\begin{aligned}P(R_3|R_1R_2) &= \frac{P(H)P(R_1|H)P(R_2|H)P(R_3|H) + P(H^c)P(R_1|H)P(R_2|H)P(R_3|H^c)}{P(H)P(R_1|H)P(R_2|H) + P(H^c)P(R_1|H^c)P(R_2|H^c)}\\ &= \frac{\left(\frac{1}{2}\right)\left(\frac{4}{6}\right)^3 + \left(\frac{1}{2}\right)\left(\frac{2}{6}\right)^3}{\left(\frac{1}{2}\right)\left(\frac{4}{6}\right)^2 + \left(\frac{1}{2}\right)\left(\frac{2}{6}\right)^2} = \frac{\left(\frac{2}{3}\right)^3 + \left(\frac{1}{3}\right)^3}{\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2} = \frac{\frac{8}{27} + \frac{1}{27}}{\frac{4}{9} + \frac{1}{9}} = \frac{\frac{9}{27}}{\frac{5}{9}} = \frac{3}{5}\,.\end{aligned}$$

Part (c): For this part we want $P(H|R_1R_2)$. We have

$$\begin{aligned}P(H|R_1R_2) &= \frac{P(HR_1R_2)}{P(R_1R_2)}\\ &= \frac{P(H)P(R_1|H)P(R_2|H)}{P(H)P(R_1|H)P(R_2|H) + P(H^c)P(R_1|H^c)P(R_2|H^c)} = \frac{\frac{4}{9}}{\frac{5}{9}} = \frac{4}{5}\,.\end{aligned}$$

Problem 84 (4 white balls in an urn)

Some definitions that will be used in both parts of this problem. Let i be the index of the trial where one ball is drawn. Depending on what happens during the game A will be drawing on the trials $i = 1, 4, 7, 10, \ldots$, B will be drawing on the trials $i = 2, 5, 8, 11, \ldots$ and C will be drawing on the trials $i = 3, 6, 9, 12, \ldots$. Let W_i be the event that a white ball is selected on the ith draw (by whoever is drawing at the time). Finally, let A be the event A draws the first white ball and therefore wins. The same for the events B and C.

Part (a): In this case there is no memory as we place and the players keep drawing balls until there is a winner. We can compute the probability that A wins by conditioning on the result of the first set of draws. We find

$$\begin{aligned}P(A) &= P(A|W_1)P(W_1) + P(A|W_1^cW_2^cW_3^c)P(W_1^cW_2^cW_3^c)\\ &= P(W_1)P(A|W_1) + P(A|W_1^cW_2^cW_3^c)P(W_1^c)P(W_2^c)P(W_3^c)\\ &= P(W_1)(1) + P(A)P(W_1^c)P(W_2^c)P(W_3^c)\,.\end{aligned}$$

We can solve the above for $P(A)$ where we find

$$P(A) = \frac{P(W_1)}{1 - P(W_1^c)P(W_2^c)P(W_3^c)} = \frac{\frac{4}{12}}{1 - \left(\frac{8}{12}\right)^3} = \frac{\frac{1}{3}}{1 - \left(\frac{2}{3}\right)^3} = \frac{9}{19} .$$

We can evaluate $P(B)$ in the same way. We have

$$\begin{aligned} P(B) &= P(B|W_1^c W_2)P(W_1^c W_2) + P(W_4^c B|W_1^c W_2^c W_3^c)P(W_1^c W_2^c W_3^c) \\ &= P(B|W_1^c W_2)P(W_1^c)P(W_2) + P(W_4^c B|W_1^c W_2^c W_3^c)P(W_1^c)P(W_2^c)P(W_3^c) \\ &= P(W_1^c)P(W_2)(1) + P(W_1^c)P(W_2^c)P(W_3^c)P(B) . \end{aligned}$$

Solving for $P(B)$ we get

$$P(B) = \frac{P(W_1^c)P(W_2)}{1 - P(W_1^c)P(W_2^c)P(W_3^c)} = \frac{(\frac{8}{12})(\frac{4}{12})}{1 - (\frac{8}{12})^3} = \frac{(\frac{2}{3})(\frac{1}{3})}{1 - (\frac{2}{3})^3} = \frac{6}{19} .$$

For $P(C)$ we have

$$\begin{aligned} P(C) &= P(C|W_1^c W_2^c W_3)P(W_1^c W_2^c W_3) + P(W_4^c W_5^c C|W_1^c W_2^c W_3^c)P(W_1^c W_2^c W_3^c) \\ &= P(C|W_1^c W_2^c W_3)P(W_1^c)P(W_2^c)P(W_3) + P(W_4^c W_5^c C|W_1^c W_2^c W_3^c)P(W_1^c)P(W_2^c)P(W_3^c) \\ &= P(W_1^c)P(W_2^c)P(W_3)(1) + P(W_1^c)P(W_2^c)P(W_3^c)P(C) . \end{aligned}$$

Solving for $P(C)$ we get

$$P(C) = \frac{P(W_1^c)P(W_2^c)P(W_3)}{1 - (P(W_1^c)P(W_2^c)P(W_3^c))} = \frac{(\frac{8}{12})^2(\frac{4}{12})}{1 - (\frac{8}{12})^3)} = \frac{(\frac{2}{3})^2(\frac{1}{3})}{1 - (\frac{2}{3})^3)} = \frac{4}{19}$$

We can check that $P(A) + P(B) + P(C) = 1$ using the above numbers as it should.

Part (b): In this case we remove balls as each draw is made. Because of this the probabilities change after each ball is removed. In this case the game cannot run forever since the longest it can run would be the case where the white balls are drawn after all of the others. In other words we can draw at most 8 balls before getting a white ball. To compute the probability that A wins we can write this event as

$$W_1 \cup W_1^c W_2^c W_3^c W_4 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7 .$$

Note we don't have the possibility of W_8 since if A does not draw a white ball in the first 7 draws there is no way he can (some other player will draw it). Using the product rule of probability to evaluate the above union of independent events we have

$$\begin{aligned} P(A) &= P(W_1) + P(W_1^c W_2^c W_3^c W_4) + P(W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7) \\ &= P(W_1) + P(W_1^c)P(W_2^c|W_1^c)P(W_3^c|W_1^c W_2^c)P(W_4|W_1^c W_2^c W_3^c) \\ &+ P(W_1^c)P(W_2^c|W_1^c)P(W_3^c|W_1^c W_2^c)P(W_4^c|W_1^c W_2^c W_3^c) \\ &\times P(W_5^c|W_1^c W_2^c W_3^c W_4^c)P(W_6^c|W_1^c W_2^c W_3^c W_4^c W_5^c)P(W_7|W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c) \\ &= \frac{4}{12} + (\frac{8}{12})(\frac{7}{11})(\frac{6}{10})(\frac{4}{9}) + (\frac{8}{12})(\frac{7}{11})(\frac{6}{10})(\frac{5}{9})(\frac{4}{8})(\frac{3}{7})(\frac{4}{6}) = \frac{7}{15} . \end{aligned}$$

To compute the probability that B wins we write this event as

$$W_1^c W_2 \cup W_1^c W_2^c W_3^c W_4^c W_5 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8 \,.$$

From which we can compute the probability using

$$\begin{aligned}
P(B) &= P(W_1^c W_2 \cup W_1^c W_2^c W_3^c W_4^c W_5 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8) \\
&= P(W_1^c W_2) + P(W_1^c W_2^c W_3^c W_4^c W_5) + P(W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8) \\
&= P(W_1^c)P(W_2|W_1^c) \\
&+ P(W_1^c)P(W_2^c|W_1^c)P(W_3^c|W_1^c W_2^c)P(W_4^c|W_1^c W_2^c W_3^c)P(W_5|W_1^c W_2^c W_3^c W_4^c) \\
&+ P(W_1^c)P(W_2^c|W_1^c)P(W_3^c|W_1^c W_2^c)P(W_4^c|W_1^c W_2^c W_3^c)P(W_5^c|W_1^c W_2^c W_3^c W_4^c) \\
&\times P(W_6^c|W_1^c W_2^c W_3^c W_4^c W_5^c)P(W_7^c|W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c)P(W_8|W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c) \\
&= (\frac{8}{12})(\frac{4}{11}) + (\frac{8}{12})(\frac{7}{11})(\frac{6}{10})(\frac{5}{9})(\frac{4}{8}) + (\frac{8}{12})(\frac{7}{11})(\frac{6}{10})(\frac{5}{9})(\frac{4}{8})(\frac{3}{7})(\frac{2}{6})(\frac{4}{5}) = \frac{53}{165} \,.
\end{aligned}$$

To compute the probability that C wins we write this event as

$$W_1^c W_2^c W_3 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8^c W_9 \,.$$

From which we can compute the probability using

$$\begin{aligned}
P(C) &= P(W_1^c W_2^c W_3 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8^c W_9) \\
&= P(W_1^c W_2^c W_3) + P(W_1^c W_2^c W_3^c W_4^c W_5^c W_6) + P(W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8^c W_9) \\
&= P(W_1^c)P(W_2|W_1^c)P(W_3|W_1^c W_2^c) \\
&+ P(W_1^c)P(W_2^c|W_1^c)P(W_3^c|W_1^c W_2^c) \\
&\times P(W_4^c|W_1^c W_2^c W_3^c)P(W_5^c|W_1^c W_2^c W_3^c W_4^c)P(W_6|W_1^c W_2^c W_3^c W_4^c W_5^c) \\
&+ P(W_1^c)P(W_2^c|W_1^c)P(W_3^c|W_1^c W_2^c) \\
&\times P(W_4^c|W_1^c W_2^c W_3^c)P(W_5^c|W_1^c W_2^c W_3^c W_4^c)P(W_6^c|W_1^c W_2^c W_3^c W_4^c W_5^c) \\
&\times P(W_7^c|W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c)P(W_8^c|W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c)P(W_9|W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8^c) \\
&= (\frac{8}{12})(\frac{7}{11})(\frac{4}{10}) + (\frac{8}{12})(\frac{7}{11})(\frac{6}{10})(\frac{5}{9})(\frac{4}{8})(\frac{4}{7}) + (\frac{8}{12})(\frac{7}{11})(\frac{6}{10})(\frac{5}{9})(\frac{4}{8})(\frac{3}{7})(\frac{2}{6})(\frac{1}{5})(1) = \frac{7}{33} \,.
\end{aligned}$$

We can again check that $P(A) + P(B) + P(C) = 1$ using the above numbers as it should.

Problem 85 (4 white balls in 3 urns)

Part (a): This is the same as Problem 84 Part (a).

Part (b): First note that each tail A flips reduces the number of non-white balls by one. Thus we can have at most 8 non-white flips before we must get a white ball. Using the same notation from Problem 84 we have that the event that A wins given by

$$\begin{aligned}
A &= W_1 \cup W_1^c W_2^c W_3^c W_4 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7 \dots \\
&= W_1 \cup \left(\prod_{i=1}^{3} W_i^c \right) W_4 \cup \left(\prod_{i=1}^{6} W_i^c \right) W_7 \cup \left(\prod_{i=1}^{9} W_i^c \right) W_{10} \\
&\cup \left(\prod_{i=1}^{12} W_i^c \right) W_{13} \cup \left(\prod_{i=1}^{15} W_i^c \right) W_{16} \cup \left(\prod_{i=1}^{18} W_i^c \right) W_{19} \\
&\cup \left(\prod_{i=1}^{21} W_i^c \right) W_{22} \cup \left(\prod_{i=1}^{24} W_i^c \right) W_{25} \\
&= \cup_{n=0}^{8} \left(\prod_{i=1}^{3n} W_i^c \right) W_{3n+1} \,.
\end{aligned}$$

Which are the events that A wins by flipping no tails, one tail, two tails etc. We are using the convention that $\prod_{i=1}^{0} \cdot = 1$. Thus the probability A wins is given by

$$\begin{aligned}
P(A) &= \frac{4}{12} + \left(\frac{8}{12}\right)^3 \left(\frac{4}{11}\right) + \left(\frac{8}{12}\right)^3 \left(\frac{7}{11}\right)^3 \left(\frac{4}{10}\right) + \dots \\
&= \sum_{n=0}^{8} \left(\prod_{k=1}^{n} \left(\frac{8-(k-1)}{12-(k-1)} \right)^3 \right) \left(\frac{4}{12-n} \right) \\
&= \sum_{n=0}^{8} \left(\prod_{k=1}^{n} \left(\frac{9-k}{13-k} \right)^3 \right) \left(\frac{4}{12-n} \right) .
\end{aligned}$$

The event B wins is given by

$$B = W_1^c W_2 \cup W_1^c W_2^c W_3^c W_4^c W_5 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8 \cup \dots \cup \left(\prod_{i=1}^{3n+1} W_i^c \right) W_{3n+2} \cup \dots .$$

Which are the events that B wins by flipping no tails, one tail, two tails etc. Since A can fail to draw a white ball at most 8 times, the last set in the above union can be when $n = 7$. Thus we have

$$B = \cup_{n=0}^{7} \left(\prod_{i=1}^{3n+1} W_i^c \right) W_{3n+2} \,.$$

Thus the probability B wins is given by

$$\begin{aligned}
P(B) &= \left(\frac{8}{12}\right)\left(\frac{4}{12}\right) + \left(\frac{8}{12}\right)^3 \left(\frac{7}{11}\right)\left(\frac{4}{11}\right) + \left(\frac{8}{12}\right)^3 \left(\frac{7}{11}\right)^3 \left(\frac{6}{10}\right)\left(\frac{4}{10}\right) + \dots \\
&= \sum_{n=0}^{7} \left(\prod_{k=1}^{n} \left(\frac{9-k}{13-k} \right)^3 \right) \left(\frac{8-n}{12-n} \right) \left(\frac{4}{12-n} \right) .
\end{aligned}$$

Finally, the event C wins is given by

$$\begin{aligned}
C &= W_1^c W_2^c W_3 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6 \cup W_1^c W_2^c W_3^c W_4^c W_5^c W_6^c W_7^c W_8^c W_9 \cup \cdots \\
&= \cup_{n=0}^{7} \left(\prod_{i=1}^{3n+2} W_i^c \right) W_{3n+3} \,.
\end{aligned}$$

Which are the events that C wins by flipping no tails, one tail, two tails etc. Thus the probability C wins is given by

$$
\begin{aligned}
P(C) &= \left(\frac{8}{12}\right)^2\left(\frac{4}{12}\right) + \left(\frac{8}{12}\right)^3\left(\frac{7}{11}\right)^2\left(\frac{4}{11}\right) + \left(\frac{8}{12}\right)^3\left(\frac{7}{11}\right)^3\left(\frac{6}{10}\right)^3\left(\frac{4}{10}\right) + \dots \\
&= \sum_{n=0}^{7}\left(\prod_{k=1}^{n}\left(\frac{9-k}{13-k}\right)^3\right)\left(\frac{8-n}{12-n}\right)^2\left(\frac{4}{12-n}\right).
\end{aligned}
$$

We evaluate all of these expressions in the python code `chap_3_prob_85.py` where we get the values $P(A) = 0.48058$, $P(B) = 0.3139$, and $P(C) = 0.20543$. These numbers satisfy $P(A) + P(B) + P(C) = 1$. as they should.

Problem 86 ($A \subset B$)

Part (a): First note that there are $\binom{n}{i}$ subsets with i elements. There are a total of 2^n subsets of S. Thus

$$
P(N(B) = i) = \frac{\binom{n}{i}}{2^n}.
$$

To evaluate $P(A \subset B|N(B) = i)$ note that originally the set A can be any of the 2^n subsets of S. Since B has i elements to have $A \subset B$ means that all the elements of A must actually also be elements of B (of which there are i). Thus A must be one of the 2^i subsets of B and we have

$$
P(A \subset B|N(B) = i) = \frac{2^i}{2^n}.
$$

Using these two results we now have

$$
\begin{aligned}
P(E) &= \sum_{i=0}^{n} P(A \subset B|N(B) = i)P(N(B) = i) = \sum_{i=0}^{n}\left(\frac{\binom{n}{i}}{2^n}\right)\left(\frac{2^i}{2^n}\right) \\
&= \left(\frac{1}{2^n}\right)\left(\frac{1}{2^n}\right)\sum_{i=0}^{n}\binom{n}{i}2^i = \left(\frac{1}{2^n}\right)\left(\frac{1}{2^n}\right)\left(1+2\right)^n = \left(\frac{3}{4}\right)^n.
\end{aligned}
$$

Part (b): Note that $AB = \emptyset$ is equivalent to the statement $A \subset B^c$. Since B^c is also a subset of S. From the previous part we have

$$
P(A \subset B^c) = \left(\frac{3}{4}\right)^n.
$$

Problem 87 (Laplace's rule of succession I)

As shown in example 5e, where C_i is the event we draw the ith coin $0 \le i \le k$ and F_n is the event that the first n flips give head we have

$$
P(C_i) = \frac{1}{k+1}, \quad \text{and} \quad P(F_n|C_i) = \left(\frac{i}{k}\right)^n,
$$

Thus the conditional probability requested is then

$$P(C_i|F_n) = \frac{P(C_iF_n)}{P(F_n)} = \frac{P(C_i)P(F_n|C_i)}{\sum_{j=0}^{k} P(C_j)P(F_n|C_j)}$$
$$= \frac{\left(\frac{1}{k+1}\right)\left(\frac{i}{k}\right)^n}{\sum_{j=0}^{k}\left(\frac{1}{k+1}\right)\left(\frac{j}{k}\right)^n} = \frac{\left(\frac{i}{k}\right)^n}{\sum_{j=0}^{k}\left(\frac{j}{k}\right)^n}.$$

Problem 88 (Laplace's rule of succession II)

The outcomes of the successive flips are not independent. The can be argued by noting the fact that we get a head on the first flip make it more likely that we drew a coin that was biased towards landing heads. We can show that by showing that $P(H_1H_2) \neq P(H_1)P(H_2)$. Again conditioning on the initial coin selected we have

$$\begin{aligned} P(H_1H_2) &= \sum_{i=0}^{k} P(H_1H_2C_i) = \sum_{i=0}^{k} P(C_i)P(H_1H_2|C_i) \\ &= \sum_{i=0}^{k} P(C_i)P(H_1|C_i)P(H_2|C_i) \\ &= \sum_{i=0}^{k}\left(\frac{1}{k+1}\right)\left(\frac{i}{k}\right)^2 = \left(\frac{1}{k+1}\right)\left(\frac{1}{k^2}\right)\sum_{i=0}^{k} i^2 \\ &= \left(\frac{1}{k+1}\right)\left(\frac{1}{k^2}\right)\left(\frac{1}{6}k(k+1)(2k+1)\right) = \frac{2k+1}{6k}. \end{aligned}$$

The probability of a single head is

$$\begin{aligned} P(H_1) &= \sum_{i=0}^{k} P(H_1C_i) = \sum_{i=0}^{k} P(C_i)P(H_1|C_i) \\ &= \sum_{i=0}^{k}\left(\frac{1}{k+1}\right)\left(\frac{i}{k}\right) = \left(\frac{1}{k+1}\right)\left(\frac{1}{k}\right)\sum_{i=0}^{k} i = \left(\frac{1}{k+1}\right)\left(\frac{1}{k}\right)\frac{(k+1)k}{2} = \frac{1}{2}. \end{aligned}$$

Notice that $P(H_1)P(H_2) = \frac{1}{4} \neq P(H_1H_2)$ and the events are not independent.

Problem 89 (3 judges vote guilty)

As suggested from the book let E_i be the event that judge i votes guilty for $i = 1, 2, 3$. Let G be the event the person is actually guilty. Then the problem tells us that $P(E_i|G) = .7$, $P(E_i|G^c) = .2$, and $P(G) = .7$.

Part (a): We want to evaluate $P(E_3|E_1E_2)$ which we do by the definition of conditional independence as

$$P(E_3|E_1E_2) = \frac{P(E_1E_2E_3)}{P(E_1E_2)}.$$

We can evaluate each of the probabilities above by conditioning on whether or not the person is guilty. That is

$$\begin{aligned} P(E_1E_2) &= P(E_1E_2|G)P(G) + P(E_1E_2|G^c)P(G^c) \\ &= P(E_1|G)P(E_2|G)P(G) + P(E_1|G^c)P(E_2|G^c)P(G^c) \\ &= (0.7)(0.7)^2 + (0.3)(0.2)^2 , \end{aligned}$$

and

$$\begin{aligned} P(E_1E_2E_3) &= P(E_1E_2E_3|G)P(G) + P(E_1E_2E_3|G^c)P(G^c) \\ &= P(E_1|G)P(E_2|G)P(E_3|G)P(G) + P(E_1|G^c)P(E_2|G^c)P(E_3|G^c)P(G^c) \\ &= (0.7)(0.7)^3 + (0.3)(0.2)^3 . \end{aligned}$$

Thus

$$P(E_3|E_1E_2) = \frac{(0.7)(0.7)^3 + (0.3)(0.2)^3}{(0.7)(0.7)^2 + (0.3)(0.2)^2} = \frac{0.2401 + 0.0024}{0.343 + 0.012} = \frac{0.2425}{0.355} = \frac{97}{142}.$$

Part (b): One way to interpret the problem is to say that we are asked for $P(E_3|E_1E_2^c)$ and then this part can be worked just like the previous one. As above we would need to evaluate $P(E_1E_2^cE_3)$ and $P(E_1E_2^c)$ and then take their ratio. If, instead, we interpret the problem to be: compute the probability that judge 3 voted guilty given that *one* of the two previous judges voted guilty. In this case we don't know which of the two previous judge's voted guilty and which did not. Thus the event we are conditioning on is $E_1E_2^c \cup E_1^cE_2$. We thus have

$$P(E_3|E_1E_2^c \cup E_1^cE_2) = \frac{P(E_1E_2^cE_3 \cup E_1^cE_2E_3)}{P(E_1E_2^c \cup E_1^cE_2)} = \frac{P(E_1E_2^cE_3) + P(E_1^cE_2E_3)}{P(E_1E_2^c) + P(E_1^cE_2)}.$$

We now calculate each of these probabilities in tern

$$\begin{aligned} P(E_1E_2^cE_3) &= P(G)P(E_1E_2^cE_3|G) + P(G^c)P(E_1E_2^cE_3|G^c) \\ &= P(G)P(E_1|G)P(E_2^c|G)P(E_3|G) + P(G^c)P(E_1|G^c)P(E_2^c|G^c)P(E_3|G^c) \\ &= (0.7)(0.7)(1-0.7)(0.7) + (0.3)(0.2)(1-0.2)(0.2) . \end{aligned}$$

We do the same thing for $P(E_1^cE_2E_3)$ and find

$$P(E_1^cE_2E_3) = (0.7)(1-0.7)(0.7)^2 + (0.3)(0.2)(1-0.2)(0.2)^2 .$$

Now for $P(E_1E_2^c)$ we have

$$\begin{aligned} P(E_1E_2^c) &= P(G)P(E_1E_2^c|G) + P(G^c)P(E_1E_2^c|G^c) \\ &= P(G)P(E_1|G)P(E_2^c|G) + P(G^c)P(E_1|G^c)P(E_2^c|G^c) \\ &= (0.7)(0.7)(1-0.7) + (0.3)(0.2)(1-0.2) . \end{aligned}$$

We do the same thing for $P(E_1 E_2)$ and find

$$P(E_1^c E_2) = (0.7)(1 - 0.7)(0.7) + (0.3)(1 - 0.2)(0.2) \,.$$

Thus we get

$$\begin{aligned} P(E_3 | E_1 E_2^c \cup E_1^c E_2) &= \frac{2[(0.7)(0.7)^2(1 - 0.7) + (0.3)(0.2)^2(1 - 0.2)]}{2[(0.7)(0.7)(1 - 0.7) + (0.3)(0.2)(1 - 0.2)]} \\ &= \frac{0.1029 + 0.0096}{0.147 + 0.048} = \frac{0.1125}{0.195} = \frac{15}{26} \,. \end{aligned}$$

Part (c): We want to evaluate $P(E_3 | E_1^c E_2^c)$ which we also do by the definition of conditional independence from

$$P(E_3 | E_1^c E_2^c) = \frac{P(E_1^c E_2^c E_3)}{P(E_1^c E_2^c)} \,.$$

As before

$$\begin{aligned} P(E_1^c E_2^c E_3) &= P(G) P(E_1^c E_2^c E_3 | G) + P(G^c) P(E_1^c E_2^c E_3 | G^c) \\ &= P(G) P(E_1^c | G) P(E_2^c | G) P(E_3 | G) + P(G^c) P(E_1^c | G^c) P(E_2^c | G^c) P(E_3 | G^c) \\ &= (0.7)(1 - 0.7)^2(0.7) + (0.3)(1 - 0.2)^2(0.2) \,, \end{aligned}$$

and

$$\begin{aligned} P(E_1^c E_2^c) &= P(G) P(E_1^c E_2^c | G) + P(G^c) P(E_1^c E_2^c | G^c) \\ &= P(G) P(E_1^c | G) P(E_2^c | G) + P(G^c) P(E_1^c | G^c) P(E_2^c | G^c) \\ &= (0.7)(1 - 0.7)^2 + (0.3)(1 - 0.2)^2 \,. \end{aligned}$$

Thus we have

$$P(E_3 | E_1 E_2) = \frac{(0.7)(1 - 0.7)^2(0.7) + (0.3)(1 - 0.2)^2(0.2)}{(0.7)(1 - 0.7)^2 + (0.3)(1 - 0.2)^2} = \frac{0.0441 + 0.0384}{0.063 + 0.192} = \frac{0.0825}{0.255} = \frac{33}{102} \,.$$

Problem 90 (n trials, 3 outcomes)

We want to evaluate the probability of the event E defined such that

outcomes 1 and 2 both occur at least once.

This is equivalent to the event

outcome 1 occurs at least once and outcome 2 occurs at least once.

The event is also equal to the contra-positive of the above statement which is

not (1 never occurs or 2 never occurs).

This last event is in tern is equivalent to the event

not (0 or 2 always occur or 0 or 1 always occur).

We will evaluate the probability of this last event and then from it derive the probability of the event of interest or E. Let U_i, V_i, W_i be the events outcome 0,1,2 occurs on the ith trial. From the above we have argued that

$$E = \left(\bigcap_{i=1}^{n} (U_i \cup W_i) \ \cup \ \bigcap_{i=1}^{n} (U_i \cup V_i) \right)^c .$$

We first find $P(E^c)$ given by

$$\begin{aligned}
P(E^c) &= P\Big(\bigcap_{i=1}^{n} (U_i \cup W_i) \ \cup \ \bigcap_{i=1}^{n} (U_i \cup V_i) \Big) \\
&= P\Big(\bigcap_{i=1}^{n} (U_i \cup W_i) \Big) + P\Big(\bigcap_{i=1}^{n} (U_i \cup V_i) \Big) - P\Big(\bigcap_{i=1}^{n} (U_i \cup W_i) \bigcap_{i=1}^{n} (U_i \cup V_i) \Big) \\
&= P\Big(\bigcap_{i=1}^{n} (U_i \cup W_i) \Big) + P\Big(\bigcap_{i=1}^{n} (U_i \cup V_i) \Big) - P\Big(\bigcap_{i=1}^{n} (U_i \cup W_i)(U_i \cup V_i) \Big) \\
&= P\Big(\bigcap_{i=1}^{n} (U_i \cup W_i) \Big) + P\Big(\bigcap_{i=1}^{n} (U_i \cup V_i) \Big) - P\Big(\bigcap_{i=1}^{n} U_i \Big) \\
&= \prod_{i=1}^{n} P(U_i \cup W_i) + \prod_{i=1}^{n} P(U_i \cup V_i) - \prod_{i=1}^{n} P(U_i) \\
&= (p_0 + p_2)^n + (p_0 + p_1)^n - p_0^n .
\end{aligned}$$

Thus using this result we have for $P(E)$ that

$$\begin{aligned}
P(E) &= P(\Big(\bigcap_{i=1}^{n} (U_i \cup W_i) \ \cup \ \bigcap_{i=1}^{n} (U_i \cup V_i) \Big)^c) \\
&= 1 - P\Big(\bigcap_{i=1}^{n} (U_i \cup W_i) \ \cup \ \bigcap_{i=1}^{n} (U_i \cup V_i) \Big) \\
&= 1 - (p_0 + p_2)^n - (p_0 + p_1)^n + p_0^n .
\end{aligned}$$

Chapter 3: Theoretical Exercises

Problem 1 (conditioning on more information)

We have

$$P(A \cap B | A) = \frac{P(A \cap B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)} .$$

and

$$P(A \cap B | A \cup B) = \frac{P((A \cap B) \cap (A \cup B))}{P(A \cup B)} = \frac{P(A \cap B)}{P(A \cup B)} .$$

But since $A \cup B \supset A$, the probabilities $P(A \cup B) \geq P(A)$, so

$$\frac{P(A \cap B)}{P(A)} \geq \frac{P(A \cap B)}{P(A \cup B)}$$

giving

$$P(A \cap B|A) \geq P(A \cap B|A \cup B)\,,$$

the desired result.

Problem 2 (simplifying conditional expressions)

Using the definition of conditional probability we can compute

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}\,.$$

since $A \subset B$. In words $P(A|B)$ is the amount of A in B. For $P(A|\neg B)$ we have

$$P(A|\neg B) = \frac{P(A \cap \neg B)}{P(\neg B)} = \frac{P(\phi)}{P(\neg B)} = 0\,.$$

Since if $A \subset B$ then $A \cap \neg B$ is empty or in words given that $\neg B$ occurred and $A \subset B$, A cannot have occurred and therefore has zero probability. For $P(B|A)$ we have

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1\,,$$

or in words since A occurs and B contains A, B must have occurred giving probability one. For $P(B|\neg A)$ we have

$$P(B|\neg A) = \frac{P(B \cap \neg A)}{P(\neg A)}\,,$$

which cannot be simplified further.

Problem 3 (biased selection of the first born)

We define n_1 to be the number of families with one child, n_2 the number of families with two children, and in general n_k to be the number of families with k children. In this problem we want to compare two different methods for selecting children. In the first method, M_1, we pick one of the m families and then randomly choose a child from that family. In the second method, M_2, we directly pick one of the $\sum_{i=1}^{k} i n_i$ children randomly. Let E be the event that a first born child is chosen. Then the question seeks to prove that

$$P(E|M_1) > P(E|M_2)\,.$$

We will solve this problem by conditioning no the number of families with i children. For example under M_1 we have (dropping the conditioning on M_1 for notational simplicity) that

$$P(E) = \sum_{i=1}^{k} P(E|F_i)P(F_i)\,,$$

where F_i is the event that the chosen family has i children. This later probability is given by

$$P(F_i) = \frac{n_i}{m},$$

for we have n_i families with i children from m total families. Also

$$P(E|F_i) = \frac{1}{i},$$

since the event F_i means that our chosen family has i children and the event E means that we select the first born, which can be done in $\frac{1}{i}$ ways. In total then we have under M_1 the following for $P(E)$

$$P(E) = \sum_{i=1}^{k} P(E|F_i)P(F_i) = \sum_{i=1}^{k} \frac{1}{i}\left(\frac{n_i}{m}\right) = \frac{1}{m}\sum_{i=1}^{k} \frac{n_i}{i}.$$

Now under the second method again $P(E) = \sum_{i=1}^{k} P(E|F_i)P(F_i)$ but under the second method $P(F_i)$ is the probability we have selected a family with i children and is given by

$$\frac{in_i}{\sum_{i=1}^{k} in_i},$$

since in_i is are the number of children from families with i children and the denominator is the total number of children. Now $P(E|F_i)$ is still the probability of having selected a family with ith children we select the first born child. This is

$$\frac{n_i}{in_i} = \frac{1}{i},$$

since we have in_i total children from the families with i children and n_i of them are first born. Thus under the second method we have

$$P(E) = \sum_{i=1}^{k}\left(\frac{1}{i}\right)\left(\frac{in_i}{\sum_{l=1}^{k} ln_l}\right) = \frac{1}{\left(\sum_{l=1}^{k} ln_l\right)}\sum_{i=1}^{k} n_i.$$

Then our claim that $P(E|M_1) > P(E|M_2)$ is equivalent to the statement that

$$\frac{1}{m}\sum_{i=1}^{k}\frac{n_i}{i} \geq \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} in_i}$$

or remembering that $m = \sum_{i=1}^{k} n_i$ that

$$\left(\sum_{i=1}^{k} in_i\right)\left(\sum_{j=1}^{k}\frac{n_j}{j}\right) \geq \left(\sum_{i=1}^{k} n_i\right)\left(\sum_{j=1}^{k} n_j\right).$$

To show that this is true (and that all earlier results are true) expand each expression. First the left hand side LHS, we obtain

$$\begin{aligned}
\text{LHS} &= (n_1 + 2n_2 + 3n_3 + \cdots + kn_k)\left(n_1 + \frac{n_2}{2} + \frac{n_3}{3} + \cdots + \frac{n_k}{k}\right) \\
&= n_1^2 + \frac{n_1 n_2}{2} + \frac{n_1 n_3}{3} + \frac{n_1 n_4}{4} + \cdots + \frac{n_1 n_k}{k} \\
&+ 2n_2 n_1 + n_2^2 + \frac{2n_2 n_3}{3} + \cdots + \frac{2n_2 n_k}{k} + \cdots \\
&+ kn_k n_1 + \frac{kn_k n_2}{2} + \cdots + n_k^2.
\end{aligned}$$

By grouping the polynomial terms we find that the above is equivalent to

$$\begin{aligned}
\text{LHS} &= n_1^2 + \left(\frac{1}{2} + 2\right) n_1 n_2 + \left(\frac{1}{3} + 3\right) n_1 n_3 + \cdots + \left(\frac{1}{k} + k\right) n_1 n_k \\
&+ n_2^2 + \left(\frac{2}{3} + \frac{3}{2}\right) n_2 n_3 + \left(\frac{4}{2} + \frac{2}{4}\right) n_2 n_4 + \cdots + \left(\frac{2}{k} + \frac{k}{2}\right) n_2 n_k \\
&+ n_3^2 + \left(\frac{3}{4} + \frac{4}{3}\right) n_3 n_4 + \cdots + \left(\frac{3}{k} + \frac{k}{3}\right) n_3 n_k + \cdots
\end{aligned}$$

Thus in general the $n_i n_j$ term has a coefficient given by 1 if $i = j$ and by

$$\frac{i}{j} + \frac{j}{i} \,,$$

if $i < j \le k$. While the expansion of the right hand side RHS will have a coefficient given by 1 if $i = j$ and 2 if $i < j \le k$. Thus a sufficient condition for the left hand side to be greater than the right hand side is for

$$\frac{i}{j} + \frac{j}{i} > 2 \quad \text{when} \quad i \neq j \quad \text{and} \quad i < j \le k \,.$$

By multiplying by the product ij, we have the above equivalent to

$$i^2 + j^2 > 2ij \,,$$

which in turn is equivalent to

$$i^2 - 2ij + j^2 > 0 \,,$$

or

$$(i - j)^2 > 0 \,,$$

which we know to be true for all i and j. Because (using reversible transformations) we have reduced our desired inequality to one that we know to be true we have shown the desired identity.

Problem 4 (fuzzy searching for balls in a box)

Let E_i be the event that the ball is present in box i. Let S_i be the event that the search of box i yields a success or "finds" the ball. Then the statement of the problem tells us that $P(E_i) = P_i$ and

$$P(S_i|E_j) = \begin{cases} \alpha_i & j = i \\ 0 & j \neq i \end{cases} .$$

We desire to compute $P(E_j|S_i^c)$ which by Bayes' rule is equal to

$$P(E_j|S_i^c) = \frac{P(S_i^c|E_j)P(E_j)}{P(S_i^c)} = \frac{(1 - P(S_i|E_j))P(E_j)}{1 - P(S_i)} \,.$$

Lets begin by computing $P(S_i)$. We have

$$P(S_i) = \sum_{k=1}^{n} P(S_i|E_k)P(E_k) = \alpha_i p_i \,.$$

Using the above we can compute our desired expression. We find

$$P(E_j|S_i^c) = \begin{cases} \frac{(1-\alpha_i)p_i}{1-\alpha_i p_i} & j = i \\ \frac{p_j}{1-\alpha_i p_i} & j \neq i \end{cases},$$

which is the desired result.

Problem 5 (negative information)

An event F is said to carry negative information about an event E and is written $F \searrow E$ if $P(E|F) \leq P(E)$.

Part (a): If $F \searrow E$ then $P(E|F) \leq P(E)$ so that using Bayes' rule for $P(E|F)$ we see that this is equivalent to the expression

$$\frac{P(F|E)P(E)}{P(F)} \leq P(E)\,,$$

and assuming $P(E) \neq 0$. This implies that $P(F|E) \leq P(F)$ so that $E \searrow F$.

Problem 6 (the union of independent events)

We recognize that $E_1 \cup E_2 \cup \cdots E_n$ as the event that at least one of the events E_i occurs. Consider the event $\neg E_1 \cap \neg E_2 \cap \neg E_3 \cdots \cap \neg E_n = \neg(E_1 \cup E_2 \cup E_3 \cup \cdots \cup E_n)$, which is the event that none of the E_i occur. Then we have

$$P(E_1 \cup E_2 \cup \cdots \cup E_n) = 1 - P(\neg(E_1 \cup E_2 \cup \cdots \cup E_n)) = 1 - P(\neg E_1 \cap \neg E_2 \cap \cdots \neg E_n)\,.$$

As a lemma for this problem assume we have only two independent events E_1 and E_2 and consider

$$\begin{aligned} P(\neg E_1 \cap \neg E_2) &= P(\neg(E_1 \cup E_2)) = 1 - P(E_1 \cup E_2) \\ &= 1 - (P(E_1) + P(E_2) - P(E_1 \cap E_2)) \\ &= 1 - P(E_1) - P(E_2) + P(E_1)P(E_2) \\ &= (1 - P(E_1))(1 - P(E_2)) = P(\neg E_1)P(\neg E_2)\,, \end{aligned}$$

using the independence of the sets E_1 and E_2. This result shows that for independent events the product rule works for the negation of the sets as well as the direct sets themselves. Thus we have for the problem at hand that

$$\begin{aligned} P(E_1 \cup E_2 \cup \cdots E_n) &= 1 - P(\neg E_1 \cap \neg E_2 \cap \cdots \neg E_n) \\ &= 1 - P(\neg E_1)P(\neg E_2)P(\neg E_3) \cdots P(\neg E_n) \\ &= 1 - (1 - P(E_1))(1 - P(E_2))(1 - P(E_3)) \cdots (1 - P(E_n))\,, \end{aligned}$$

the required result.

Problem 7 (extinct fish)

Part (a): We desire to compute P_w the probability that the last ball drawn is white. This probability will be

$$P_w = \frac{n}{n+m},$$

because we have n white balls that can be selected from $n+m$ total balls that can be placed in the last spot.

Part (b): Let R be the event that the red fish species are the *first* species to become extinct. Then following the hint we write $P(R)$ as

$$P(R) = P(R|G_l)P(G_l) + P(R|B_l)P(B_l).$$

Here G_l is the event that the green fish species are the *last* fish species to become extinct and B_l the event that the blue fish species are the *last* fish species to become extinct. Now we conclude that

$$P(G_l) = \frac{g}{r+b+g},$$

and

$$P(B_l) = \frac{b}{r+b+g}.$$

We can see these by considering the blue fish as an example. If the blue fish are the last ones extinct then we have b possible blue fish to select from the $r+b+g$ total number of fish to be the last fish. Now we need to compute the conditional probabilities $P(R|G_l)$. This can be thought of as the event that the red fish go extinct and then the blue fish. This is the same type of experiment as in Part (a) of this problem in that we must have a blue fish go extinct (i.e. a draw with a blue fish last). This can happen with probability

$$\frac{b}{r+b+g-1},$$

where the denominator is one less than $r+b+g$ since the last fish drawn must be a green fish by the required conditioning. In the same way we have that

$$P(R|B_l) = \frac{g}{r+b+g-1}.$$

So that the total probability $P(R)$ is then given by

$$\begin{aligned} P(R) &= \left(\frac{b}{r+b+g-1}\right)\left(\frac{g}{r+b+g}\right) + \left(\frac{g}{r+b+g-1}\right)\left(\frac{b}{r+b+g}\right) \\ &= \frac{2bg}{(r+b+g-1)(r+b+g)}. \end{aligned}$$

Problem 8 (some inequalities)

Part (a): If $P(A|C) > P(B|C)$ and $P(A|C^c) > P(B|C^c)$, then consider $P(A)$ which by conditioning on C and C^c becomes

$$\begin{aligned} P(A) &= P(A|C)P(C) + P(A|C^c)P(C^c) \\ &> P(B|C)P(C) + P(B|C^c)P(C^c) = P(B)\,. \end{aligned}$$

Where the second line follows from the given inequalities.

Part (b): Following the hint, let C be the event that the sum of the pair of dice is 10, A the event that the first die lands on a 6 and B the event that the second die lands a 6. Then $P(A|C) = \frac{1}{3}$, and $P(A|C^c) = \frac{5}{36-3} = \frac{5}{33}$. So that $P(A|C) > P(A|C^c)$ as expected. Now $P(B|C)$ and $P(B|C^c)$ will have the same probabilities as for A. Finally, we see that $P(A \cap B|C) = 0$, while $P(A \cap B|C^c) = \frac{1}{33} > 0$. So we have found an example where $P(A \cap B|C) < P(A \cap B|C^c)$ and a counter example has been found.

Problem 9 (pairwise independence)

Let A be the event that the first toss lands heads and let B be the event that the second toss lands heads, and finally let C be the event that both lands on the same side. Now $P(A, B) = \frac{1}{4}$, and $P(A) = P(B) = \frac{1}{2}$, so A and B are independent. Now

$$P(A, C) = P(C|A)P(A) = \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4}\,.$$

but $P(C) = \frac{1}{2}$ so $P(A, C) = P(A)P(C)$ and A and C are independent. Finally

$$P(B, C) = P(C|B)P(B) = \frac{1}{4}\,,$$

so again B and C are independent. Thus A, B, and C are pairwise independent but for three sets to be fully independent we must have in addition that

$$P(A, B, C) = P(A)P(B)P(C)\,.$$

The right hand side of this expression is $\left(\frac{1}{2}\right)^3$ while the left hand side is the event that both tosses land heads and so $P(A, B, C) = \frac{1}{4} \neq P(A)P(B)P(C)$ and the three sets are not independent.

Problem 10 (pairwise independence does not imply independence)

Let $A_{i,j}$ be the event that person i and j have the same birthday. We desire to show that these events are pairwise independent. That is the two events $A_{i,j}$ and $A_{r,s}$ are independent but the totality of all $\left(\begin{array}{c} n \\ 2 \end{array} \right)$ events are not independent. Now

$$P(A_{i,j}) = P(A_{r,s}) = \frac{1}{365},$$

since for the specification of either one persons birthday the probability that the other person will have that birthday is 1/365. Now

$$P(A_{i,j} \cap A_{r,s}) = P(A_{i,j}|A_{r,s})P(A_{r,s}) = \left(\frac{1}{365}\right)\left(\frac{1}{365}\right) = \frac{1}{365^2}.$$

This is because $P(A_{i,j}|A_{r,s}) = P(A_{i,j})$ i.e. the fact that people r and s have the same birthday has no effect on whether the event $A_{i,j}$ is true. This is true even if one of the people in the pairs (i,j) and (r,s) is the same. When we consider the intersection of *all* the sets $A_{i,j}$, the situation changes. This is because the event $\cap_{(i,j)} A_{i,j}$ (where the intersection is over all pairs (i,j)) is the event that *every* pair of people have the same birthday, i.e. that everyone has the same birthday. This will happen with probability

$$\left(\frac{1}{365}\right)^{n-1},$$

while if the events $A_{i,j}$ were independent the required probability would be

$$\prod_{(i,j)} P(A_{i,j}) = \left(\frac{1}{365}\right)^{\left(\begin{array}{c} n \\ 2 \end{array} \right)} = \left(\frac{1}{365}\right)^{\frac{n(n-1)}{2}}.$$

Since $\left(\begin{array}{c} n \\ 2 \end{array} \right) \neq n-1$, these two results are not equal and the totality of events $A_{i,j}$ are not independent.

Problem 11 (at least one head)

The probability that we obtain at least on head is one minus the probability that we obtain all tails in n flips. Thus this probability is $1-(1-p)^n$. If this is to be made greater than $\frac{1}{2}$ we have

$$1-(1-p)^n > \frac{1}{2},$$

or solving for n we have $n > \frac{\ln(1/2)}{\ln(1-p)}$, so since the this expression can be non-integer take n we need to take the next integer larger than or equal to this number. That is take n to be

$$n = \left\lceil \frac{\ln(1/2)}{\ln(1-p)} \right\rceil.$$

Problem 12 (an infinite sequence of flips)

Let a_i be the probability that the ith coin lands heads. The consider the random variable N, specifying the location where the first head occurs. This problem then is like a geometric random variable where we want to determine the first time a success occurs. Then we have for a distribution of $P\{N\}$ the following

$$P\{N = n\} = a_n \prod_{i=1}^{n-1} (1 - a_i)\,.$$

This states that the first $n-1$ flips must land tails and the last flip (the nth) then lands heads. Then when we add this probability up for $n = 1, 2, 3, \cdots, \infty$ i.e.

$$\sum_{n=1}^{\infty} \left[a_n \prod_{i=1}^{n-1} (1 - a_i) \right]\,,$$

is the probability that a head occurs *somewhere* in the infinite sequence of flips. The other possibility would be for a head to *never* appear. This will happen with a probability of

$$\prod_{i=1}^{\infty} (1 - a_i)\,.$$

Together these two expressions consist of all possible outcomes and therefore must sum to one. Thus we have proven the identity

$$\sum_{n=1}^{\infty} \left[a_n \prod_{i=1}^{n-1} (1 - a_i) \right] + \prod_{i=1}^{\infty} (1 - a_i) = 1\,,$$

or the desired result.

Problem 13 (winning by flipping)

Let $P_{n,m}$ be the probability that A who starts the game accumulates n head before B accumulates m heads. We can evaluate this probability by conditioning on the outcome of the first flip made by A. If this flip lands heads, then A has to get $n-1$ more flips before B's obtains m. If this flip lands tails then B obtains control of the coin and will receive m flips before A receives n with probability $P_{m,n}$ by the implicit symmetry in the problem. Thus A will accumulate the correct number of heads with probability $1 - P_{m,n}$. Putting these two outcomes together (since they are the mutually exclusive and exhaustive) we have

$$P_{n,m} = pP_{n-1,m} + (1 - p)(1 - P_{m,n})\,,$$

or the desired result.

Problem 14 (gambling against the rich)

Let P_i be the probability you eventually go broke when your initial fortune is i. Then conditioning on the result of the first wager we see that P_i satisfies the following difference equation

$$P_i = pP_{i+1} + (1-p)P_{i-1}\,.$$

This simply states that the probability you go broke when you have a fortune of i is p times P_{i+1} if you win the first wager (since if you win the first wager you now have $i+1$ as your fortune) *plus* $1-p$ times P_{i-1} if you loose the first wager (since if you loose the first wager you will have $i-1$ as your fortune). To solve this difference equation we recognize that its solution must be given in terms of a constant raised to the ith power i.e. α^i. Using the ansatz that $P_i = \alpha^i$ and inserting this into the above equation we find that α must satisfy the following

$$p\alpha^2 - \alpha + (1-p) = 0\,.$$

Using the quadratic equation to solve this equation for α we find α given by

$$\alpha = \frac{1 \pm \sqrt{1-4p(1-p)}}{2p} = \frac{1 \pm \sqrt{(2p-1)^2}}{2p} = \frac{1 \pm (2p-1)}{2p}\,.$$

Taking the plus sign gives $\alpha^+ = 1$, while taking the minus sign in the above gives $\alpha^- = \frac{q}{p}$. Now the general solution to this difference equation is then given by

$$P_i = C_1 + C_2\left(\frac{q}{p}\right)^i \quad \text{for} \quad i \geq 0\,.$$

Problem 15 (n trials and r successes)

The event that we want the probability of is the sum of independent events where in each of these events we have $r-1$ success in the $n-1$ trials and the n trial is also a success (and the last one needed). Based on combining a binomial probability with this last success, the probability we seek is then

$$\left(\binom{n-1}{r-1}p^{r-1}(1-p)^{n-1-(r-1)}\right)p = \binom{n-1}{r-1}p^r(1-p)^{n-r}\,.$$

Recall that the problem of the points is the situation where we obtain a success with probability p and we want the probability we have n successes before m failures. To place the problem of the points in the framework of this problem we can consider extending the number of trials we perform in such a way that we always perform $n+m$ games. Since the probability we want to compute is the probability that from these total $n+m$ games we get n success before m failures this event must be one of the following independent events

- We get our n successes in the first n games and the remaining m games are all failures.
- We get our n successes in the first $n+1$ games (so have one failure) and the remaining $m-1$ games are all failures.

- We get our n successes in the first $n+2$ games (and thus have two failures) and the remaining $m-2$ are all failures.
- etc.
- We get our n successes in the first $n+m-2$ games (and thus have $m-2$ failures) and the remaining 2 games are all failures
- We get our n successes in the first $n+m-1$ games (and thus have $m-1$ failures) and the remaining 1 game is a failure

We have computed the probability of each of these events the first part of this problem. Thus by adding the contributions from each independent event we have

$$\sum_{i=n}^{n+m-1} \binom{i-1}{n-1} p^n (1-p)^{i-n} .$$

Note: I'm not sure how to show that this is equivalent to the Fermat's solution which is

$$\sum_{i=n}^{n+m-1} \binom{m+n-1}{i} p^i (1-p)^{m+n-1-i} ,$$

if anyone knows how to show this please contact me.

Problem 16 (the probability of an even number of successes)

Let P_n be the probability that n Bernoulli trials result in an even number of successes. Then the given difference equation can be obtained by conditioning on the result of the first trial as follows. If the first trial is a success then we have $n-1$ trials to go and to obtain an even *total* number of tosses we want the number of successes in this $n-1$ trials to be *odd* This occurs with probability $1-P_{n-1}$. If the first trial is a failure then we have $n-1$ trials to go and to obtain an even total number of tosses we want the number of successes in this $n-1$ trials to be *even* This occurs with probability P_{n-1}. Thus we find that

$$P_n = p(1-P_{n-1}) + (1-p)P_{n-1} \quad \text{for} \quad n \geq 1 .$$

Some special point cases are easily computed. We have by assumption that $P_0 = 1$, and $P_1 = q$ since with only one trial, this trial must be a failure to get a total even number of successes. Given this difference equation and a potential solution we can verify that this solution satisfies our equation and therefore know that it is a solution. One can easily check that the given P_n satisfies $P_0 = 1$ and $P_1 = q$. In addition, for the given assumed solution we have that

$$P_{n-1} = \frac{1+(1-2p)^{n-1}}{2} ,$$

From which we find (using this expression in the right hand side of the difference equation above)

$$\begin{aligned} p(1-P_{n-1}) + (1-p)P_{n-1} &= p + (1-2p)P_{n-1} \\ &= p + (1-2p)\left(\frac{1+(1-2p)^{n-1}}{2}\right) \\ &= p + \frac{1-2p}{2} + \frac{(1-2p)^n}{2} \\ &= \frac{1}{2} + \frac{(1-2p)^n}{2} = P_n\,. \end{aligned}$$

Showing that P_n is a solution the given difference equation.

Problem 17 (odd number of successes)

Let S_i be the event the ith trial results in success and let E_n be the event the number of successes in n total trials is odd.

Part (a): The probability $P(E_1)$ is the probability the first trial is a success so we have

$$P(E_1) = P(S_1) = \frac{1}{2(1)+1} = \frac{1}{3}\,.$$

The probability $P(E_2)$ is the union of the independent events where the first trial is a success and the second trial is a failure or the first trial is a failure and the second trial is a success. Thus

$$\begin{aligned} E_2 &= S_1S_2^c \cup S_1^cS_2 \quad \text{so} \\ P(E_2) &= P(S_1S_2^c) + P(S_1^cS_2) = (\frac{1}{3})(\frac{4}{5}) + (\frac{2}{3})(\frac{1}{5}) = \frac{6}{(3)(5)} = \frac{2}{5} \end{aligned}$$

The probability $P(E_3)$ is the union of the independent events where we have three total successes or one success and two failures. Thus

$$\begin{aligned} E_3 &= S_1S_2S_3 \cup S_1S_2^cS_3^c \cup S_1^cS_2S_3^c \cup S_1^cS_2S_3 \quad \text{so} \\ P(E_3) &= P(S_1S_2S_3) + P(S_1S_2^cS_3^c) + P(S_1^cS_2S_3^c) + P(S_1^cS_2^cS_3) \\ &= (\frac{1}{3})(\frac{1}{5})(\frac{1}{7}) + (\frac{1}{3})(\frac{4}{5})(\frac{6}{7}) + (\frac{2}{3})(\frac{1}{5})(\frac{6}{7}) + (\frac{2}{3})(\frac{4}{5})(\frac{1}{7}) = \frac{45}{(3)(5)(7)} = \frac{3}{7}\,. \end{aligned}$$

In the same way the probability $P(E_4)$ is the union of the independent events where we have three successes with one failure or one success with three failures. Thus the the probability $P(E_4)$ is given by

$$\begin{aligned} P(E_4) &= P(S_1^cS_2S_3S_4) + P(S_1S_2^cS_3S_4) + P(S_1S_2S_3^cS_4) + P(S_1S_2S_3S_4^c) \\ &+ P(S_1S_2^cS_3^cS_4^c) + P(S_1^cS_2S_3^cS_4^c) + P(S_1^cS_2^cS_3S_4^c) + P(S_1^cS_2^cS_3^cS_4) \\ &= (\frac{2}{3})(\frac{1}{5})(\frac{1}{7})(\frac{1}{9}) + (\frac{1}{3})(\frac{4}{5})(\frac{1}{7})(\frac{1}{9}) + (\frac{1}{3})(\frac{1}{5})(\frac{6}{7})(\frac{1}{9}) + (\frac{1}{3})(\frac{1}{5})(\frac{1}{7})(\frac{8}{9}) \\ &+ (\frac{1}{3})(\frac{4}{5})(\frac{6}{7})(\frac{8}{9}) + (\frac{2}{3})(\frac{1}{5})(\frac{6}{7})(\frac{8}{9}) + (\frac{2}{3})(\frac{4}{5})(\frac{1}{7})(\frac{8}{9}) + (\frac{2}{3})(\frac{4}{5})(\frac{6}{7})(\frac{1}{9}) \\ &= \frac{420}{(3)(5)(7)(9)} = \frac{4}{9}\,. \end{aligned}$$

Finally, the probability $P(E_5)$ is the union of the independent events where we have all success, three successes with two failures or one success with four failures. Thus the the probability $P(E_5)$ is given by

$$\begin{aligned}
P(E_5) &= P(S_1S_2S_3S_4S_5) \\
&+ P(S_1^cS_2^cS_3S_4S_5) + P(S_1^cS_2S_3^cS_4S_5) + P(S_1^cS_2S_3S_4^cS_5) \\
&+ P(S_1^cS_2S_3S_4S_5^c) + P(S_1S_2^cS_3^cS_4S_5) + P(S_1S_2^cS_3S_4^cS_5) \\
&+ P(S_1S_2^cS_3S_4S_5^c) + P(S_1S_2S_3^cS_4^cS_5) + P(S_1S_2S_3^cS_4S_5^c) \\
&+ P(S_1S_2S_3S_4^cS_5^c) + P(S_1S_2^cS_3^cS_4^cS_5^c) + P(S_1^cS_2S_3^cS_4^cS_5^c) \\
&+ P(S_1^cS_2^cS_3S_4^cS_5^c) + P(S_1^cS_2^cS_3^cS_4S_5^c) + P(S_1^cS_2^cS_3^cS_4^cS_5) \\
&= (\frac{1}{3})(\frac{1}{5})(\frac{1}{7})(\frac{1}{9})(\frac{1}{11}) \\
&+ (\frac{2}{3})(\frac{4}{5})(\frac{1}{7})(\frac{1}{9})(\frac{1}{11}) + (\frac{2}{3})(\frac{1}{5})(\frac{6}{7})(\frac{1}{9})(\frac{1}{11}) + (\frac{2}{3})(\frac{1}{5})(\frac{1}{7})(\frac{8}{9})(\frac{1}{11}) \\
&+ (\frac{2}{3})(\frac{1}{5})(\frac{1}{7})(\frac{1}{9})(\frac{10}{11}) + (\frac{1}{3})(\frac{4}{5})(\frac{6}{7})(\frac{1}{9})(\frac{1}{11}) + (\frac{1}{3})(\frac{4}{5})(\frac{1}{7})(\frac{8}{9})(\frac{1}{11}) \\
&+ (\frac{1}{3})(\frac{4}{5})(\frac{1}{7})(\frac{1}{9})(\frac{10}{11}) + (\frac{1}{3})(\frac{1}{5})(\frac{6}{7})(\frac{8}{9})(\frac{1}{11}) + (\frac{1}{3})(\frac{1}{5})(\frac{6}{7})(\frac{1}{9})(\frac{10}{11}) \\
&+ (\frac{1}{3})(\frac{1}{5})(\frac{1}{7})(\frac{8}{9})(\frac{10}{11}) + (\frac{1}{3})(\frac{4}{5})(\frac{6}{7})(\frac{8}{9})(\frac{10}{11}) + (\frac{2}{3})(\frac{1}{5})(\frac{6}{7})(\frac{8}{9})(\frac{10}{11}) \\
&+ (\frac{2}{3})(\frac{4}{5})(\frac{1}{7})(\frac{8}{9})(\frac{10}{11}) + (\frac{2}{3})(\frac{4}{5})(\frac{6}{7})(\frac{1}{9})(\frac{10}{11}) + (\frac{2}{3})(\frac{4}{5})(\frac{6}{7})(\frac{8}{9})(\frac{1}{11}) \\
&= \frac{4725}{(3)(5)(7)(9)(11)} = \frac{5}{11}\,.
\end{aligned}$$

Part (b): From the above expressions we hypothesize that

$$P_n \equiv P(E_n) = \frac{n}{2n+1}\,.$$

Part (c): Now P_n is the probability that we have an odd number of successes in n trials and P_{n-1} must be the probability we have an odd number of success in $n-1$ trials. We can compute P_n in terms of P_{n-1} by conditioning on the result of the last trial. If the result of the nth trial is success, the number of successes in the total n trials will be odd if and only if the number of successes in the first $n-1$ trials is even. This last event happens with probability $1 - P_{n-1}$. At the same time if the result of the nth trial is failure, the number of successes in n trials will be odd if and only if the number of successes in the first $n-1$ trials is odd. Thus we have just argued that

$$\begin{aligned}
P_n &= \left(\frac{1}{2n+1}\right)(1 - P_{n-1}) + \left(1 - \frac{1}{2n+1}\right)P_{n-1} \\
&= \Big(\frac{1}{2n+1}\Big)(1 - P_{n-1}) + \Big(\frac{2n}{2n+1}\Big)P_{n-1} \\
&= \frac{1}{2n+1} + \Big(\frac{2n-1}{2n+1}\Big)P_{n-1}\,.
\end{aligned} \tag{19}$$

Part (d): For this part we want to show that $P_n = \frac{n}{2n+1}$ is a solution to the above difference equation. Note that

$$P_{n-1} = \frac{n-1}{2(n-1)+1} = \frac{n-1}{2n-1}\,.$$

Thus the right-hand-side of Equation 19 gives

$$\begin{aligned}
\text{RHS} &= \left(\frac{1}{2n+1}\right)\left(1 - \frac{n-1}{2n-1}\right) + \left(\frac{2n}{2n+1}\right)\left(\frac{n-1}{2n-1}\right) \\
&= \frac{1}{2n+1} - \left(\frac{1}{2n+1}\right)\left(\frac{n-1}{2n-1}\right) + \left(\frac{2n}{2n+1}\right)\left(\frac{n-1}{2n-1}\right) \\
&= \frac{2n-1}{(2n+1)(2n-1)} - \frac{n-1}{(2n+1)(2n-1)} + \frac{2n(n-1)}{(2n+1)(2n-1)} \\
&= \frac{n+2n^2-2n}{(2n+1)(2n-1)} = \frac{n(2n-1)}{(2n+1)(2n-1)} = \frac{n}{2n+1}\,,
\end{aligned}$$

which is the same as the functional form for P_n showing the equivalence.

Problem 18 (3 consecutive heads)

Let Q_n be the probability that in n tosses of a fair coin no run of three consecutive heads appears. Since when $n = 0, 1, 2$ is is not possible to flip three heads we have $Q_0 = Q_1 = Q_2 = 1$. We can compute Q_n by conditioning on the result of the first few flips. If the first flip is a tail then we cannot have this flip as the beginning of a sequence of three consecutive heads and the probability that no run appears is the probability that none appear in the remaining $n-1$ flips which is the value Q_{n-1}. If the first flip is in fact a head then we need to consider what happens in the next flips. If the second flip is then a tail this first flip cannot be in a run of heads and the probability that no run of three heads occur in the next $n-2$ flips is Q_{n-2}. If the second flip is a head then again we need to look at the next flip. If is a tail the same logic above applies and probability that no run of three heads occur in the next $n-3$ flips is Q_{n-3}. If in fact the third flip is a head then we *do* have a run of heads and the probability of no run of heads is 0. When we combine the above with the probability of the individual flips, we can summarize this discussion as the probability of interest Q_n can be decomposed into the following independent events and their probabilities:

- $T\cdots$ with a probability of $\frac{1}{2}Q_{n-1}$
- $HT\cdots$ with a probability of $\frac{1}{4}Q_{n-2}$
- $HHT\cdots$ with a probability of $\frac{1}{8}Q_{n-3}$
- $HHH\cdots$ with a probability of 0.

Here the $\cdots$ notation represents unrealized coin flips. Thus we can add these probabilities to get Q_n and we have shown

$$Q_n = \frac{1}{2}Q_{n-1} + \frac{1}{4}Q_{n-2} + \frac{1}{8}Q_{n-3}\,.$$

Given the initial conditions on the first three values of Q_n of can use the above to compute Q_8.

Problem 19 (the n-game gambler's ruin)

With the probability p that A wins (and in a slightly different notation) by conditioning on the outcome of the ith flip we have

$$P(n,i) = pP(n-1,i+1) + (1-p)P(n-1,i-1)\,.$$

Next note that if $i = N$, then gambler A has all the money and must certainly win and we have $P(n,N) = 1$. In the same way if $i = 0$, gambler B has all the money and A must certainly loose and we have $P(n,0) = 0$. If $n = 0$, then they stop playing and A cannot win. Thus we take $P(0,i) = 0$ for $0 \le i < N$. To find the probability of interest let $N = 5$, $i = 3$, $n = 7$. Using the above relationship we find some results we will need later

$$\begin{aligned}
P(1,3) &= pP(0,4) + qP(0,2) = 0 \\
P(1,1) &= pP(0,2) + qP(0,0) = 0 \\
\\
P(2,4) &= pP(1,5) + qP(1,3) = p \\
P(2,2) &= pP(1,3) + qP(1,1) = 0 \\
\\
P(3,3) &= pP(2,4) + qP(2,2) = p^2 \\
P(3,1) &= pP(2,2) + qP(2,0) = 0 \\
\\
P(4,4) &= pP(3,5) + qP(3,3) = p + qp^2 \\
P(4,2) &= pP(3,3) + qP(3,1) = p^3\,,
\end{aligned}$$

and also

$$\begin{aligned}
P(5,3) &= pP(4,4) + qP(4,2) = p(p+qp^2) + qp^3 = p^2 + 2qp^3 \\
P(5,1) &= pP(4,2) + qP(4,0) = p^4 \\
\\
P(6,4) &= pP(5,5) + qP(5,3) = p + q(p^2 + 2qp^3) = p + qp^2 + 2q^2p^3 \\
P(6,2) &= pP(5,3) + qP(5,1) = p(p^2 + 2qp^3) + qp^4 = p^3 + 3qp^4 \\
\\
P(7,3) &= pP(6,4) + qP(6,2) \\
&= p(p + qp^2 + 2q^2p^3) + q(p^3 + 3qp^4) = p^2 + 2qp^3 + 5q^2p^4\,.
\end{aligned}$$

Problem 20 (white and black balls, 2 urns)

With probability α a ball is chosen from the first urn. All subsequent selections are made based on the fact that if a black ball is drawn we "switch" urns and begin to draw from the other/alternative urn. Let α_n be the probability that the nth ball is drawn from the first urn. We take $\alpha_1 = \alpha$. The to calculate α_{n+1} we can condition on whether the nth ball was drawn from the first urn or not. If it was, then with probability p we would draw a while ball from that urn and we would not have switch urns. Thus ball $n+1$ is from urn number 1. If it was not, then with probability $1-p'$ we would have drawn a black ball from the second urn and would have to switch urns on the $n+1$st draw to the first urn. Thus

$$\alpha_{n+1} = p\alpha_n + (1-p')(1-\alpha_n) = \alpha_n(p+p'-1) + 1 - p' \,. \tag{20}$$

For $n \geq 1$ and with $\alpha_1 = \alpha$. The solution to this recursion relationship is given by

$$\alpha_n = \frac{1-p'}{2-p-p'} + \left(\alpha - \frac{1-p'}{2-p-p'}\right)(p+p'-1)^{n-1} \,. \tag{21}$$

To prove this not that when n = 1 we get $\alpha_1 = \alpha$ as we should in the above. Assume the functional form for α_n given by Equation 21 holds up to some n. Then put that expression into the right-hand-side of Equation 20. We get

$$\begin{aligned}
\text{RHS} &= \left[\frac{1-p'}{2-p-p'} + \left(\alpha - \frac{1-p'}{2-p-p'}\right)(p+p'-1)^{n-1}\right](p+p'-1) + 1 - p' \\
&= \frac{1-p'}{2-p-p'}(p+p'-1) + \left(\alpha - \frac{1-p'}{2-p-p'}\right)(p+p'-1)^{n} + 1 - p' \\
&= (1-p')\left(\frac{p+p'-1}{2-p-p'} + 1\right) + \left(\alpha - \frac{1-p'}{2-p-p'}\right)(p+p'-1)^{n} \\
&= (1-p')\left(\frac{p+p'-1+2-p-p'}{2-p-p'}\right) + \left(\alpha - \frac{1-p'}{2-p-p'}\right)(p+p'-1)^{n} \\
&= \frac{1-p'}{2-p-p'} + \left(\alpha - \frac{1-p'}{2-p-p'}\right)(p+p'-1)^{n} \,,
\end{aligned}$$

which is Equation 21 evaluated at $n+1$ proving the expression.

Next Let P_n be the probability that the nth ball selected is white. This depends on the urn from which we are selecting from. We have

$$P_n = p\alpha_n + p'(1-\alpha_n) = p' + (p-p')\alpha_n \quad \text{for} \quad n \geq 2 \,.$$

From what we know about α_n given in Equation 21 we have a solution for P_n.

To calculate the requested limits note that if $0 < p, p' < 1$, then

$$0 < p + p' < 2, \quad \text{thus} \quad -1 < p + p' - 1 < 1 \,.$$

Using this we have

$$\lim_{n\to\infty} (p+p'-1)^n = 0 \,,$$

and we have

$$\lim_{n\to\infty} \alpha_n = \frac{1-p'}{2-p-p'} = \frac{1-p'}{(1-p)+(1-p')} \,.$$

For P_n we have

$$\lim_{n\to\infty} P_n = p' + \frac{1-p'}{(1-p)+(1-p')}(p-p') = \frac{p(1-p')+p'(1-p)}{(1-p)+(1-p')} \,.$$

Problem 21(counting votes)

Part (a): We are told for this problem that candidate A receives n votes and candidate B receives m votes where we assume that $n > m$. Lets represent the each sequential vote by a sequence of A's and B's. The number of AB sequences is given by

$$\frac{(n+m)!}{n!\,m!} \tag{22}$$

We will assume all sequences are equally likely. Let $E(n, m)$ be the event A is always ahead in the counting of votes and $P_{n,m}$ the events probability.

To calculate $P_{2,1}$ the only sequence where A leads B for all counts is $\{AAB\}$. Using Equation 22 we then have $P_{2,1} = \frac{1}{3}$.

To calculate $P_{3,1}$ the sequences where A leads B for all counts are: $\{AAAB, AABA\}$. Using Equation 22 we then have $P_{3,1} = \frac{2}{4} = \frac{1}{2}$.

To calculate $P_{4,1}$ the sequences where A leads B for all counts are: $\{AAAAB, AAABA, AABAA\}$. Using Equation 22 we then have $P_{4,1} = \frac{3}{5}$.

To calculate $P_{3,2}$ the sequences where A leads B for all counts are: $\{AAABB, AABAB\}$. Using Equation 22 we have $\frac{(3+2)!}{3!2!} = 10$ and thus $P_{3,2} = \frac{1}{5}$.

To calculate $P_{4,2}$ the sequences where A leads B for all counts are:

$$\{AAAABB, AAABAB, AAABBA, AABAAB, AABABA\}\,.$$

Using Equation 22 we have $\frac{(4+2)!}{4!2!} = 15$ and thus $P_{4,2} = \frac{5}{15} = \frac{1}{3}$.

To calculate $P_{4,3}$ the sequences where A leads B for all counts are:

$$\{AAAABBB, AAABABB, AAABBAB, AABAABB, AABABAB\}$$

Using Equation 22 again we find $P_{4,3} = \frac{5}{35} = \frac{1}{7}$.

Part (b): For $P_{n,1}$ note that from Equation 22 there are $\frac{(n+1)!}{n!\,1!} = n + 1$ sequences. Each sequence has one vote for B. Each of these sequences start with the characters AA, AB or B. Note that the sequences that start with B or AB will not have A leading in votes for all time, while for the sequences that start with AA the number of A votes will always be larger than the number of votes for B by at least 1. The number of sequences starting with AA (since we have specified that the first two votes are for A we have $n + 1 - 2$ votes to yet place) is

$$\frac{(n+1-2)!}{(n-2)!\,1!} = n - 1\,.$$

Thus $P_{n,1} = \frac{n-1}{n+1}$.

For $P_{n,2}$ note that there are $\frac{(n+2)!}{n!\,2!} = \frac{(n+2)(n+1)}{2}$ sequences of A and B's and each sequence will have two B's. All our sequences start with either AAA, $AABA$, $AABB$, AB or B. The only of these sequences where A is leading for all counts are AAA and $AABA$. We thus need to count the number of sequences starting with each of these prefixes. The number of sequences starting with AAA (since now we have $n+2-3$ votes to distribute) is

$$\frac{(n+2-3)!}{(n-3)!2!} = \frac{(n-1)(n-2)}{2},$$

while the number of sequences starting with $AABA$ (since now we have $n+2-4$ votes to distributed) is

$$\frac{(n+2-4)!}{(n-3)!1!} = n-2.$$

The total number of sequences where A is leading in votes for all time is then the sum of these two events or

$$\frac{(n-1)(n-2)}{2} + n - 2 = \frac{(n+1)(n-2)}{2}$$

Using this we find

$$P_{n,2} = \frac{\frac{(n+1)(n-2)}{2}}{\frac{(n+2)(n+1)}{2}} = \frac{n-2}{n+2}.$$

Part (c): It looks like the pattern for $P_{n,m}$ is $P_{n,m} = \frac{n-m}{n+m}$.

Part (d): By conditioning on who received the last vote (either A or B) we find

$$P_{n,m} = \left(\frac{n}{n+m}\right) P_{n-1,m} + \left(\frac{m}{n+m}\right) P_{n,m-1}. \tag{23}$$

Part (e): We will show that the fraction $P_{n,m} = \frac{n-m}{n+m}$ satisfies the right-hand-side of Equation 23 and when simplified gives the left-hand-side of the equation.

$$\begin{aligned} \text{RHS} &= \left(\frac{n}{n+m}\right)\left(\frac{n-1-m}{n-1+m}\right) + \left(\frac{m}{n+m}\right)\left(\frac{n-m+1}{n+m-1}\right) \\ &= \frac{n^2 - n - nm + mn - m^2 + m}{(n+m)(n+m-1)} = \frac{(n-m)(n+m-1)}{(n+m)(n+m-1)} = \frac{n-m}{n+m} = \text{LHS}. \end{aligned}$$

Problem 22 (is it rainy or dry?)

Let D_n be the event the weather is dry on day n and the weather (either wet or dry) tomorrow will be the same as the weather today with probability p (so that it is different weather with a probability of $1-p$). This means that

$$P(D_n|D_{n-1}) = P(D_n^c|D_{n-1}^c) = p.$$

We are told that the weather is dry on day 0 which means that $P(D_0) = 1$, and let $n \geq 1$. Then it can become dry on day n in two ways. If it was dry on the day $n-1$ and it stays dry, or it was wet on day $n-1$ and it became dry. Conditioning on what the weather was yesterday (dry or wet) and the necessary transition we have

$$\begin{aligned} P(D_n) &= P(D_{n-1})P(D_n|D_{n-1}) + P(D_{n-1}^c)P(D_n|D_{n-1}^c) \\ &= P(D_{n-1})p + P(D_{n-1}^c)(1-p) \\ &= P(D_{n-1})p + (1-P(D_{n-1}))(1-p) \\ &= P(D_{n-1})(2p-1) + (1-p) \quad \text{for} \quad n \geq 1 \,. \end{aligned}$$

We next want to show that the solution to the above recurrence is given by

$$P(D_n) = \frac{1}{2} + \frac{1}{2}(2p-1)^n , \quad n \geq 0 \,. \tag{24}$$

We will show that by induction on n. First for $n = 0$ we have

$$P(D_0) = \frac{1}{2} + \frac{1}{2}(2p-1)^0 = 1 \,,$$

as it should. Let $n \geq 1$ and assume that Equation 24 is true for $n-1$. We then have

$$\begin{aligned} P(D_n) &= P(D_{n-1})(2p-1) + (1-p) \\ &= \left(\frac{1}{2} + \frac{1}{2}(2p-1)^{n-1}\right)(2p-1) + (1-p) \\ &= \frac{1}{2}(2p-1) + \frac{1}{2}(2p-1)^n + (1-p) = \frac{1}{2} + \frac{1}{2}(2p-1)^n \,, \end{aligned}$$

showing that Equation 24 is true for n as well.

Problem 24 (round robin tournaments)

In this problem we specify an integer k and then ask whether it is possible for every set of k players to have there exist a member from the *other* $n-k$ players that beat these k players when competing against these k. To show that this is possible if the given inequality is true, we follow the hint. In the hint we enumerate the $\binom{n}{k}$ sets of k players and let B_i be the event that *no* of the other $n-k$ contestant beats every one of the k players in the i set of k. Then $P(\cup_i B_i)$ is the probability that at least one of the subsets of size k has no external player that beats everyone. Then $1 - P(\cup_i B_i)$ is the probability that every subset of size k has an external player that beats everyone. Since this is the event we want to be possible we desire that

$$1 - P(\cup_i B_i) > 0 \,,$$

or equivalently

$$P(\cup_i B_i) < 1 \,.$$

Now Boole's inequality states that $P(\cup_i B_i) \leq \sum_i P(B_i)$, so if we pick our k such that $\sum_i P(B_i) < 1$, we will necessarily have $P(\cup_i B_i) < 1$ possible. Thus we will focus on ensuring that $\sum_i P(B_i) < 1$.

Lets now focus on evaluating $P(B_i)$. Since this is the probability that no contestant from outside the ith cluster beats all players inside, we can evaluate it by considering a particular player outside the k member set. Denote the other player by X. Then X would beat all k members with probability $\left(\frac{1}{2}\right)^k$, and thus with probability $1 - \left(\frac{1}{2}\right)^k$ does *not* beat all players in this set. As the set B_i, requires that all $n-k$ players *not* beat the k players in this ith set, each of the $n-k$ exterior players must fail at beating the k players and we have

$$P(B_i) = \left(1 - \left(\frac{1}{2}\right)^k\right)^{n-k}.$$

Now $P(B_i)$ is in fact independent of i (there is no reason it should depend on the particular subset of players) we can factor this result out of the sum above and simply multiply by the number of terms in the sum which is $\left(\begin{array}{c} n \\ k \end{array}\right)$ giving the requirement for possibility of

$$\left(\begin{array}{c} n \\ k \end{array}\right)\left(1 - \left(\frac{1}{2}\right)^k\right)^{n-k} < 1,$$

as was desired to be shown.

Problem 25 (a direct proof of conditioning)

Consider $P(E|F)$ which is equal to $P(E,G|F) + P(E,G^c|F)$, since the events (E,G) and (E,G^c) are independent. Now these component events can be simplified as

$$\begin{aligned} P(E,G|F) &= P(E|F,G)P(G|F) \\ P(E,G^c|F) &= P(E|F,G^c)P(G^c|F), \end{aligned}$$

and the above becomes

$$P(E|F) = P(E|F,G)P(G|F) + P(E|F,G^c)P(G^c|F),$$

as expected.

Problem 26 (conditional independence)

Equations 5.11 and 5.12 from the book are two equivalent statements of conditional independence. Namely E_1 and E_2 are conditionally independent given F occurs if

$$P(E_1|E_2,F) = P(E_1|F), \tag{25}$$

or equivalently

$$P(E_1E_2|F) = P(E_1|F)P(E_2|F). \tag{26}$$

To prove this equivalence, consider the left-hand-side of Equation 26. We have when we use Equation 25 in the second step that

$$P(E_1E_2|F) = P(E_1|E_2F)P(E_2|F) = P(E_1|F)P(E_2|F).$$

Proving the equivalence.

Problem 27 (extension of conditional independence)

A set $\{E_i : i \in I\}$ is said to be *conditionally independent*, given F, if for every finite subset of events $\{E_{i_k} : 1 \le k \le n\}$ with two or more members ($n \ge 2$) we have

$$P\left(\bigcap_{k=1}^{n} E_{i_k} \mid F\right) = \prod_{k=1}^{n} P(E_{i_k} \mid F)$$

Problem 28 (does independence imply conditional independence)

This statement is false and can be shown with the following example. Consider the event where we toss two fair coins. The outcomes are thus $\{HH, HT, TH, TT\}$. Then let H_1 be the event the first coin lands heads up and let H_2 be the event the second coin lands heads up. Then we have the events H_1, H_2, and H_1H_2 given by

$$H_1 = \{HH, HT\}\,, \quad H_2 = \{HH, TH\}\,, \quad H_1H_2 = \{HH\}\,.$$

Thus we have $P(H_1) = P(H_2) = \frac{1}{2}$, and $P(H_1H_2) = \frac{1}{4}$. Now let F be the event the coins land the same way. Then this event F is given by $F = \{HH, TT\}$ and we have $H_1F = \{HH\}$, $H_2F = \{HH\}$, with $H_1H_2F = \{HH\}$. From these we can calculate probabilities. We have $P(F) = \frac{1}{2}$, and $P(H_1F) = P(H_2F) = P(H_1H_2F) = \frac{1}{4}$. We know that H_1 and H_2 are independent events from

$$P(H_1H_2) = \frac{1}{4} = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = P(H_1)P(H_2)\,.$$

With condition probabilities given by

$$\begin{aligned}
P(H_1H_2|F) &= \frac{P(H_1H_2F)}{P(F)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} \\
P(H_1|F) &= \frac{P(H_1F)}{P(F)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} \\
P(H_2|F) &= \frac{P(H_2F)}{P(F)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}\,.
\end{aligned}$$

Thus

$$P(H_1H_2|F) = \frac{1}{2} \neq \frac{1}{4} = (\frac{1}{2})(\frac{1}{2}) = P(H_1|F)P(H_2|F)\,,$$

and we have that H_1 and H_2 are not conditionally independent given F even thought they are independent.

Problem 30 (extensions on Laplace's rule of succession)

In Laplace's rule of succession we assume we have $k+1$ coins, the ith one of which yields heads when flipped with probability $\frac{i}{k}$. Then from this version of the experiment the first n

flips of the chosen coin results in r heads and $n-r$ tails. Let H denote the event that the $n+1$ flip will land heads. then conditioning on the chosen coin C_i for $0 \le i \le k$ we have the following

$$P(H|F_n) = \sum_{i=0}^{n} P(H|C_i, F_n)P(C_i|F_n)$$

Then $P(H|C_i, F_n) = P(H|C_i) = \frac{i}{k}$ and by Bayes' rule

$$P(C_i|F_n) = \frac{P(F_n|C_i)P(C_i)}{\sum_{i=0}^{n} P(F_n|C_i)P(C_i)}$$

so since we are told that flipping our coin n times generates r heads and $n-r$ tails we have that

$$P(F_n|C_i) = \left(\begin{array}{c} n \\ r \end{array} \right) \left(\frac{i}{k}\right)^r \left(1 - \frac{i}{k}\right)^{n-r},$$

and that

$$P(C_i) = \frac{1}{k+1},$$

so that $P(C_i|F_n)$ becomes

$$\begin{aligned} P(C_i|F_n) &= \frac{\left(\begin{array}{c} n \\ r \end{array} \right) \left(\frac{i}{k}\right)^r \left(1 - \frac{i}{k}\right)^{n-r} \cdot \left(\frac{1}{k+1}\right)}{\sum_{i=0}^{n} \left(\begin{array}{c} n \\ r \end{array} \right) \left(\frac{i}{k}\right)^r \left(1 - \frac{i}{k}\right)^{n-r} \cdot \left(\frac{1}{k+1}\right)} \\ &= \frac{\left(\frac{i}{k}\right)^r \left(1 - \frac{i}{k}\right)^{n-r}}{\sum_{i=0}^{n} \left(\frac{i}{k}\right)^r \left(1 - \frac{i}{k}\right)^{n-r}}, \end{aligned}$$

so that our probability of a head becomes

$$P(H|F_n) = \frac{\sum_{i=0}^{n} \left(\frac{i}{k}\right)^{r+1} \left(1 - \frac{i}{k}\right)^{n-r}}{\sum_{i=0}^{n} \left(\frac{i}{k}\right)^r \left(1 - \frac{i}{k}\right)^{n-r}}$$

If k is large then we can write (the integral identity is proven below)

$$\frac{1}{k} \sum_{i=0}^{n} \left(\frac{i}{k}\right)^r \left(1 - \frac{i}{k}\right)^{n-r} \approx \int_0^1 x^r (1-x)^{n-r} dx = \frac{r!(n-r)!}{(n+1)!}.$$

Thus for large k our probability $P(H|F_n)$ becomes

$$P(H|F_n) = \frac{\frac{(r+1)!(n-r)!}{(n+2)!}}{\frac{r!(n-r)!}{(n+1)!}} = \frac{(r+1)!}{(n+2)!} \cdot \frac{(n+1)!}{r!} = \frac{r+1}{n+2}.$$

Where we have used the identity

$$\int_0^1 y^n (1-y)^m dy = \frac{n!\, m!}{(n+m+1)!}.$$

To prove this identity we will define $C(n,m)$ to be this integral and use integration by parts to derive a difference equation for $C(n,m)$. Remembering the integration by parts formula $\int udv = uv - \int vdu$ we see that

$$\begin{aligned} C(n,m) &\equiv \int_0^1 y^n(1-y)^m dy \\ &= (1-y)^m \frac{y^{n+1}}{n+1}\Big|_0^1 + \int_0^1 \frac{y^{n+1}}{n+1} m(1-y)^{m-1} dy \\ &= 0 + \frac{m}{n+1}\int_0^1 y^{n+1}(1-y)^{m-1} dy \\ &= \frac{m}{n+1} C(n+1, m+1)\,. \end{aligned}$$

Using this recurrence relationship we will prove the proposed representation for C by using mathematical induction. We begin by determining some initial conditions for $C(\cdot,\cdot)$. We have that

$$C(n,0) = \int_0^1 y^n dy = \frac{y^{n+1}}{n+1}\Big|_0^1 = \frac{1}{n+1}\,.$$

Note that this incidental equals $\frac{n!0!}{(n+1)!}$ as it should. Using our recurrence relation derived above we then find that

$$\begin{aligned} C(n,1) &= \frac{1}{n+1} C(n+1,0) = \frac{1}{(n+2)(n+1)} \quad \text{and} \\ C(n,2) &= \frac{2}{n+1} C(n+1,1) = \frac{2}{(n+1)(n+3)(n+2)}\,. \end{aligned}$$

Note that these two expressions equal $\frac{n!1!}{(n+2)!}$ and $\frac{n!2!}{(n+2+1)!}$ respectively as they should. We have shown that

$$C(n,m) = \frac{n!m!}{(n+m+1)!} \quad \text{for} \quad m \le 2\,,$$

so to prove this by induction we will assume that it holds in general and prove that it holds for $C(n,m+1)$. For the expression using our recurrence relationship (and our induction hypothesis) we have that

$$C(n,m+1) = \frac{m+1}{n+1} C(n+1,m) = \frac{m+1}{n+1}\left(\frac{(n+1)!m!}{(n+m+2)!}\right) = \frac{n!(m+1)!}{(n+(m+1)+1)!}\,,$$

which proves this result for $m+1$ and therefore by induction it is true for all m.

Chapter 3: Self-Test Problems and Exercises

Problem 9 (watering the plant)

Let W be the event the neighbor waters the plant and let D be the event the plant dies. We are told that $P(D|W^c) = 0.8$, $P(D|W) = 0.15$, and $P(W) = 0.90$.

	A	a
A	AA	Aa
a	aA	aa

Table 14: The possible genotypes from two hybrid parents.

Part (a): We want to compute $P(D^c)$. We have

$$\begin{aligned} P(D) &= P(D|W)P(W) + P(D|W^c)P(W^c) \\ &= P(D|W)P(W) + P(D|W^c)(1 - P(W)) \\ &= (0.15)(0.90) + (0.8)(1 - 0.90) = 0.135 + 0.080 = 0.215\,. \end{aligned}$$

Thus $P(D^c) = 1 - P(D) = 1 - 0.215 = 0.785$.

Part (b): We want to compute $P(W^c|D)$. We find

$$P(W^c|D) = \frac{P(D|W^c)P(W^c)}{P(D)} = \frac{P(D|W^c)(1 - P(W))}{P(D)} = \frac{(0.80)(1 - 0.90)}{0.215} = \frac{16}{43}\,.$$

Problem 10 (black rat genotype)

We are told that in a certain species of rats, black dominates over brown. Let A be the dominant black allele and let a be the recessive brown allele. Since the sibling rat is brown and brown is recessive, the sibling rat must have a genotype aa. The only way for two black parents to have a brown offspring is if both parents have the genotype Aa.

Part (a): As each parent contributes one allele to the genotype of the offspring, the possible offspring of our two Black parents are given by the results in Table 14. There we see that any offspring of these two parents will have genotypes of AA (with probability 1/4), Aa (with probability 1/2), or aa (with probability 1/4). Since we know that this rat is black it must be of genotypes AA or Aa. The probability it is of genotype Aa (or hybrid) is then

$$\frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{4}} = \frac{2}{3}\,.$$

The probability it is pure is then $1 - \frac{2}{3} = \frac{1}{3}$.

Part (b): This black rat then mates with a brown rat. Since brown is recessive, the brown rat must have a genotype of aa. As each parent contributes one allele to the genotype of the offspring, if the black rat has genotype AA, then offspring will have genotype Aa. If the black rat has genotype Aa, the offspring will have either the genotype Aa or aa. See Table 15 for the possible offspring.

Let E be the event the rat has genotype AA or is "pure", then E^c is the event the rat has genotype Aa or is "hybrid". Let C_i be the event the ith offspring has genotype Aa. Then from Table 15 we see that

$$P(C_i|E) = 1, \quad \text{and} \quad P(C_i|E^c) = \frac{1}{2}\,.$$

	a	a
A	Aa	Aa
A	Aa	Aa

	a	a
A	Aa	Aa
a	aa	aa

Table 15: Possible offspring of a black rat mated with a brown rat.

Then we want to evaluate $P(E \mid C_1C_2C_3C_4C_5)$. From Bayes' rule we have

$$P(E \mid C_1C_2C_3C_4C_5) = \frac{P(C_1C_2C_3C_4C_5 \mid E)P(E)}{P(C_1C_2C_3C_4C_5)}.$$

Assume the 5 events the offspring have genotype Aa are conditionally independent, given the rat has genotype AA or the rat has genotype Aa. We then have

$$P(C_1C_2C_3C_4C_5 \mid E) = \prod_{i=1}^{5} P(C_i|E) = (1)^5 \quad \text{and}$$
$$P(C_1C_2C_3C_4C_5 \mid E^c) = \prod_{i=1}^{5} P(C_i|E^c) = \left(\frac{1}{2}\right)^5.$$

Thus we have

$$\begin{aligned} P(C_1C_2C_3C_4C_5) &= P(C_1C_2C_3C_4C_5|E)P(E) + P(C_1C_2C_3C_4C_5|E^c)P(E^c) \\ &= \left(\frac{1}{3}\right)(1)^5 + \left(\frac{2}{3}\right)\left(\frac{1}{2}\right)^5, \end{aligned}$$

so our desired probability is given by

$$\begin{aligned} P(E \mid C_1C_2C_3C_4C_5) &= \frac{P(C_1C_2C_3C_4C_5 \mid E)P(E)}{P(C_1C_2C_3C_4C_5)} = \frac{\left(\frac{1}{3}\right)(1)^5}{\left(\frac{1}{3}\right)(1)^5 + \left(\frac{2}{3}\right)\left(\frac{1}{2}\right)^5} \\ &= \frac{1}{1+\frac{1}{16}} = \frac{16}{17}. \end{aligned}$$

Problem 11 (circuit flow)

Let C_i be the event the ith relay is closed (so that current can flow through that connection) and let E be the event current flows between A and B. If relay 1 is closed then the event current flows is

$$E|C_1 = C_4 \cup C_3C_5 \cup C_2C_5.$$

Note in the solution in the back of the book I think the expression there is missing the term C_2C_5. If on the other hand relay 1 is open the event current flows is

$$E|C_1^c = C_2C_3C_4 \cup C_2C_5.$$

Thus the probability that current flows is

$$P(E) = P(E|C_1)P(C_1) + P(E|C_1^c)P(C_1^c).$$

Since all relays are independent the condition that C_1 (or C_1^c) hold true does not affect the evaluation of the probabilities of the other sets of open/closed relays. Thus

$$\begin{aligned} P(C_4 \cup C_2C_5 \cup C_3C_5|C_1) &= P(C_4) + P(C_2C_5) + P(C_3C_5) \\ &- P(C_2C_3C_5) - P(C_2C_4C_5) - P(C_3C_4C_5) + P(C_2C_3C_4C_5) \quad \text{and} \\ P(C_2C_3C_4 \cup C_2C_5|C_1^c) &= P(C_2C_3C_4) + P(C_2C_5) - P(C_2C_3C_4C_5)\,. \end{aligned}$$

By independence each of term in the above can be expanded in terms of products of p_i. Using these two results we find

$$\begin{aligned} P(E) &= p_1(p_4 + p_2p_5 + p_3p_5 - p_2p_3p_5 - p_2p_4p_5 - p_3p_4p_5 + p_2p_3p_4p_5) \\ &+ (1 - p_1)(p_2p_3p_4 + p_2p_5 - p_2p_3p_4p_5)\,. \end{aligned}$$

Problem 12 (k-out-of-n-system)

Let C_i be the event the ith component is working and for each part let E be the event at least k of n components are working.

Part (a): To have one of two components working means that $E = C_1 \cup C_2$. Thus we have

$$\begin{aligned} P(E) = P(C_1 \cup C_2) &= P(C_1) + P(C_2) - P(C_1C_2) \\ &= P(C_1) + P(C_2) - P(C_1)P(C_2) \\ &= p_1 + p_2 - p_1p_2\,. \end{aligned}$$

The the probability of the event of interest or $P(C_1|E)$ can be computed as

$$P(C_1|E) = \frac{P(C_1E)}{P(E)} = \frac{P(C_1)}{P(E)} = \frac{p_1}{p_1 + p_2 - p_1p_2}\,,$$

since if C_1 is true then E must also be true.

Part (b): We will evaluate $P(C_1|E) = \frac{P(C_1E)}{P(E)}$. We thus need $P(C_1E)$ and $P(E)$. To have two of three components working means that

$$E = C_1C_2 \cup C_1C_3 \cup C_2C_3\,.$$

Thus

$$\begin{aligned} P(E) &= P(C_1C_2 \cup C_1C_3 \cup C_2C_3) \\ &= P(C_1C_2) + P(C_1C_3) + P(C_2C_3) \\ &- P(C_1C_2C_1C_3) + P(C_1C_2C_2C_3) + P(C_1C_3C_2C_3) + P(C_1C_2C_1C_3C_2C_3) \\ &= P(C_1C_2) + P(C_1C_3) + P(C_2C_3) + 2P(C_1C_2C_3) \\ &= p_1p_2 + p_1p_3 + p_2p_3 - 2p_1p_2p_3\,. \end{aligned}$$

From the above expression for E we have C_1E given by

$$C_1E = C_1C_2 \cup C_1C_3 \cup C_1C_2C_3 = C_1C_2 \cup C_1C_3\,,$$

since $C_1C_2C_3 \subset C_1C_2$. Thus

$$\begin{aligned} P(C_1E) &= P(C_1C_2 \cup C_1C_3) = P(C_1C_2) + P(C_1C_3) - P(C_1C_2C_3) \\ &= p_1p_2 + p_1p_3 - p_1p_2p_3 \,, \end{aligned}$$

and we find

$$P(C_1|E) = \frac{P(C_1E)}{P(E)} = \frac{p_1p_2 + p_1p_3 - p_1p_2p_3}{p_1p_2 + p_1p_3 + p_2p_3 - 2p_1p_2p_3} \,.$$

Problem 13 (roulette)

This is a classic example of what is called "gamblers fallacy". We can show that in fact the probability of getting red has not changed given the 10 times that black has appeared assuming independence of the sequential spins. Let R_i be the event the ball lands on red on the ith spin. Let B_i be the event the ball lands on black on the ith spin. Then the probability of the ball landing on red on spin i given the 10 other blacks is

$$\begin{aligned} P(R_i|B_{i-1}B_{i-2}\dots B_{i-10}) &= \frac{P(R_iB_{i-1}B_{i-2}\dots B_{i-10})}{P(B_{i-1}B_{i-2}\dots B_{i-10})} = \frac{P(R_i)P(B_{i-1})P(B_{i-2})\dots P(B_{i-10})}{P(B_{i-1})P(B_{i-2})\dots P(B_{i-10})} \\ &= P(R_i) \,, \end{aligned}$$

showing that there has been no change in his chance. It is interesting to note that even if one believed in this strategy very strongly (which we argue above is not a sound idea) the strategy itself would be onerous to implement since the event of 10 blacks in a row would not happen very frequently, giving rise to long waiting times between bets.

Problem 14 (the odd man out)

On each coin toss the player A will be the odd man if he gets heads while the others get tails or he gets tails while the others get heads. Each of these events happens with probability

$$p_1(1-p_2)(1-p_3) \quad \text{and} \quad (1-p_1)p_2p_3 \,, \tag{27}$$

respectively. The game continues if all players get heads or all players get tails. Each of these events happen with probability

$$p_1p_2p_3 \quad \text{and} \quad (1-p_1)(1-p_2)(1-p_3) \,, \tag{28}$$

respectively. The game stops with A *not* the odd man out with probability one minus the sum of the four probabilities above. Let E be the event A is the eventual odd man out and the probability we want to compute. Then we can compute $P(E)$ by conditioning on the result of the first set of coin tosses. Let E_1 be the event that A is the odd man out on one coin toss and O_1 the event that one of the players B or C is the odd man out on one coin toss. We then have

$$\begin{aligned} P(E) &= P(E|E_1)P(E_1) + P(E|O_1)P(O_1) + P(E|(E_1 \cup O_1)^c)P((E_1 \cup O_1)^c) \\ &= P(E_1) + P(E)P((E_1 \cup O_1)^c) \,. \end{aligned} \tag{29}$$

Where we have used the facts that

$$P(E|E_1) = 1\,, \quad P(E|O_1) = 0\,, \quad \text{and} \quad P(E|(E_1 \cup O_1)^c) = P(E)\,,$$

since in the last case no one was the odd man out and the game effectively starts over. From Equation 29 we can solve for $P(E)$. We find

$$P(E) = \frac{P(E_1)}{1 - P((E_1 \cup O_1)^c)}\,.$$

We can evaluate $P(E_1)$ by summing the two terms in Equation 27 and we can evaluate $1 - P((E_1 \cup O_1)^c)$ by recognizing that this is the probability no one is the odd man out on the first toss and then use Equation 28. Thus we get

$$P(E) = \frac{p_1(1-p_2)(1-p_3) + (1-p_1)p_2p_3}{1 - [p_1p_2p_3 + (1-p_1)(1-p_2)(1-p_3)]}\,.$$

Problem 15 (the second trial is larger)

Let N and M be the outcome of the first and second experiment respectively. We want $P\{M > N\}$. We can do this by conditioning on the outcome of M. We have

$$\begin{aligned} P\{M > N\} &= \sum_{i=1}^{n} P\{M > N = i | N = i\} P\{N = i\} \\ &= \sum_{i=1}^{N} \left(\sum_{j=i+1}^{n} p_j \right) p_i = \sum_{i=1}^{n} \sum_{j=i+1}^{n} p_i p_j\,. \end{aligned}$$

As another way to solve this problem let E be the event that the first experiment is smaller than the second, let F the event that the two experiments have the same value and let G be the event that the first experiment is larger than the second. Then by symmetry $P(E) = P(G)$ and we have

$$1 = P(E) + P(F) + P(G) = 2P(E) + P(F)\,.$$

Now we can explicitly evaluate $P(F)$ since $P(F) = \sum_{i=1}^{n} p_i^2$. Thus

$$P(E) = \left(\frac{1}{2}\right)(1 - \sum_{i=1}^{n} p_i^2)\,.$$

These two expressions can be shown to be equal by squaring the relationship $\sum_{i=1}^{n} p_i = 1$.

Problem 16 (more heads)

Let A be the event A gets more heads than B after each has flipped n coins. Let B be the event A gets fewer heads than B after each has flipped n coins. Let C be the event A and

B get the same number of heads after each has flipped n coins. Let E be the event A gets more total heads than B after his $n+1$st flip. The following the hint we have

$$\begin{aligned} P(E) &= P(E|A)P(A) + P(E|B)P(B) + P(E|C)P(C) \\ &= 1P(A) + 0p(B) + \frac{1}{2}P(C) \\ &= P(A) + \frac{1}{2}P(C)\,. \end{aligned}$$

But on the nth flip we have $P(A)+P(B)+P(C)=1$, and by symmetry $P(A)=P(B)$ thus

$$2P(A) + P(C) = 1 \quad \text{so} \quad P(C) = 1 - 2P(A)\,.$$

When we put this into the above expression for $P(E)$ we find

$$P(E) = P(A) + \frac{1}{2}(1 - 2P(A)) = \frac{1}{2}\,.$$

Problem 17 (independence with E, $F\cup G$, FG)

Part (a): This statement is False. Consider the following counter example. A die is rolled and let the events E, F, and G be defined as the following outcomes from this roll $E=\{1,6\}$, $F=\{1,2,3\}$, and $G=\{1,4,5\}$. These base events and their derivatives have the following probabilities

$$\begin{aligned} P(E) = \frac{2}{6}, \quad P(F) = \frac{3}{6}, \quad P(G) = \frac{3}{6} \\ P(EF) = \frac{1}{6}, \quad P(E|F) = \frac{1}{3} = P(E) \\ P(EG) = \frac{1}{6}, \quad P(E|G) = \frac{1}{3} = P(E)\,. \end{aligned}$$

Note that since $P(EF)=P(E)P(F)$ we have that E and F are independent. In the same way E and G are independent. Now consider the events $E(F\cup G)$ and $F\cup G$. We have

$$P(F\cup G) = \frac{5}{6}, \quad \text{and} \quad P(E(F\cup G)) = \frac{1}{6}\,.$$

Since $P(E(F\cup G)) = \frac{1}{6} \neq P(E)P(F\cup G) = \frac{1}{3}\cdot\frac{5}{6} = \frac{5}{18}$ we have that E and $F\cup G$ are not independent.

Part (b): This is true. Since E and F are independent, $P(EF)=P(E)P(F)$ and since E and G are independent we have $P(EG)=P(E)P(G)$. Now consider

$$\begin{aligned} P(E(F\cup G)) &= P(EF\cup EG) = P(EF) + P(EG) - P(EFEG) = P(EF) + P(EG) - P(EFG) \\ &= P(E)P(F) + P(E)P(G)\,, \end{aligned}$$

using independence and the fact that since $FG=\emptyset$ we have $EFG=\emptyset$. In the same way since $FG=\emptyset$ we have $P(F\cup G)=P(F)+P(G)$. Thus

$$P(E)P(F\cup G) = P(E)P(F) + P(E)P(G)\,.$$

Since this is the same expression as $P(E(F \cup G))$ we have that $P(E(F \cup G)) = P(E)P(F \cup G)$ or the pair E and $F \cup G$ are independent.

Part (c): This is true. Since E and FG are independent, we have $P(E(FG)) = P(E)P(FG)$. Since F and G are independent we have $P(FG) = P(F)P(G)$. Then

$$P(G(EF)) = P(E(FG)) = P(E)P(FG) = P(E)P(F)P(G) .$$

Since E and F are independent we have $P(EF) = P(E)P(F)$ and thus

$$P(G)P(EF) = P(G)P(E)P(F) .$$

Since these are both the same expressions we have shown that, $P(G(EF)) = P(G)P(EF)$, and the pair G and EF are independent.

Problem 18 ($\emptyset$ and independence)

Part (a): This is always false. The reason is that if $AB = \emptyset$, then $P(AB) = P(\emptyset) = 0$, but we are told that $P(A) > 0$ and $P(B) > 0$ thus the product the the two probabilities is nonnegative $P(A)P(B) > 0$. Thus it is not possible for $P(AB) = P(A)P(B)$ and A and B are not independent.

Part (b): This is always false. The reason is that if we assume A and B are independent, then $P(AB) = P(A)P(B)$. Since we assume that $P(A) > 0$ and $P(B) > 0$ we must have$P(A)P(B) > 0$ and thus $P(AB) \neq 0$, which would be required if A and B were mutually exclusive.

Part (c): This is always false. If we assume $P(A) = P(B) = 0.6$ and assume that A and B could be mutually exclusive we would then conclude

$$P(A \cup B) = P(A) + P(B) - P(AB) = 0.6 + 0.6 - 0 = 1.2 .$$

But we can't have the probability of the event $A \cup B$ greater than 1. Thus A and B cannot be mutually exclusive.

Part (d): This can possibly be true. Let an urn have 6 red balls and 4 white balls and draw sequentially twice with replacement a ball. Let A and B be the event that we draw a red ball. Then $P(A) = P(B) = 0.6$ and these two events are independent (since we are drawing with replacement).

Problem 19 (ranking trials)

Let H be the event the coin toss is heads. Let E_i be the event the result of the ith trial is success. The probabilities of each event in the order they are listed are

- $P(H) = \dfrac{1}{2} = 0.5$
- $P(E_1E_2E_3) = P(E_1)P(E_2)P(E_3) = (0.8)^3 = 0.512$, when $P(E_i) = 0.8$.
- $P(\cap_{i=1}^7 E_i) = \prod_{i=1}^7 P(E_i) = (0.9)^7 = 0.4789269$, when $P(E_i) = 0.9$.

Problem 20 (defective radios)

To start this problem we define several events, let A be the event the that the radios were produced at factory A, B be the event they were produced at factory B and let D_i be the event the ith radio is defective. Then from the problem statement we are told that $P(A) = P(B) = \frac{1}{2}$, $P(D_i|A) = 0.05$, and $P(D_i|B) = 0.01$. We observe that event D_2 and want to calculate $P(D_1|D_2)$. To compute this probability we will condition on whether or not the two radios came from factory A or B. We have

$$\begin{aligned} P(D_2|D_1) &= P(D_2|D_1, A)P(A|D_1) + P(D_2|D_1, B)P(B|D_1) \\ &= P(D_2|A)P(A|D_1) + P(D_2|B)P(B|D_1)\,, \end{aligned}$$

where we have assumed D_1 and D_2 are conditionally independent given A or B. Now to evaluate $P(A|D_1)$ and $P(B|D_1)$ we will use Bayes' rule. For example for either D_1 or D_2 we have

$$P(A|D) = \frac{P(D|A)P(A)}{P(D)} = \frac{P(D|A)P(A)}{P(D|A)P(A) + P(D|B)P(B)}\,.$$

Using the numbers given for this problem we have

$$\begin{aligned} P(A|D) &= \frac{0.05(0.5)}{0.05(0.5) + 0.01(0.5)} = 0.833 \\ P(B|D) &= \frac{0.01(0.5)}{0.05(0.5) + 0.01(0.5)} = 0.166\,. \end{aligned}$$

Thus we find

$$P(D_2|D_1) = 0.05(0.833) + 0.01(0.166) = 0.0433\,.$$

Problem 21 ($P(A|B) = 1$ means $P(B^c|A^c) = 1$)

We are told that $P(A|B) = 1$, thus using the definition of conditional probability we have $\frac{P(A,B)}{P(B)} = 1$ or $P(A, B) = P(B)$. The event AB is always a subset of the event B thus $P(AB) < P(B)$ if $AB \neq B$. Thus $AB = B$ and $B \subseteq A$. Complementing this relationship gives $B^c \supseteq A^c$ or that $P(B^C|A^c) = 1$ since if A^c occurs then B^c must have also occurred as $A^c \subseteq B^c$.

Problem 22 (i red balls)

For this problem let $E(i, n)$ be the event there are exactly i red balls in the urn after n stages and let R_n be the event a red ball is selected at stage n to generate the configuration of balls in the next stage, where we take $n \geq 0$. Then we want to show

$$P(E(i, n)) = \frac{1}{n+1} \quad \text{for} \quad 1 \leq i \leq n+1\,. \tag{30}$$

At stage $n = 0$ the urn initially contains 1 red and 1 blue ball

$$P(E(1, 0)) = 1 = \frac{1}{0+1}\,.$$

Thus we have the needed initial condition for the induction hypothesis. Now let $n > 0$ and assume that Equation 30 holds up to some stage n and we want to show that it then also holds on stage $n + 1$.

To evaluate $P(E(i, n + 1))$ we consider how we could get i red balls on the $n + 1$st stage. There are two ways this could happen, either we had i red balls during stage n and we drew a blue ball, or we had $i - 1$ red balls during stage n and we drew a red ball. Since initially there are 2 balls in the urn and one ball is added to the urn at each stage. Thus after stage n, there are $n + 2$ balls in the urn. Thus on stage n the probability we draw a red or blue ball is given by

$$\begin{aligned} P(R_n) &= \frac{i-1}{n+2} \\ P(R_{n-1}^c) &= 1 - P(R_{n-1}) = 1 - \frac{i-1}{n+2} = \frac{n+2-i}{n+2}\,. \end{aligned}$$

Thus we have

$$\begin{aligned} P(E(i, n+1)) &= P(E(i-1, n))P(R_n) + P(E(i, n))P({R_n}^c) \\ &= P(E(i-1, n))\left(\frac{i-1}{n+2}\right) + P(E(i, n))\left(\frac{n+2-i}{n+2}\right) \\ &= \frac{1}{n+1}\left(\frac{i-1}{n+2}\right) + \frac{1}{n+1}\left(\frac{n+2-i}{n+2}\right) \\ &= \frac{n+1}{(n+1)(n+2)} = \frac{1}{n+2}\,, \end{aligned}$$

where we have used the induction hypothesis to conclude $P(E(i-1, n)) = P(E(i, n)) = \frac{1}{n+1}$. Since we have shown that Equation 30 is true at stage $n+1$ by induction it is true for all n.

Problem 25 (a conditional inequality)

Now following the hint we have

$$P(E|E \cup F) = P(E|E \cup F, F)P(F) + P(E|E \cup F, \neg F)P(\neg F)\,.$$

But $P(E|E\cup F, F) = P(E|F)$, since $E\cup F \supset F$, and $P(E|E\cup F, \neg F) = P(E|E\cap\neg F) = 1$, so the above becomes

$$P(E|E\cup F) = P(E|F)P(F) + (1 - P(F))\,.$$

Dividing by $P(E|F)$ we have

$$\frac{P(E|E\cup F)}{P(E|F)} = P(F) + \frac{1-P(F)}{P(E|F)}\,.$$

Since $P(E|F) \leq 1$ we have that $\frac{1-P(F)}{P(E|F)} \geq 1 - P(F)$ and the above then becomes

$$\frac{P(E|E\cup F)}{P(E|F)} \geq P(F) + (1 - P(F)) = 1$$

giving the desired result of $P(E|E\cup F) \geq P(E|F)$. In words this says that the probability that E occurs given E *or* F occurs must be larger than if we just know that only F occurs.

Balls	(W,W)	(W,B)	(W,0)	(B,B)	(B,0)	(0,0)
X	-2	+1	-1	4	2	0

Table 16: Possible values for our winnings X when two colored balls are selected from the urn in Problem 1.

Chapter 4 (Random Variables)

Chapter 4: Problems

Problem 1 (winning by drawing balls from an urn)

The possibilities of the various X's we can obtain are given in Table 16. We find that the probabilities of the various X values are given by

$$\begin{aligned}
P\{X=-2\} &= \frac{\binom{8}{2}}{\binom{14}{2}} = \frac{4}{13} \\
P\{X=-1\} &= \frac{\binom{8}{1}\binom{2}{1}}{\binom{14}{2}} = \frac{16}{91} \\
P\{X=0\} &= \frac{\binom{2}{2}}{\binom{14}{2}} = \frac{1}{91} \\
P\{X=1\} &= \frac{\binom{8}{1}\binom{4}{1}}{\binom{14}{2}} = \frac{32}{91} \\
P\{X=2\} &= \frac{\binom{4}{1}\binom{2}{1}}{\binom{14}{2}} = \frac{8}{91} \\
P\{X=3\} &= 0 \\
P\{X=4\} &= \frac{\binom{4}{2}}{\binom{14}{2}} = \frac{6}{91}\,.
\end{aligned}$$

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

Table 17: The possible values for the product of two dice when two dice are rolled.

Problem 2 (the product of two dice)

We begin by constructing the sample space of possible outcomes. These numbers are computed in table 17, where the row corresponds to the first die and the column corresponds to the second die. In each square we have placed the product of the two dice. Each pair has probability of 1/36, so by enumeration we find that

$$
\begin{aligned}
P\{X=1\} &= \frac{1}{36}, \quad P\{X=2\} = \frac{2}{36} \\
P\{X=3\} &= \frac{2}{36}, \quad P\{X=4\} = \frac{3}{36} \\
P\{X=5\} &= \frac{2}{36}, \quad P\{X=6\} = \frac{4}{36} \\
P\{X=8\} &= \frac{2}{36}, \quad P\{X=9\} = \frac{1}{36} \\
P\{X=10\} &= \frac{2}{36}, \quad P\{X=12\} = \frac{4}{36} \\
P\{X=15\} &= \frac{2}{36}, \quad P\{X=16\} = \frac{1}{36} \\
P\{X=18\} &= \frac{2}{36}, \quad P\{X=20\} = \frac{2}{36} \\
P\{X=24\} &= \frac{2}{36}, \quad P\{X=25\} = \frac{1}{36} \\
P\{X=30\} &= \frac{2}{36}, \quad P\{X=36\} = \frac{1}{36},
\end{aligned}
$$

with any other integer having zero probability.

Problem 4 (ranking five men and five women)

Note: In contrast to the explicitly stated instructions provided by the problem where $X = 1$ would correspond to the event that the highest ranked woman is ranked first (best), I choose to solve this problem with the backwards convention that the best ranking corresponds to $X = 10$, effectively the reverse of the standard convention. I'm sorry if this causes any confusion.

As the variable X represents the ranking of the the highest female when we have five total females and using the backwards ranking convention discussed above, the lowest the highest ranking can be is five, so $P\{X = i\} = 0$ if $1 \leq i \leq 4$. Now $P\{X = 5\}$ is proportional to the number of ways to get five women first in a line

$$P\{X = 5\} = \frac{(5!)(5!)}{10!} = \frac{1}{7 \cdot 6^2} = \frac{1}{252}\,.$$

Since if the first five positions are taken up by women and the last five positions are taken up by men we have 5! orderings of the women and 5! orderings of the men. Giving $5! \cdot 5!$ possible arrangements.

We now want to evaluate $P\{X = 6\}$. To do this we must place a women in the sixth place which can be done in five possible ways (from among the five women). We then must place four more women and one man in the positions $1, 2, 3, 4$ and 5. We can pick the man in five ways (from among all possible men) and his position in another five ways. We then have 4! orderings of the remaining four women and 4! orderings of the remaining four men. Thus our probability is then

$$P\{X = 6\} = \frac{5 \cdot 5 \cdot 5 \cdot 4! \cdot 4!}{10!} = \frac{5}{252}\,.$$

Now we need to evaluate $P\{X = 7\}$ which has a numerator consisting of the following product of terms.

$$\binom{5}{1} \cdot \binom{5}{2} \cdot (2!) \cdot \binom{6}{2} \cdot (4!) \cdot (3!)\,.$$

The first term $\binom{5}{1}$, is the ways to pick the women in the seventh spot. The term $\binom{5}{2}$ is the number of ways to pick the two men that will go to the left of this woman. The term 2! represents all possible permutations of these two men. The term $\binom{6}{2}$, is the number of ways we can select the specific spots these two men go into. The term 4! is the ways to pick the orderings of the remaining women. Finally the 3!, represents the number of ways to pick the ordering of the three remaining men. We then need to divide the product by 10! to convert this into a probability which gives

$$P\{X = 7\} = \frac{5}{84}\,.$$

We now need to evaluate $P\{X = 8\}$. To do this we reason as follows. We have $\binom{5}{1}$ ways to pick the women to place in the eighth spot. Then $\binom{7}{4}$ spots to the left of this women to pick as the spots where the four remaining women will be placed. Then 4! different placements of the women in these spots. Once all the women are placed we have $7 - 4 = 3$ slots to place three men who will be to the left of the initial women at position eight. The men to go in these spots can be picked in $\binom{5}{3}$ ways and their ordering selected in 3!.

Finally, we have 2! arrangements of the remaining two men, giving a total count of the number of instances where $X = 8$ of

$$\binom{5}{1} \cdot \binom{7}{4} \cdot (4!) \cdot \binom{5}{3} \cdot (3!) \cdot (2!) = 504000\,,$$

which gives a probability

$$P\{X = 8\} = \frac{5}{36}\,.$$

To evaluate $P\{X = 9\}$ we have $\binom{5}{1}$ ways to pick the women at position nine. Then $\binom{8}{4}$ ways to pick the remaining slots to the left of this women to place the remaining other women into, and 4! ways to rearrange them. We then have $8 - 4 = 4$ slots for men to go into and $\binom{5}{4}$ ways to pick four men to fill these spots and 4! ways to rearrange them. So the number of instances when $X = 9$ is given by

$$\binom{5}{1} \cdot \binom{8}{4} \cdot (4!) \cdot \binom{5}{4} \cdot (4!) = 1008000\,,$$

which gives a probability of

$$P\{X = 9\} = \frac{5}{18}\,.$$

Finally, to evaluate $P\{X = 10\}$ we have $\binom{5}{1}$ ways to pick the women, $\binom{9}{4}$ ways to pick spots for the four remaining women and 4! ways to rearrange them. With the women placed, we have five slots remaining for the men and 5! ways of arraignment them. This gives

$$\binom{5}{1} \cdot \binom{9}{4} \cdot (4!) \cdot (5!)\,,$$

Giving a probability of

$$P\{X = 10\} = \frac{1}{2}\,.$$

One can further check that if we add all of these probabilities up we obtain

$$\frac{1}{252} + \frac{5}{252} + \frac{5}{84} + \frac{5}{36} + \frac{5}{18} + \frac{1}{2} = 1\,,$$

as we should.

We now present a simpler method that uses combinatorial counting to evaluate these probabilities. As before we have $P\{X = i\} = 0$ for $1 \le i \le 4$. Lets now assume $5 \le i \le 10$ and we can compute $P\{X = i\}$ in the following way. We first select one of the five women to occupy the leading position i. This can be done in 5 ways. Next we have a total of $i - 1$ positions behind the leading woman occupying position i in which we can to place the four remaining women. We can select the spots in to place the women in $\binom{i-1}{4}$ ways and their specific ordering in 4! ways. Next we can place the men in the remaining five spots in 5! ways. Using the multiplication principle we conclude that the probability for $X = i$ is thus given by

$$P\{X = i\} = \frac{5\left(\begin{array}{c} i-1 \\ 4 \end{array}\right)(4!)(5!)}{10!}.$$

Since the product of $\left(\begin{array}{c} i-1 \\ 4 \end{array}\right)$ and 4! simplifies as

$$\left(\begin{array}{c} i-1 \\ 4 \end{array}\right)(4!) = \frac{(i-1)!}{(i-5)!},$$

the expression for $P\{X = i\}$ above simplifies to

$$P\{X = i\} = \frac{5}{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}\left(\frac{(i-1)!}{(i-5)!}\right) \quad \text{for} \quad 5 \leq i \leq 10 .$$

Evaluating the above for each value of i duplicated the results from above. One thing to notice about the above formula is that the location of the men in this problem statement becomes irrelevant. This can be seen if we write the final expression above as $\frac{5\left(\frac{(i-1)!}{(i-5)!}\right)}{10\cdot9\cdot8\cdot7\cdot6}$. In that expression the denominator of $10 \cdot 9 \cdot 8 \cdot 7 \cdot 6$ represents the number of ways one can place five women in ten spots, while the numerator in the above expression

$$5\left(\frac{(i-1)!}{(i-5)!}\right) = 5!\frac{(i-1)!}{(i-1-4)!4!} = 5!\left(\begin{array}{c} i-1 \\ 4 \end{array}\right),$$

for $5 \leq i \leq 10$ represents the number of ways to place five women where the top woman is in the i-th spot. See the Matlab/Octave file `chap_4_prob_4.m` for the fractional simplifications needed in this problem.

Problem 5 (the difference between heads and tails)

Define $X = n_H - n_T$ with n_H the number of heads and n_T the number of tails. Then if our sequence of n flips results in all heads (n of them) with no tails we have $X = n$. If we have $n - 1$ heads (and thus one tail) the variable X is given by $X = n - 1 - 1 = n - 2$. Continuing, if we have $n - 2$ heads and therefore two tails our variable X then becomes, $X = n - 2 - 2 = n - 4$. In general, we see by induction that

$$X \in \{n, n-2, n-4, \cdots, 4-n, 2-n, -n\},$$

or as a formula $X = n - 2i$ with i taken from $0, 1, 2, \cdots, n$. This result can be easily be derived algebraically by recognizing the constrain that $n = n_H + n_T$, which implies when we solve for n_H that $n_H = n - n_T$, so that

$$\begin{aligned} X &\equiv n_H - n_T = (n - n_T) - n_T \\ &= n - 2n_T , \end{aligned}$$

where $0 \leq n_T \leq n$.

Problem 6 (the probabilities of heads minus tails)

From Problem 5 we see that the probability that X takes on a specific value is directly related to the probability of obtaining some number n_T of tails. The probability of obtaining n_T tails in n flips is a binomial random variable with parameters $(n, p = 1/2)$ and thus probability $\left(\begin{array}{c} n \\ n_T \end{array}\right) p^{n_T}(1-p)^{n-n_T}$. Thus for a fair coin (where $p = 1/2$) we have

$$\begin{aligned}
P\{X = n\} &= P\{n_T = 0\} = \frac{\left(\begin{array}{c} n \\ 0 \end{array}\right)}{2^n} = \frac{1}{2^n} \\
P\{X = n-2\} &= P\{n_T = 1\} = \frac{\left(\begin{array}{c} n \\ 1 \end{array}\right)}{2^n} = \frac{n}{2^n} \\
P\{X = n-4\} &= P\{n_T = 2\} = \frac{\left(\begin{array}{c} n \\ 2 \end{array}\right)}{2^n} = \frac{n(n-1)}{2^{n+1}},
\end{aligned}$$

etc. So in general we have

$$P\{X = n - 2i\} = P\{n_T = i\} = \frac{1}{2^n}\left(\begin{array}{c} n \\ i \end{array}\right).$$

So if $n = 3$ we have

$$\begin{aligned}
P\{X = 3\} &= \frac{1}{2^3} = \frac{1}{8} \\
P\{X = 1\} &= \frac{1}{2^3}\left(\begin{array}{c} 3 \\ 1 \end{array}\right) = \frac{3}{8} \\
P\{X = -1\} &= \frac{1}{2^3}\left(\begin{array}{c} 3 \\ 2 \end{array}\right) = \frac{3}{8} \\
P\{X = -3\} &= \frac{1}{2^3}\left(\begin{array}{c} 3 \\ 3 \end{array}\right) = \frac{1}{8}
\end{aligned}$$

Problem 7 (the functions of two dice)

In table 18 we construct a table of all possible outcomes associated with the two dice rolls. In that table the row corresponds to the first die and the column corresponds to the second die. Then for each part of the problem we find that

Part (a): $X \in \{1, 2, 3, 4, 5, 6\}$.

Part (b): $X \in \{1, 2, 3, 4, 5, 6\}$.

Part (c): $X \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

Part (d): $X \in \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$.

	1	2	3	4	5	6
1	(1,1,2,0)	(2,1,3,-1)	(3,1,4,-2)	(4,1,5,-3)	(5,1,6,-4)	(6,1,7,-5)
2	(2,1,3,1)	(2,2,4,0)	(3,2,5,-1)	(4,2,6,-2)	(5,2,7,-3)	(6,2,8,-4)
3	(3,1,4,2)	(3,2,5,1)	(3,3,6,0)	(4,3,7,-1)	(5,3,8,-2)	(6,3,9,-3)
4	(4,1,5,3)	(4,2,6,2)	(4,3,7,1)	(4,4,8,0)	(5,4,9,-1)	(6,4,10,-2)
5	(5,1,6,4)	(5,2,7,3)	(5,3,8,2)	(5,4,9,1)	(5,5,10,0)	(6,5,11,-1)
6	(6,1,7,5)	(6,2,8,4)	(6,3,9,3)	(6,4,10,2)	(6,5,11,1)	(6,6,12,0)

Table 18: The possible values for the maximum, minimum, sum, and first minus second die observed when two dice are rolled.

Problem 8 (probabilities on dice)

The solution to this problem involves counting up the number of times that X equals the given value and then dividing by $6^2 = 36$. For each part we have the following

Part (a): From table 18, for this part we find that

$$\begin{aligned} P\{X=1\} &= \frac{1}{36}, \quad P\{X=2\} = \frac{1}{12}, \quad P\{X=3\} = \frac{5}{36}, \\ P\{X=4\} &= \frac{7}{36}, \quad P\{X=5\} = \frac{1}{4}, \quad P\{X=6\} = \frac{11}{36} \end{aligned}$$

Part (b): From table 18, for this part we find that

$$\begin{aligned} P\{X=1\} &= \frac{11}{36}, \quad P\{X=2\} = \frac{1}{4}, \quad P\{X=3\} = \frac{7}{36}, \\ P\{X=4\} &= \frac{5}{36}, \quad P\{X=5\} = \frac{7}{36}, \quad P\{X=6\} = \frac{1}{36} \end{aligned}$$

Part (c): From table 18, for this part we find that

$$\begin{aligned} P\{X=2\} &= \frac{1}{36}, \quad P\{X=3\} = \frac{1}{18}, \quad P\{X=4\} = \frac{1}{12}, \\ P\{X=5\} &= \frac{1}{9}, \quad P\{X=6\} = \frac{5}{36}, \quad P\{X=7\} = \frac{1}{6}, \\ P\{X=8\} &= \frac{5}{36}, \quad P\{X=9\} = \frac{1}{9}, \quad P\{X=10\} = \frac{1}{12}, \\ P\{X=11\} &= \frac{1}{18} \quad P\{X=12\} = \frac{1}{36}. \end{aligned}$$

Part (d): From table 18, for this part we find that

$$\begin{aligned} P\{X=-5\} &= \frac{1}{36}, \quad P\{X=-4\} = \frac{1}{18}, \quad P\{X=-3\} = \frac{1}{9}, \\ P\{X=-2\} &= \frac{1}{9}, \quad P\{X=-1\} = \frac{5}{36}, \quad P\{X=0\} = \frac{1}{6}, \\ P\{X=1\} &= \frac{5}{36}, \quad P\{X=2\} = \frac{1}{9}, \quad P\{X=3\} = \frac{1}{12}, \\ P\{X=4\} &= \frac{1}{18}, \quad P\{X=5\} = \frac{1}{36}. \end{aligned}$$

Problem 9 (sampling balls from an urn with replacement)

For this problem balls selected with replacement from an urn and we define the random variable X as $X = \max(x_1, x_2, x_3)$ where x_1, x_2, and x_3 are the numbers on the balls from each three draws. We know that

$$P\{X = 1\} = \frac{1}{20^3}.$$

Now $P\{X = 2\}$ can be computed as follows. To count the number of sets of three draws that contain at least one two and so the max will be 2 (one such set is $(1, 1, 2)$) we consider all sets of three we could build from the components 1 and 2. Then for each slot we have two choices so we have $2 \cdot 2 \cdot 2 = 2^3$ possible choices. But one of these (the one selected by assembling an ordered set of three elements from only the element one) so the number of sets with a two as the largest element is $2^3 - 1^3 = 2^3 - 1 = 7$. To compute

$$P\{X = 3\}$$

we consider all ordered sets we can construct from the elements $1, 2, 3$ since we have three choices for the first spot, three for the second spot, and three for the third we have $3^3 = 27$. The number of sets that have a three in them are this number minus the number of sets that have only two's and one's in them. Which is given by 2^3 thus we have $3^3 - 2^3 = 27 - 8 = 19$. The general pattern then is

$$P\{X = i\} = \frac{i^3 - (i-1)^3}{20^3}. \tag{31}$$

As a more general way to derive the result above consider that with replacement the sample space for three draws is $\{1, 2, \cdots, 20\}^3$. Then each possible draw of three numbers has the same probability $\frac{1}{20^3}$. If we let H_i denote the event that the highest numbered ball drawn (from the three) has a value of i then the event H_i can be broken down into several mutually independent events. The first is the draw (i, i, i) where all three balls show the number i. This can happen in only one way. The second type of draw under which event H_i can be said to have occurred are draws where there are two balls that have the highest number i and the third draw has a lower number. An example draw like this would be (i, i, X), with $X < i$. This can happen in $\left(\begin{array}{c} 3 \\ 1 \end{array}\right) = 3$ ways and each draw has a probability of

$$\left(\frac{1}{20}\right)\left(\frac{1}{20}\right)\left(\frac{i-1}{20}\right),$$

of happening. The third type of draw we can get and have event H_i is one of the type (i, X, Y) where the two numbers X and Y are such that $X < i$ and $Y < i$. This draw has a probability of happening given by

$$\left(\frac{1}{20}\right)\left(\frac{i-1}{20}\right)\left(\frac{i-1}{20}\right),$$

and there are $\left(\begin{array}{c} 3 \\ 1 \end{array}\right) = 3$ ways that draws like this can happen. Thus to compute $P\{X = i\}$ we sum the three results above to get

$$P\{X = i\} = \frac{1}{20^3} + \frac{3(i-1)}{20^3} + \frac{3(i-1)^2}{20^3} = \frac{3i^2 - 3i + 1}{20^3}.$$

By expanding the numerator of Equation 31 we can show that these two expressions are equivalent. An *much* simpler method to get Equation 31 can be obtained by using the "trichotomy" property by noting that

$$P\{X \le i\} = P\{X = i\} + P\{X < i\} = P\{X = i\} + P\{X \le i - 1\} ,$$

or

$$P\{X = i\} = P\{X \le i\} - P\{X \le i - 1\} = \left(\frac{i}{20}\right)^3 - \left(\frac{i-1}{20}\right)^3 .$$

Using the above to calculate $P\{X \ge 17\}$ or the probability that we win the bet we find

$$\frac{17^3 - 16^3 + 18^3 - 17^3 + 19^3 - 18^3 + 20^3 - 19^3}{20^3} = \frac{20^3 - 16^3}{20^3} = \frac{61}{125} .$$

A *much* simpler method of solving this problem is simply to note that

$$P\{\text{winning}\} = P\{X \ge 17\} = 1 - P\{X < 17\} = 1 - (16/20)^3 ,$$

the same result as earlier.

Problem 10 (if we win i dollars)

For this problem we desire to compute the conditional probability we win i dollars given we win something. Let E be the event that we win something and we want to evaluate $P\{X = i|E\}$ using Bayes' rule we find that

$$P\{X = i|E\} = \frac{P\{E|X = i\}P\{X = i\}}{P\{E\}} .$$

Now $P\{E\} = \sum_{i=1,2,3} P\{E|X = i\}P\{X = i\}$ for these are the i that we make a profit on and therefore have $P\{E|X = i\} \ne 0$ (the other i's all have $P\{E|X = i\} = 0$). For these i's we have that

$$P\{E|X = i\} = 1 ,$$

so we get for $P\{E\}$ given by

$$P\{E\} = \frac{39}{165} + \frac{15}{165} + \frac{1}{165} = \frac{1}{3} .$$

So that we have $P\{X = i|E\} = 0$, when $i = 0, -1, -2, -3$, and that

$$\begin{aligned} P\{X = 1|E\} &= \frac{P\{X = 1\}}{P\{E\}} = \frac{\frac{39}{165}}{\frac{1}{3}} = 0.709 \\ P\{X = 2|E\} &= \frac{P\{X = 2\}}{P\{E\}} = \frac{\frac{15}{165}}{\frac{1}{3}} = 0.2727 \\ P\{X = 3|E\} &= \frac{P\{X = 3\}}{P\{E\}} = \frac{\frac{1}{165}}{\frac{1}{3}} = 0.01818 . \end{aligned}$$

Problem 11 (the Riemann hypothesis)

Part (a): Note that there are $\lfloor \frac{10^3}{3} \rfloor = 333$ multiples of three in the set $\{1, 2, 3, ., 10^3\}$. The multiples are specifically

$$3 \cdot 1, 3 \cdot 2, \cdots, 3 \cdot 333 .$$

Note that the last element equals 999. Since we are given that any number N is equally likely to to be chosen from the 1000 numbers we see that

$$P(N \text{ is a multiple of } 3) = \frac{333}{10^3} .$$

In the same way we would compute

$$\begin{aligned}
P(N \text{ is a multiple of } 5) &= \frac{\lfloor \frac{10^3}{5} \rfloor}{10^3} = \frac{200}{10^3} \\
P(N \text{ is a multiple of } 7) &= \frac{\lfloor \frac{10^3}{7} \rfloor}{10^3} = \frac{142}{10^3} \\
P(N \text{ is a multiple of } 15) &= \frac{\lfloor \frac{10^3}{15} \rfloor}{10^3} = \frac{66}{10^3} \\
P(N \text{ is a multiple of } 105) &= \frac{\lfloor \frac{10^3}{105} \rfloor}{10^3} = \frac{9}{10^3} .
\end{aligned}$$

Note that the expression used above "N is a multiple of the number r" is the same statement as "N is divisible by r". In each of the above cases we see that as k gets larger and larger we expect

$$P(N \text{ is divisible by } r) = \lim_{k \to \infty} \frac{1}{10^k} \left\lfloor \frac{10^k}{r} \right\rfloor = \frac{1}{r} . \tag{32}$$

Part (b): From Equation 32 we have that N is *not* divisible by r with probability

$$P(N \text{ is not divisible by } r) = 1 - \frac{1}{r} = \frac{r-1}{r} .$$

From the background given in this problem we want to evaluate $P\{\mu(N) = 0\}$ or equivalently

$$P\{N \text{ is not divisible by any } P^2 \text{ with } P \text{ a prime}\} .$$

Lets enumerate the primes from smallest to largest $P_1, P_2, \cdots$ where $P_1 = 2$, $P_2 = 3$, $P_3 = 5$ etc. For two distinct primes P_i and P_j let E_i and E_j be the events that N is divisible by P_i^2 and P_j^2 respectively. The event E_i^c is the event that N is not disable by P_i^2, $P(E_i) = \frac{1}{P_i^2}$, and $P(E_i^c) = 1 - \frac{1}{P_i^2} = \frac{P_i^2 - 1}{P_i^2}$ with equivalent expressions for E_j^c and $P(E_j)$. We then want to evaluate the event that N is not divisible by any prime squared or in terms of the events defined

$$P\{\mu(N) = 0\} = P\{E_1^c E_2^c E_3^c \cdots\} .$$

Lets now consider the independent events E_i and E_j. We want to show that E_i^c and E_j^c are independent so that the above product event can be evaluated by taking the product of the individual events. Consider the divisibility of two distinct primes (squared) P_i^2 and P_j^2.

Then the total sample space of probability events is the union of the disjoint sets $E_i \cup E_j$ and $(E_i \cup E_j)^c$. Thus

$$1 = P(E_i \cup E_j) + P((E_i \cup E_j)^c) = P(E_i \cup E_j) + P(E_i^c \cap E_j^c) \, .$$

Thus solving for $P(E_i^c \cap E_j^c)$ and using what we know for $P(E_i)$ we have

$$\begin{aligned} P(E_i^c \cap E_j^c) &= P(E_i^c E_j^c) = 1 - P(E_i \cup E_j) = 1 - (P(E_i) + P(E_j) - P(E_i E_j)) \\ &= 1 - \frac{1}{P_i^2} - \frac{1}{P_j^2} + \frac{1}{P_i^2 P_j^2} = \left(1 - \frac{1}{P_i^2}\right)\left(1 - \frac{1}{P_j^2}\right) \\ &= P(E_i^c) P(E_j^c) \, . \end{aligned}$$

This last result shows that E_i^c and E_j^c are independent. Generalizing this to all of the terms in the above expression for $P\{\mu(N) = 0\}$ we find that

$$\begin{aligned} P\{\mu(N) = 0\} &= P\{E_1^c E_2^c E_3^c \cdots\} = P\{E_1^c\} P\{E_2^c\} P\{E_3^c\} \cdots \\ &= \left(1 - \frac{1}{2^2}\right)\left(1 - \frac{1}{3^2}\right)\left(1 - \frac{1}{5^2}\right) \cdots \\ &= \prod_{i=1}^{\infty} \left(1 - \frac{1}{P_i^2}\right) = \prod_{i=1}^{\infty} \frac{1 - P_i^2}{P_i^2} = \frac{6}{\pi^2} \approx 0.6079271 \, . \end{aligned}$$

The expression we wanted to show.

Player		Opponent		Random Variable
Guesses	Shows	Guesses	Shows	Amount Won: X
1	1	1	1	0
1	1	1	2	−3
1	1	2	1	2
1	1	2	2	0
1	2	1	1	3
1	2	1	2	0
1	2	2	1	0
1	2	2	2	−4
2	1	1	1	−2
2	1	1	2	0
2	1	2	1	0
2	1	2	2	3
2	2	1	1	0
2	2	1	2	4
2	2	2	1	−3
2	2	2	2	0

Table 19: Two-Finger Morra Outcomes

Problem 12 (two-finger Morra)

Part (a): All possible outcomes for a round of play of two-finger Morra are shown in Table 19. Under the given assumptions, each row of the table is equally likely and can therefore be assigned a probability of $\frac{1}{16}$. Using that table the associated probabilities for the possible values of X are given by

$$\begin{aligned}
P\{X=0\} &= \frac{8}{16} = \frac{1}{2} \\
P\{X=2\} &= P\{X=-2\} = \frac{1}{16} \\
P\{X=3\} &= P\{X=-3\} = \frac{2}{16} = \frac{1}{8} \\
P\{X=4\} &= P\{X=-4\} = \frac{1}{16}\,.
\end{aligned}$$

Part (b): In this case the strategy corresponds to using only rows 1, 4, 13, and 16 of Table 19. We see that either both players guess correctly or both players guess incorrectly on every play. Thus the only output is $X = 0$ and we have $P\{X = 0\} = 1$.

Problem 13 (selling encyclopedias)

There are 9 possible outcomes, as summarized in Table 20. Summing all possible ways to get the various values of X we find

$$\begin{aligned}
P\{X=0\} &= .28 \\
P\{X=500\} &= .21 + .06 = .27 \\
P\{X=1000\} &= .21 + .045 + .06 = .315 \\
P\{X=1500\} &= .045 + .045 = .09 \\
P\{X=2000\} &= .045\,.
\end{aligned}$$

Sale from Customer 1	Sales from Customer 2	X	Probability
0	0	0	$(1-.3)(1-.6) = .28$
0	500	500	$(1-.3)(.6)(.5) = .21$
0	1000	1000	$(1-.3)(.6)(.5) = .21$
500	0	500	$(.3)(.5)(1-.6) = .06$
500	500	1000	$(.3)(.5)(.6)(.5) = .045$
500	1000	1500	$(.3)(.5)(.6)(.5) = .045$
1000	0	1000	$(.3)(.5)(1-.6) = .06$
1000	500	1500	$(.3)(.5)(.6)(.5) = .045$
1000	1000	2000	$(.3)(.5)(.6)(.5) = .045$

Table 20: Encyclopedia Sales

Problem 14 (getting the highest number)

To begin we note that there are 5! equally likely possible possible orderings of the numbers $1-5$ that could be dealt to the five players. Now player 1 will win 4 times if he has the highest of the five numbers. Thus the first number must be a 5 followed by any of the 4! possible orderings of the other numbers. This gives a probability

$$P\{X=4\} = \frac{1\cdot 4!}{5!} = \frac{1}{5}.$$

Next player 1 will win 3 times if his number exceeds the numbers of players 2, 3, and 4, but is less than the number of player 5. In other words, player 1 must have the second highest number and player 5 the highest. This means that player 5 must have been given the number 5 and player 1 must have been given the number 4 and the other 3 numbers can be in any order among the remaining players. This gives a probability of

$$P\{X=3\} = \frac{1\cdot 1\cdot 3!}{5!} = \frac{1}{20}.$$

For player 1 to win twice he must have a number greater than the numbers of players 2 and 3 but less than that of player 4; i.e., of the first four players, player four has the highest number and player 1 has the second highest. We select the four numbers to assign to the first four player in $\binom{5}{4}$ ways. This leaves a single number for player 5. We then select the largest number from this group of four, for player four (in one way), and then select the second largest number (in one way) for the first player. This gives two remaining numbers which can be ordered in two ways. This gives a probability of

$$P\{X=2\} = \frac{\binom{5}{4}\cdot 1\cdot 1\cdot 2!\cdot 1}{5!} = \frac{1}{12}.$$

Player 1 wins exactly once if his number is higher than that of player 2 and lower than that of player 3. Following the logic when we win twice we select the three numbers for players $1-3$ in $\binom{5}{3}$ ways. This gives two numbers for the players 4 and 5 which can be ordered in 2 ways. From the initial set of three, we assign the largest of this set to player 3 and the next largest to player 1. The last number goes to player 2. Taken together this gives a probability of

$$P\{X=1\} = \frac{\binom{5}{3}\cdot 1\cdot 1\cdot 1\cdot 2!}{5!} = \frac{1}{6}.$$

Finally, player 1 never wins if his number is less than that of player 2. The same logic as above gives for this probability the following

$$P\{X=0\} = \frac{\binom{5}{2}\cdot 1\cdot 1\cdot 3!}{5!} = \frac{1}{2}$$

Problem 15 (the NBA draft pick)

Notice first that once a ball belonging to a team has been drawn, any other balls belonging to that team are subsequently ignored, so we may treat the problem as if all balls belonging to a team are removed from the urn once any ball belonging to that team is drawn. Let us adopt the following terminology in analyzing the problem:

$$
\begin{aligned}
F_i &= \text{“First pick goes to team with } i\text{-th worst record”} \\
S_i &= \text{“Second pick goes to team with } i\text{-th worst record”} \\
T_i &= \text{“Third pick goes to team with } i\text{-th worst record”}
\end{aligned}
$$

In the above notation we have $1 \le i \le 11$. Note that $i = 1$ is the worst team and has 11 balls in the urn initially, $i = 2$ is the second worst team and has 10 balls in the urn initially, etc. In general, the ith worst team has $12 - i$ balls in the urn until it is selected. With this shorthand, much of this problem can be solved by conditioning on what happens "first" i.e. that the event F_i comes before the event S_i and both come before T_i.

$$
\begin{aligned}
P\{X=1\} &= P\{F_1\} = \frac{11}{66} = 0.1667\,. \\
P\{X=2\} &= P\{S_1\} \\
&= P\{F_2 S_1 \vee F_3 S_1 \vee \cdots \vee F_{11} S_1\} \\
&= P\{F_2\} P\{S_1 \mid F_2\} + P\{F_3\} P\{S_1 \mid F_3\} + \cdots + P\{F_{11}\} P\{S_1 \mid F_{11}\} \\
&= \frac{10}{66} \cdot \frac{11}{66-10} + \frac{9}{66} \cdot \frac{11}{66-9} + \cdots + \frac{1}{66} \cdot \frac{11}{66-1} \\
&= \sum_{k=2}^{11} \frac{12-k}{66} \cdot \frac{11}{66-(12-k)} \\
&= \sum_{k=2}^{11} \frac{12-k}{66} \cdot \frac{11}{54+k} = 0.15563\,. \\
P\{X=3\} &= P\{T_1\} \\
&= \sum_{k=2}^{11} P\{S_k T_1\} = \sum_{k=2}^{11} \left(\sum_{\substack{j=2 \\ j \ne k}}^{11} P\{F_j S_k T_1\} \right) \\
&= \sum_{k=2}^{11} \left(\sum_{\substack{j=2 \\ j \ne k}}^{11} \frac{12-j}{66} \cdot \frac{12-k}{66-(12-j)} \cdot \frac{11}{66-(12-j)-(12-k)} \right) \\
&= \sum_{k=2}^{11} \sum_{\substack{j=2 \\ j \ne k}}^{11} \frac{12-j}{66} \cdot \frac{12-k}{54+j} \cdot \frac{11}{42+j+k} = 0.1435 \\
P\{X=4\} &= 1 - \sum_{i=1}^{3} P\{X=i\} = 0.53423 \\
P\{X=i\} &= 0, \quad i \notin \{1,2,3,4\}
\end{aligned}
$$

Note that $P\{X = i\} = 0$ for $i \ge 5$ since if it is not drawn in the first 3 draws it will be *given* the fourth draft pick according to the rules. These sums are computed in the MATLAB file `chap_4_prob_15.m`.

Problem 16 (more draft picks)

Following the notation introduced in the previous problem, we have

Part (a):

$$P\{Y_1 = i\} = P\{F_i\} = \frac{12-i}{66} \quad \text{for} \quad 1 \le i \le 11\,.$$

Part (b):

$$\begin{aligned} P\{Y_2 = i\} = P\{S_i\} &= \sum_{j=1; j\neq i}^{11} P\{F_j S_i\} \\ &= \sum_{j=1; j\neq i}^{11} \frac{12-j}{66}\left(\frac{12-i}{66-(12-j)}\right) = \sum_{j=1; j\neq i}^{11} \frac{12-j}{66}\left(\frac{12-i}{54+j}\right)\,. \end{aligned}$$

Part (c):

$$\begin{aligned} P\{Y_3 = i\} &= P\{T_i\} \\ &= \sum_{j=1; j\neq i}^{11} P\{S_j T_i\} = \sum_{j=1; j\neq i}^{11}\left(\sum_{k=1; k\neq i,j}^{11} P\{F_k S_j T_i\}\right) \\ &= \sum_{j=1; j\neq i}^{11}\left(\sum_{k=1; k\neq i,j}^{11} \frac{12-k}{66}\cdot\frac{12-j}{66-(12-k)}\cdot\frac{12-i}{66-(12-j)-(12-k)}\right)\,. \end{aligned}$$

Problem 17 (probabilities from the distribution function)

Part (a): We find that

$$\begin{aligned} P\{X=1\} &= P\{X \le 1\} - P\{X < 1\} = \frac{1}{2} - \frac{1}{4} = \frac{1}{4} \\ P\{X=2\} &= P\{X \le 2\} - P\{X < 2\} = \frac{11}{12} - \left(\frac{1}{2} + \frac{2-1}{4}\right) = \frac{1}{6} \\ P\{X=3\} &= P\{X \le 3\} - P\{X < 3\} = 1 - \frac{11}{12} = \frac{1}{12}\,. \end{aligned}$$

Part (b): We find that

$$\begin{aligned} P\{\frac{1}{2} < X < \frac{3}{2}\} &= \lim_{n\to\infty} P\{X \le \frac{3}{2} - \frac{1}{n}\} - P\{X \le \frac{1}{2}\} \\ &= \left(\frac{1}{2} + \frac{\frac{3}{2}-1}{4}\right) - \frac{\frac{1}{2}}{4} = \frac{1}{2}\,. \end{aligned}$$

Problem 18 (the probabilities mass from the number of heads)

The event H_k, that in n flips of a coin we get k heads is given by a binomial random variable and so we have

$$P\{H_k\} = \binom{n}{k} p^k (1-p)^{n-k}.$$

When $n = 4$ and $p = \frac{1}{2}$ as for this problem we have $P\{H_k\} = \binom{4}{k}\left(\frac{1}{2}\right)^4$. Thus

$$\begin{aligned}
P\{X = -2\} &= P\{H_k = 0\} = \left(\frac{1}{2}\right)^4 \\
P\{X = -1\} &= P\{H_k = 1\} = 4\left(\frac{1}{2}\right)^4 \\
P\{X = 0\} &= P\{H_k = 2\} = 6\left(\frac{1}{2}\right)^4 \\
P\{X = 1\} &= P\{H_k = 3\} = 4\left(\frac{1}{2}\right)^4 \\
P\{X = 2\} &= P\{H_k = 4\} = \left(\frac{1}{2}\right)^4,
\end{aligned}$$

are the values of the probability mass function.

Problem 19 (probabilities from the distribution function)

Since $F(b) = \sum_{x \le b} p(x)$ from the given expression for $F(b)$ we see that

$$\begin{aligned}
p(0) &= \frac{1}{2} \\
p(1) &= \frac{3}{5} - \frac{1}{2} = \frac{1}{10} \\
p(2) &= \frac{4}{5} - \frac{3}{5} = \frac{1}{5} \\
p(3) &= \frac{9}{10} - \frac{4}{5} = \frac{1}{10} \\
p(3.5) &= 1 - \frac{9}{10} = \frac{1}{10}.
\end{aligned}$$

Problem 20 (a winning roulette strategy)

This problem can more easily be worked by imagining a tree like structure representing the possible outcomes and their probabilities. For example, from the problem in the text, if we

First Game	Second Game	Third Game	X	Probability
win	N.A.	N.A.	+1	p
loss	win	win	+1	p^2q
loss	win	loss	-1	pq^2
loss	loss	win	-1	pq^2
loss	loss	loss	-3	q^3

Table 21: Playing "no loose" roulette

win on the first play (with probability $p = \frac{9}{19}$) we stop playing and have won +1. If we loose on the first play we will play the game two more times. In these two games we can win twice, win once and lose once or loose twice. Ignoring for the time begin the initial loss these three outcomes occur with probabilities given by a binomial distribution with $n = 3$ and $p = \frac{9}{19}$ or

$$p^2\,,\quad 2pq\,,\quad q^2\,.$$

The reward (payoff) for each of the outcomes is given by

$$+2\,,\quad 0\,,\quad -2\,.$$

Since these second two games are only played if we loose the first game we must condition them on the output from that event. Thus the total probabilities (and win amounts X) then are given in Table 21 Using these we can answer the questions given.

Part (a): We find

$$P\{X > 0\} = p + p^2q = \frac{9}{19} + \frac{10}{19}\left(\frac{9}{19}\right)^2 = 0.5917\,.$$

Part (b): No, there are two paths where we win but 3 where we loose. One of these paths has a loss of -3 which is relatively large given the problem.

Part (c): We find

$$E[X] = 1p + 1p^2q - 1pq^2 - 1pq^2 - 3q^3 = -0.108\,,$$

when we evaluate.

Problem 21 (selecting students or buses)

Part (a): The probability of selecting a student on a bus is proportional to the number of students on that bus while the probability we select a given bus driver is simply $1/4$, since there is no weighting based on the number of students in each bus. Thus $E[X]$ should be larger than $E[Y]$.

Part (b): We have that

$$E[X] = \sum_{i=1}^{4} x_i p(x_i) = 40\left(\frac{40}{148}\right) + 33\left(\frac{33}{148}\right) + 25\left(\frac{25}{148}\right) + 50\left(\frac{50}{148}\right) = 39.28\,.$$

while

$$E[Y] = \sum_{i=1}^{4} y_i p(y_i) = 40\left(\frac{1}{4}\right) + 33\left(\frac{1}{4}\right) + 50\left(\frac{1}{4}\right) = 37\,.$$

So we see that $E[X] > E[Y]$ as expected.

Problem 22 (winning games)

Warning: Due to time constraints this problem has not been checked as thoroughly as others and may not be entirely complete. If anyone finds anything wrong with these please let me know.

We will consider the two specific cases where $i = 2$ and $i = 3$ before the general case. When $i = 2$ to evaluate the expected number of games played we want to evaluate $P\{N = n\}$ where N is the random variable determining the expected number of games played before a win (by either team A or B). Then $P\{N = 1\} = 0$ since we need two wins for A or B to win overall. Now $P\{N = 2\} = p^2 + q^2$, since from the four possible outcomes from the two experiments (A, A), (A, B), (B, A), and (B, B) only two result in a win. The first (A, A) occurs with probability p^2 and the last with probability q^2. Since they are mutually exclusive events we have the desired probability above. Continuing, we have that

$$P\{N = 3\} = 2pqp + 2qpq = 2p^2q + 2q^2p\,,$$

since to have A win on three games (and not win on two) we must place the last of A's wins as the third win. Thus only two sequences give wins for A in three flips i.e. (A, B, A) and (B, A, A). Each occurs with probability p^2q. The second term in the above is equivalent but with q replaced with p.

The expression $P\{N = 4\}$ is not a valid probability since one player A or B would have won before four games. We can also check that we have a complete formulation by computing the probability A or B wins after *any* number of flips i.e. consider

$$\begin{aligned}
p^2 + q^2 + 2p^2q + 2q^2p &= p^2 + (1-p)^2 + 2p^2(1-p) + 2(1-p)^2p \\
&= p^2 + 1 - 2p + p^2 + 2p^2 - 2p^3 + 2(1 - 2p + p^2)p = 1\,.
\end{aligned}$$

Thus the expected number of games to play before one team wins is

$$E[N] = 2(p^2 + q^2) + 3(2p^2q + 2q^2p) = 2 + 2p - 2p^2$$

In the general case it appears that

$$
\begin{aligned}
P\{N=i\} &= p^i + q^i \\
P\{N=i+1\} &= \binom{i}{1} qp^i + \binom{i}{1} pq^i \\
P\{N=i+2\} &= \binom{i+1}{2} q^2p^i + \binom{i+1}{2} p^2q^i \\
P\{N=i+3\} &= \binom{i+2}{3} q^3p^i + \binom{i+2}{3} p^3q^i \\
&\vdots \\
P\{N=i+(i-1)\} &= \binom{2i-2}{i-1} q^{i-1}p^i + \binom{2i-2}{i-1} p^{i-1}q^i \,.
\end{aligned}
$$

In the case $i=3$ the above procedure becomes

$$
\begin{aligned}
P\{N=3\} &= p^3+q^3 \\
P\{N=4\} &= 3qp^3+3pq^3 \\
P\{N=5\} &= 6q^2p^3+6p^2q^3 \,.
\end{aligned}
$$

Checking that we have included every term we compute the sum of all of the above terms to obtain

$$
p^3+q^3+3qp^3+3pq^3+6q^2p^3+6p^2q^3 \,,
$$

Which is simplified (to the required one) in the Mathematica file `chap_4_prob_22.nb`. To compute the expectation we have

$$
E[N] = 3(p^3+q^3) + 4(3qp^3+3pq^3) + 5(6q^2p^3+6p^2q^3) \,.
$$

We take the derivative of this expression and set it equal to zero in the Mathematical file `chap_4_prob_22.nb`.

Problem 23 (trading commodities)

Part (a): Let assume one buys x of the commodity at the start of the week, then in cash one has $C = 1000 - 2x$. Here we have x ounces of our commodity with $0 \le x \le 500$. Then at the end of the week our total value is given by

$$
V = 1000 - 2x + Yx \,,
$$

where Y is the random variable representing the cost per ounce of the commodity. We desire to maximize $E[V]$. We have

$$
\begin{aligned}
E[V] &= 1000 - 2x + x\sum_{i=1}^{2} y_i p(y_i) \\
&= 1000 - 2x + x\left(1\left(\frac{1}{2}\right) + 4\left(\frac{1}{2}\right)\right) \\
&= 1000 + \frac{x}{2} \,.
\end{aligned}
$$

Since this is an increasing linear function of x, to maximize our expected amount of money, we should buy as much as possible. Thus let $x = 500$ i.e. buy all that one can.

Part (b): We desire to maximize the expected amount of the commodity that one posses. Now by purchasing x at the beginning of the week, one is then left with $1000 - 2x$ cash to buy more at the end of the week. The amount of the commodity A, that we have at the end of the week is given by

$$A = x + \frac{1000 - 2x}{Y},$$

where Y is the random variable denoting the cost per ounce of our commodity at the end of the week. Then the expected value of A is then given by

$$\begin{aligned} E[A] &= x + \sum_{i=1}^{2} \left(\frac{1000 - 2x}{y_i} \right) p(y_i) \\ &= x + \left(\frac{1000 - 2x}{1} \right) \left(\frac{1}{2} \right) + \left(\frac{1000 - 2x}{4} \right) \left(\frac{1}{2} \right) \\ &= 625 - \frac{x}{4}. \end{aligned}$$

Which is linear and decreases with increasing x. Thus we should pick $x = 0$ i.e. buy none of the commodities now and buy it all at the end of the week.

Problem 24

Part (a): Let X_B be the gain of B when playing the game. Then if A has written down one we have

$$E[X_B] = p(1) + (1 - p) \left(\frac{-3}{4} \right) = \frac{7p - 3}{4}.$$

However if A has written down two, then our expectation becomes

$$E[X_B] = p \left(\frac{-3}{4} \right) + (1 - p)2 = \frac{8 - 11p}{4}.$$

To derive the value of p that will maximize player B's return, we incorporate the fact that the profit X_B depends on what A does by conditioning on the possible choices. Thus we have

$$E[X_B] = \begin{cases} \frac{7p-3}{4} & A \text{ picks } 1 \\ \frac{8-11p}{4} & A \text{ picks } 2 \end{cases}$$

Plotting these two linear lines we have Figure **??** (left). From this graph we recognize that we will guarantee the maximal possible expected return independent of what A does if we select p such that

$$\frac{7p - 3}{4} = \frac{8 - 11p}{4}.$$

which gives $p = \frac{11}{18}$. Thus the expected gain with this value of p is given by

$$\left.\frac{7p-3}{4}\right|_{p=\frac{11}{18}} = \frac{23}{72}\,.$$

Now consider the expected loss of player A under his randomized rule. To do so, let Y_A be the random variable specifying the loss received by player A. Then if B always picks number one we have

$$E[Y_A] = q(-1) + (1-q)\left(\frac{3}{4}\right) = \frac{3}{4} - \frac{7}{4}q\,,$$

while if B always picks number two we have

$$E[Y_A] = q\left(\frac{3}{4}\right) + (1-q)(-2) = \frac{11}{4}q - 2\,.$$

Plotting these expected losses as function of q we have Figure **??** (right). Then to find the smallest expected loss for player A independent of what player B does we have to find q such that

$$\frac{3}{4} - \frac{7}{4}q = \frac{11}{4}q - 2\,.$$

When we solve for q we find that $q = \frac{11}{18}$, which is the same as before. Now the optimal expected loss is given by

$$\frac{3}{4} - \frac{7}{4}\left(\frac{11}{18}\right) = -\frac{23}{72}\,,$$

which is the negative of the expected gain for player B.

Problem 25 (expected winnings with slots)

To compute the expected winnings when playing one game from a slot machine, we first need to compute the probabilities for each of the winning combinations. To begin with we note that we have a total of 20^3 possible three dial combinations. Now lets compute a count of each dial combination that results in a payoff. We find

$$\begin{aligned}
N(\text{Bar}, \text{Bar}, \text{Bar}) &= 3 \\
N(\text{Bell}, \text{Bell}, \text{Bell}) &= 2 \cdot 2 \cdot 3 = 12 \\
N(\text{Bell}, \text{Bell}, \text{Bar}) &= 2 \cdot 2 \cdot 1 = 4 \\
N(\text{Plum}, \text{Plum}, \text{Plum}) &= 4 \cdot 1 \cdot 6 = 24 \\
N(\text{Orange}, \text{Orange}, \text{Orange}) &= 3 \cdot 7 \cdot 6 = 126 \\
N(\text{Orange}, \text{Orange}, \text{Bar}) &= 3 \cdot 7 \cdot 1 = 21 \\
N(\text{Cherry}, \text{Cherry}, \text{Anything}) &= 7 \cdot 7 \cdot 20 = 980 \\
N(\text{Cherry}, \text{No Cherry}, \text{Anything}) &= 7 \cdot (20-7) \cdot 20 = 1820\,.
\end{aligned}$$

So the number of non winning rolls is given by $20^3 - \sum \text{Above} = 20^3 - 2990 = 5101$. Thus the expected winnings are then given by

$$\begin{aligned} E[W] &= 60\left(\frac{3}{20^3}\right) + 20\left(\frac{12}{20^3}\right) + 18\left(\frac{4}{20^3}\right) + 14\left(\frac{24}{20^3}\right) \\ &+ 10\left(\frac{126}{20^3}\right) + 8\left(\frac{21}{20^3}\right) + 2\left(\frac{980}{20^3}\right) \\ &+ 0\left(\frac{1820}{20^3}\right) - 1\left(\frac{5010}{20^3}\right) = -0.09925\,. \end{aligned}$$

Problem 26 (guess my number)

Part (a): If at stage n by asking the question "is it i", one is able to eliminate one possible choice from further consideration (assuming that we have not guessed the correct number) before stage n. Thus let E_n be the event at stage n we guess the number correctly, assuming we have not guessed it correctly in the $n-1$ earlier stages. Then

$$P(E_n) = \frac{1}{10-(n-1)} = \frac{1}{11-n}\,.$$

so we have that

$$\begin{aligned} P(E_1) &= \frac{1}{10} \\ P(E_2) &= \frac{1}{9} \\ P(E_3) &= \frac{1}{8} \\ &\vdots \\ P(E_{10}) &= 1\,. \end{aligned}$$

The expected number of guesses to make using this method is then given by

$$\begin{aligned} E[N] &= 1\left(\frac{1}{10}\right) + 2\left(1-\frac{1}{10}\right)\left(\frac{1}{9}\right) \\ &+ 3\left(1-\frac{1}{10}\right)\left(1-\frac{1}{9}\right)\frac{1}{8} + \cdots \\ &= 1\left(\frac{1}{10}\right) + 2\left(\frac{1}{10}\right) + 3\left(\frac{1}{10}\right) + \cdots \\ &= \sum_{n=1}^{10} n\left(\frac{1}{10}\right) = \frac{1}{10}\left(\frac{10(10+1)}{2}\right) = 5.5\,. \end{aligned}$$

Part (b): In this second case we will ask questions of the form: "is i less than the current midpoint of the list". For example, initially the number can be any of the numbers

	Questions (in order)	Number of Questions
1	$(<5);(<3);(<2)$	3
2	$(<5);(<3);(<2)$	3
3	$(<5);(<3);(<2)$	3
4	$(<5);(<3);$	2
5	$(<5);(<7);(<6)$	3
6	$(<5);(<7);(<6)$	3
7	$(<5);(<7);(<6)$	3
8	$(<5);(<7);(<8)$	3
9	$(<5);(<7);(<8);(<9)$	4
10	$(<5);(<7);(<8);(<9)$	4

Table 22: The sequence of questions asked and the number for the situations where the hidden number is somewhere between 1 and 10.

$1, 2, 3, \cdots, 9, 10$ so one could ask the question "is i less than five". If the answer is yes, then we repeat our search procedure on the list $1, 2, 3, 4$. If the answer is no, we repeat our search on the list $5, 6, 7, 8, 9, 10$. Thus we never know the identity of the hidden number until $O(\text{ceiling}(\lg(10)))$ steps have been taken. Since $\lg(10) = 3.32$ we require $O(4)$ steps. To determine the expected number of steps, lets enumerate the number of guesses each specific integer would require using the above method Note, that it might be better to ask the question is i less than or equal to x). Then since any given number is equally likely to be selected the expected number of question to be asked is given by

$$E[N] = \frac{1}{10}(7 \cdot 3 + 2 + 8) = 3.1\,.$$

Problem 27

The company desires to make $0.1\,A$ of a profit. Assuming the cost charged to each customer is C, the expected profit of the company then given by

$$C + p(-A) + (1-p)(0) = C - pA\,.$$

This can be seen as the fixed cost received from the paying customers minus what is lost if a claim must be paid out. For this to be $0.1A$ we should have $c - pA = 0.1A$ or solving for C we have

$$C = \left(p + \frac{1}{10}\right) A\,.$$

Problem 28

We can explicitly calculate the number of defective items obtained in the sample of twenty. We find that

$$\begin{aligned} P_0 &= \frac{\binom{16}{3}\binom{4}{0}}{\binom{20}{3}} = 0.491 \\ P_1 &= \frac{\binom{16}{2}\binom{4}{1}}{\binom{20}{3}} = 0.421 \\ P_2 &= \frac{\binom{16}{1}\binom{4}{2}}{\binom{20}{3}} = 0.084 \\ P_3 &= \frac{\binom{16}{0}\binom{4}{3}}{\binom{20}{3}} = 0.0035\,, \end{aligned}$$

so the expected number of defective items is given by

$$3P_3 + 2P_2 + 1P_1 + 0P_0 = 0.6 = \frac{3}{5}\,.$$

Problem 29 (a machine that breaks down)

Under the first strategy we would check the first possibility and if needed check the second possibility. This has an expected cost of

$$C_1 + R_1\,,$$

if the first possibility is true (which happens with probability p) and

$$C_1 + C_2 + R_2\,,$$

if the second possibility is true (which happens with probability $1-p$). Here I am explicitly assuming that if the first check is a failure we must then check the second possibility (at a cost C_2) before repair (at a cost of R_2). Another assumption would be that if the first check is a failure then we know that the second cause is the real one and we don't have to check for it. This results in a cost of $C_1 + R_2$ rather than $C_1 + C_2 + R_2$. The first assumption seems more consistent with the problem formulation and will be the one used. Thus under the first strategy we have an expected cost of

$$p(C_1 + R_1) + (1-p)(C_1 + C_2 + R_2)\,,$$

so our expected cost becomes

$$C_1 + pR_1 + (1-p)(C_2 + R_2) = C_1 + C_2 + R_2 + p(R_1 - C_2 - R_2)\,.$$

Now under the second strategy we would first check the second possibility and if needed check the first possibility. This first action has an expected cost of

$$C_2 + R_2\,,$$

if the second possibility is true cause (this happens with probability $1-p$) and

$$C_2 + C_1 + R_1\,,$$

if the first possibility is true (which happens with probability p). This gives an expected cost when using the second strategy of

$$(1-p)(C_2 + R_2) + p(C_2 + C_1 + R_1) = C_2 + R_2 + p(C_1 + R_1 - R_2)\,.$$

The expected cost under strategy number one will be less than the expected cost under strategy number if

$$C_1 + C_2 + R_2 + p(R_1 - C_2 - R_2) < C_2 + R_2 + p(C_1 + R_1 - R_2)\,.$$

When we solve for p the above simplifies to

$$p > \frac{C_1}{C_1 + C_2}\,.$$

As the threshold value to use for the different strategies. This result has the intuitive understanding in that if p is "significantly" large (meaning the break is more likely to be caused by the first possibility) we should check the first possibility first. While if p is not significantly large we should check the second possibility first.

Problem 30 (the St. Petersburg paradox)

The probability that the first tail appears on the nth flip means that the $n-1$ heads must first appear and then a tail. This gives a probability of

$$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{n-1} = \left(\frac{1}{2}\right)^{n}\,.$$

Then the expected value of our winnings is given by

$$\sum_{n=1}^{\infty} 2^n \left(\frac{1}{2}\right)^n = \sum_{n=1}^{\infty} 1 = +\infty\,.$$

Part (a): If a person payed 10^6 to play this game he would only "win" if the first tail appeared on toss greater than or equal to n^* where $n^* \geq \log_2(10^6) = 6\log_2(10) = 6\frac{\ln(10)}{\ln(2)} = 19.931$, or $n^* = 20$. In that case this event would occur with probability

$$\sum_{k=n^*}^{\infty}\left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^{n^*}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^{n^*-1}\,,$$

since $\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = 2$. With $n^* = 20$ we see that this probability is given by 9.5367×10^{-7} a rather small number. Thus many would not be willing to play under these conditions.

Part (b): In this case, if we play k games then we will definitely "win" if the first tail appears on a flip n^* (or greater) where n^* solves

$$-k\,10^6 + 2^{n^*} > 0\,,$$

or

$$n^* > 6\log_2(10) + \log_2(k) = 19.931 + \log_2(k)\,.$$

Since this target n^* grows logarithmically with k one would expect that enough random experiments were ran that eventually a very high paying result would appear. Thus many would be willing to pay this game.

Problem 31 (scoring your guess)

Since the meteorologist truly believes that it will rain with probability p^* if he quotes a probability p, then the expected score he will receive is given by

$$E[S;p] = p^*(1-(1-p)^2) + (1-p^*)(1-p^2)\,.$$

We want to pick a value of p such that we maximize this expression. To do so, consider the derivative of this expression set equal to zero and solve for the value of p. We find that

$$\frac{dE[S;p]}{dp} = p^*(2(1-p)) + (1-p^*)(-2p) = 0\,.$$

solving for p we find that $p = p^*$. Taking the second derivative of this expression we find that

$$\frac{d^2E[S;p]}{dp^2} = -2p^* - 2(1-p^*) = -2 < 0\,,$$

showing that $p = p^*$ is a maximum. This is a nice reason for using this metric, since it behooves the meteorologist to quote the probability of rain that he truly believes is true.

Problem 32 (testing diseased people)

We have one hundred people which we break up into ten groups of ten for the purposes of testing for a disease. For each group we will test the entire group of people with one test. This test will be "positive" (meaning at least one person has the disease) with probability $1-(0.9)^{10}$. Since 0.9^{10} is the probability that all people are normal and the complement of this probability is the probability that at least one person has the disease. Then the expected number of tests for each group of ten is then

$$1 + 0((0.9)^{10}) + 10(1-(0.9)^{10}) = 11 - 10(0.9)^{10} = 7.51\,.$$

Where the first 1 is because we will certainly test the pooled people and the remaining to expressions represent the case where the entire pooled test result comes back negative (no more tests needed) and the case where the entire pooled test result comes back positive (meaning we have ten individual tests to then do).

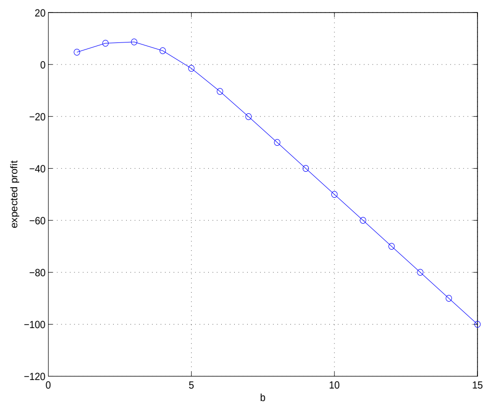

Figure 1: The expected profit for the newspaper ordering problem when b papers are ordered.

Problem 33 (the number of papers to purchase)

Let b be the variable denoting the number of papers bought and N the *random* variable denoting the number of papers demanded. Finally, let the random variable P be the newsboys' profits. Then with these definitions the newsboys' profits is given by

$$P = -10b + 15\min(N, b) \quad \text{for} \quad b \geq 1\,,$$

This is because if we only buy b papers we can only sell a maximum of b papers independent of what the demand N is. Then to calculate the expected profit we have that

$$\begin{aligned} E[P] &= -10b + 15E[\min(N, b)] \\ &= -10b + 15\sum_{n=0}^{10} \min(n, b) \binom{10}{n} \left(\frac{1}{3}\right)^n \left(\frac{2}{3}\right)^{10-n} . \end{aligned}$$

To evaluate the optimal number of papers to buy we can plot this as a function of b for $1 \leq b \leq 15$. In the Matlab file `chap_4_prob_33.m`, where this function is computed and plotted. See Figure 1, for a figure of the produced plot. There one can see that the maximum expected profit occurs when we order $b = 3$ newspapers. The expected profit in that case is given by 8.36.

Problem 35 (a game with marbles)

Part (a): Define W to be the random variable expression the winnings obtained when one plays the proposed game. The expected value of W is then given by

$$E[W] = 1.1P_{\text{sc}} - 1.0P_{\text{dc}}$$

where the notation "sc" means that the two drawn marbles are of the same color and the notation "dc" means that the two drawn marbles are of different colors. Now to calculate each of these probabilities we introduce the four possible events that can happen when we draw to marbles: RR, BB, RB, and BR. As an example the notation RB denotes the event that we first draw a red marble and then second draw a black marble. With this notation we see that P_{sc} is given by

$$\begin{aligned} P_{\text{sc}} &= P\{RR\} + P\{BB\} \\ &= \frac{5}{10}\left(\frac{4}{9}\right) + \frac{5}{10}\left(\frac{4}{9}\right) = \frac{4}{9}\,. \end{aligned}$$

while P_{dc} is given by

$$\begin{aligned} P_{\text{dc}} &= P\{RB\} + P\{BR\} \\ &= \frac{5}{10}\left(\frac{5}{9}\right) + \frac{5}{10}\left(\frac{5}{9}\right) = \frac{5}{9}\,. \end{aligned}$$

With these two results the expected profit is then given by

$$1.1\left(\frac{4}{9}\right) - 1.0\left(\frac{5}{9}\right) = -\frac{1}{15}\,.$$

Part (b): The variance of the amount one wins can be computed by the standard expression for variance in term of expectations. Specifically we have

$$\text{Var}(W) = E[W^2] - E[W]^2\,.$$

Now using the results from Part (a) above we see that

$$E[W^2] = \frac{4}{9}(1.1)^2 + \frac{5}{9}(-1.0)^2 = \frac{82}{75}\,.$$

so that

$$\text{Var}(W) = \frac{82}{75} - \left(\frac{1}{15}\right)^2 = \frac{49}{45} \approx 1.08\,.$$

Problem 36 (the variance of the number of games played)

From Problem 22 we have that (for $i = 2$)

$$E[N^2] = 4(p^2 + q^2) + 9(2p^2q + 2q^2p) = 4 + 10p - 10p^2\,.$$

Thus the variance is given by

$$\begin{aligned} \text{Var}(N) &= E[N^2] - (E[N])^2 \\ &= 4 + 10 - 10p^2 - (2 + 2p - 2p^2)^2 \\ &= 2p(1 - 3p + 4p^2 - 2p^3)\,. \end{aligned}$$

Which has an inflection point at $p = 1/2$.

Problem 38 (evaluating expectations and variances)

Part (a): We find, expanding the quadratic and using the linearity property of expectations that

$$E[(2+X)^2] = E[4+4X+X^2] = 4+4E[X]+E[X^2]\,.$$

In terms of the variance, $E[X^2]$ is given by $E[X^2] = \text{Var}(X) + E[X]^2$, both terms of which we know from the problem statement. Using this the above becomes

$$E[(2+X)^2] = 4+4(1)+(5+1^2) = 14\,.$$

Part (b): We find, using properties of the variance that

$$\text{Var}(4+3X) = \text{Var}(3X) = 9\text{Var}(X) = 9\cdot 5 = 45\,.$$

Exercise 39 (drawing two white balls in four draws)

The probability of drawing a white ball is $3/6 = 1/2$. Thus if we consider event that we draw a white ball a success, the probability requested is that in four trials, two are found to be successes. This is equal to a binomial distribution with $n = 4$ and $p = 1/2$, thus our desired probability is given by

$$\binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{4-2} = \frac{6}{4\cdot 4} = \frac{3}{8}\,.$$

Problem 40 (guessing on a multiple choice exam)

With three possible answers possible for each question we have a $1/3$ chance of guessing any specific question correctly. Then the probability that the student gets four or more correct by guessing would be the required sum of a binomial distribution. Specifically we have

$$\binom{5}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^1 + \binom{5}{5}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^0 = \frac{11}{243}\,.$$

Where the first term is the probability the student guess four questions (from five) correctly and the second term is the probability that the student guesses all five questions correctly.

Problem 41 (proof of extrasensory perception)

Randomly guessing the man would get seven correct answers (out of ten) with probability

$$\binom{10}{7}\left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right)^3 = 0.11718\,.$$

Since the book then requests the probability that he does *at least this well* we would need to consider the probability that he gets eight, nine, or ten answers correct. This would be the sum of three more numbers computed in exactly the same way as above. This sum is 0.17188.

Problem 42 (failing engines)

The number of of engines that fail (or function) is a Binomial random variable with probability $1 - p$ and p respectively. For a three engine plane the probability that it makes a successful flight if the probability that three or two engines function (for in that case we will have a majority). This probability is

$$\left(\begin{array}{c} 3 \\ 3 \end{array} \right) p^3 (1-p)^0 + \left(\begin{array}{c} 3 \\ 2 \end{array} \right) p^2 (1-p)^1 = p^2(3 - 2p) \,.$$

For a five engine plane the probability that it makes a successful flight is the probability that five, four, or three engines function. This probability is

$$\left(\begin{array}{c} 5 \\ 5 \end{array} \right) p^5 (1-p)^0 + \left(\begin{array}{c} 5 \\ 4 \end{array} \right) p^4 (1-p)^1 + \left(\begin{array}{c} 5 \\ 3 \end{array} \right) p^3 (1-p)^2 = p^3(6p^2 - 15p + 10) \,.$$

Then a five engine plane will be preferred to a three engine plane if

$$p^3(6p^2 - 15p + 10) \geq p^2(3 - 2p) \,.$$

Putting all the p variable to one side of the inequality (and defining a function $f(\cdot)$) the above is equivalent to

$$f(p) \equiv 2p^3 = 5p^2 + 4p - 1 \geq 0 \,.$$

Plotting the function $f(p)$ in Figure **??** we see that it is positive for $\frac{1}{2} \leq p \leq 1$. Thus for values of p in this range we derive a benefit by using the five engine plane.

Problem 44 (will it function?)

Let F be the event that our system functions and R the event that it rains. Then by conditioning on whether it rains or not we have

$$\begin{aligned} P(F) &= P(F|R)P(R) + P(F|R^c)P(R^c) = \alpha P(F|R) + (1-\alpha) P(F|R^c) \\ &= \alpha \sum_{i=k}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) p_1^i (1-p_1)^{n-i} + (1-\alpha) \sum_{i=k}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) p_2^i (1-p_2)^{n-i} \,. \end{aligned}$$

Problem 46 (convictions)

Let E be the event that a jury renders a correct decision and let G be the event that a person is guilty. Then

$$P(E) = P(E|G)P(G) + P(E|G^c)P(G^c) \,.$$

From the problem statement we know that $P(G) = 0.65$ and $P(G^c) = 0.35$ so $P(E|G)$ is the probability we get the correct decision given that the defendant is guilty. To reach the correct decision we must have nine or more guilty votes so

$$P(E|G) = \sum_{i=9}^{12} \left(\begin{array}{c} 12 \\ i \end{array} \right) (0.8)^i (0.2)^{12-i},$$

which is the case where the person is guilty and at least nine members vote on this persons guilt. This is because

$$P(\text{Vote Guilty}|G) = 1 - P(\text{Vote Innocent}|G) = 1 - 0.2 = 0.8 .$$

Thus we can compute $P(E|G)$ using the above sum. We find $P(E|G) = 0.79457$. Now $P(E|G^c) = 1 - P(E^c|G^c)$ or one minus the probability the jury makes a mistake and votes an innocent man guilty. This is

$$P(E|G^c) = 1 - \sum_{i=9}^{12} \left(\begin{array}{c} 12 \\ i \end{array} \right) (0.1)^i (0.9)^{12-i},$$

Since $P(\text{Vote Guilty}|G^c) = 0.1$. The above can be computed and equals $P(E|G^c) \approx 1.0$, so that

$$P(E) = 0.79457(0.65) + 1.0(0.35) = 0.86647,$$

as the probability that the jury makes the correct decision. To calculate the percentage of defendants that are convected we need to compute

$$P(E|G)P(G) + P(E^c|G^c)P(G^c) = 0.79457(0.65) + 0.0(0.35) = 0.516 .$$

Some of the calculations for this problem can be done in the octave code `chap_4_prob_46.m`.

Problem 47 (military convictions)

Part (a): With nine judges we have

$$P(G) = \sum_{i=5}^{9} \left(\begin{array}{c} 9 \\ i \end{array} \right) (0.7)^i (0.3)^{9-i} = 0.901 .$$

With eight judges we have

$$P(G) = \sum_{i=5}^{8} \left(\begin{array}{c} 8 \\ i \end{array} \right) (0.7)^i (0.3)^{8-i} = 0.805 .$$

With seven judges we have

$$P(G) = \sum_{i=4}^{7} \left(\begin{array}{c} 7 \\ i \end{array} \right) (0.7)^i (0.3)^{7-i} = 0.8739 .$$

Part (b): For nine judges

$$P(G^c) = 1 - P(G) = 1 - \sum_{i=5}^{9} \binom{9}{i} (0.3)^i (0.7)^{9-i} = 0.901\,.$$

For eight judges we have

$$P(G^c) = 1 - \sum_{i=5}^{8} \binom{8}{i} (0.3)^i (0.7)^{8-i} = 0.942\,.$$

For seven judges we have

$$P(G^c) = 1 - \sum_{i=4}^{7} \binom{7}{i} (0.3)^i (0.7)^{7-i} = 0.873\,.$$

Part (c): Assume the defense attorney would like to free his or her client. Let D be the event that the client goes free then

$$P(D) = P(D|G)P(G) + P(D|G^c)P(G^c)\,,$$

with $P(D|G)$ indexed by the number of judges then

$$\begin{aligned} P(D|n=9) &= (1 - 0.9011)(0.6) + (0.901)(0.4) = 0.419 \\ P(D|n=8) &= (1 - 0.805)(0.6) + (0.942)(0.4) = 0.493 \\ P(D|n=7) &= (1 - 0.8739)(0.6) + (0.873)(0.4) = 0.429\,. \end{aligned}$$

Thus the defense attorney has the best chance of getting his client off if there are two judges and so he should request that one be removed.

Problem 48 (defective disks)

For this problem lets take the guarantee that the company provides to mean that a package will be considered "defective" if it has *more than* one defective disk. The probability that more than one disk in a pack is defective (P_d) is given by

$$P_d = 1 - \binom{10}{0} (0.01)^0 (0.99)^{10} - \binom{10}{1} (0.01)^1 (0.99)^9 \approx 0.0042\,,$$

since $\binom{10}{0} (0.01)^0 (0.99)^{10}$ is the probability that *no* disks are defective in the package of ten disks, and $\binom{10}{1} (0.01)^1 (0.99)^9$ is the probability that one of the ten disks is defective. If a customer buys three packs of disks the probability that he returns exactly one pack is the probability that from his three packs one package is defective. This is given by a binomial distribution with parameters $n = 3$ and $p = 0.0042$. We find this to be

$$\binom{3}{1} (0.0042)^1 (1 - 0.0042)^2 = 0.0126\,.$$

Problem 49 (flipping coins)

We are told in the problem statement that the event the first coin C_1, lands heads happens with probability 0.4, while the event that the second coin C_2 lands heads happens with probability 0.7.

Part (a): Let E be the event that exactly seven of the ten flips land on heads then conditioning on the initially drawn coin (either C_1 or C_2) we have

$$P(E) = P(E|C_1)P(C_1) + P(E|C_2)P(C_2)\,.$$

Now we can evaluate each of these conditional probabilities as

$$\begin{aligned} P(E|C_1) &= \binom{10}{7}(0.4)^7(0.6)^3 = 0.0424 \\ P(E|C_2) &= \binom{10}{7}(0.7)^7(0.3)^3 = 0.2668\,. \end{aligned}$$

So $P(E)$ is given by (assuming uniform probabilities on the coin we initially select)

$$P(E) = 0.5 \cdot 0.0424 + 0.5 \cdot 0.2668 = 0.1546\,.$$

Part (b): If we are told that the first three of the ten flips are heads then we desire to compute what is the conditional probability that exactly seven of the ten flips land on heads. To compute this let A be the event that the first three flips are heads. Then we want to compute $P(E|A)$, which we can do by conditioning on the initial coin selected, i.e.

$$P(E|A) = P(E|A, C_1)P(C_1) + P(E|A, C_2)P(C_2)\,.$$

Now as before we find that

$$\begin{aligned} P(E|A, C_1) &= \binom{7}{4}(0.4)^4(0.6)^3 = 0.1935 \\ P(E|A, C_2) &= \binom{7}{4}(0.7)^4(0.3)^3 = 0.2268\,. \end{aligned}$$

So the above probability is given by

$$P(E|A) = 0.5 \cdot 0.1935 + 0.5 \cdot 0.2668 = 0.2102\,.$$

Problem 52 (airplane crashes)

Assume that N the number of monthly airplane crashes is given by a Poisson random variable. Then since $E[N] = \lambda$ we know the parameter λ in the distribution. From this we can calculate the probabilities requested.

Part (a): We have

$$P(N \geq 2) = 1 - P(N < 2) = 1 - P(N = 0) - P(N = 1)$$
$$= 1 - e^{-3.5} - \frac{e^{-3.5}(3.5)}{1!} = 0.8641\,.$$

Part (b): We have

$$P(N \leq 1) = P(0) + P(1) = e^{-3.5} + \frac{e^{-3.5}(3.5)}{1!} = 0.1359\,.$$

Problem 55 (errors when typing)

Let A and B the event that the paper is typed by typist A or typist B respectively. Let E be the event that our article has a least one error, then

$$P(E) = P(E|A)P(A) + P(E|B)P(B)\,,$$

since both typist are equally likely $P(A) = P(B) = \frac{1}{2}$ and

$$P(E|A) = \sum_{i=1}^{\infty} P\{E = i|A\} = \sum_{i=1}^{\infty} e^{-\lambda_A} \frac{\lambda_A^i}{i!} = 1 - e^{-\lambda_A} = 1 - e^{-3} = 0.9502\,.$$

and in the same way

$$P(E|B) = 1 - e^{-4.2} = 0.985\,.$$

so that $P(E) = 0.5(0.9502) + 0.5(0.985) = 0.967$ so the probability of no errors is given by $1 - P(E) = 0.03239$

Problem 56 (at least two birthdays)

The probability that at least one person will have the same birthday as myself is the complement of the probability that no other person has a birthday equivalent to myself. An individual person will not have the same birthday as myself with probability $p = \frac{364}{365} = 0.997$. Thus the probability that n people all do not have my birthday is then p^n. So the probability that at least one person does have my birthday is given by $1 - p^n$. To have this greater than $1/2$ requires that $1 - p^n \geq 1/2$ or $p^n \leq 1/2$ or

$$n \geq \frac{\ln(1/2)}{\ln(p)} = \frac{\ln(2)}{\ln(365/364)} = 252.6\,.$$

To make n a integer take $n \geq 253$.

Problem 57 (accidents on a highway)

Part (a): $P\{X \geq 3\} = 1 - P\{X = 0\} - P\{X = 1\} - P\{X = 2\}$, with

$$P\{X = i\} = \frac{e^{-\lambda}\lambda^i}{i!}.$$

Then with $\lambda = 3$ we have

$$P\{X \geq 3\} = 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{1}{2}e^{-\lambda}\lambda^2 = 0.576.$$

Part (b): $P\{X \geq 3|X \geq\} = \frac{P\{X\geq 3, X\geq 1\}}{P\{X\geq 1\}} = \frac{P\{X\geq 3\}}{P\{X\geq 1\}}$. Now $P\{X \geq 1\} = 1 - e^{-\lambda} = 0.95$. So $P\{X \geq 3|X \geq 1\} = \frac{0.576}{0.95} = 0.607$.

Problem 61 (a full house)

The probability of obtaining i full houses from n ($n = 1000$) is given by a binomial random variable with $p = 0.0014$ and $n = 1000$. Thus the probability of obtaining at least two full house is

$$\begin{aligned}\sum_{i=2}^{n} \binom{n}{i} p^i(1-p)^{n-i} &= 1 - \sum_{i=0}^{1} \binom{n}{i} p^i(1-p)^{n-i} \\ &= 1 - \binom{1000}{0} p^0(1-p)^{1000} - \binom{1000}{1} p^1(1-p)^{999}.\end{aligned}$$

In this problem since $p = 0.0014$ and $n = 1000$ we have $pn = 1.4$ is rather small and we can use the Poisson approximation to the binomial distribution as

$$P\{X = i\} \approx \frac{e^{-\lambda}\lambda^i}{i!} \quad \text{with} \quad \lambda = pn = 1.4,$$

so the above probability is approximately $1 - e^{-1.4} - e^{-1.4}(1.4) = 0.408$.

Problem 62 (the probability that no wife sits next to her husband)

From Problem 66, the probability that couple i is selected next to each other is given by $\frac{2}{2n-1} = \frac{1}{n-1/2}$. Then we can approximate the probability that the total number of couples sitting together is a Poisson distribution with parameter $\lambda = n\frac{1}{n-1/2} = \frac{2n}{2n-1}$. Thus the probability that no wife sits next to her husband is given by evaluating a Poisson distribution with count equal to zero and $\lambda = \frac{2n}{2n-1}$ or

$$\exp\{-\frac{2n}{2n-1}\}.$$

When $n = 10$ this expression is $\exp\{-\frac{20}{19}\} \approx 0.349$. The exact formula is computed in example 5n from Chapter 2, where the exact probability is given as 0.3395 showing that our approximation is rather close.

Problem 63 (entering the casino)

Part (a): Lets assume the number of people entering the casino follows a Poisson approximation with rate $\lambda = 1$ person in two minutes. Then in our five minutes interval from 12:00 to 12:05 we have a Poisson random variable with parameter of $\lambda t = 1\left(\frac{5}{2}\right) = 2.5$, so the probability that no one enters in that five minutes is given by

$$P\{N = 0\} = e^{-2.5} = 0.0821\,.$$

Part (b): The probability that at least four people enter is

$$P\{N \geq 4\} = 1 - P\{N \leq 3\} = 1 - e^{-\lambda}\left[1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{6}\right] = 0.242\,,$$

when $\lambda = 2.5$.

Problem 64 (suicides)

Assume that the number of people who commit suicide is a Poisson random variable with suicide rate one per every 10^5 inhabitants per month.

Part (a): Since we have $4\,10^5$ inhabitants the suicide rate for the population will be $\lambda = 4 \cdot 1 = 4$. The desired probability is then

$$\begin{aligned} P\{N \geq 8\} &= 1 - P\{N \leq 7\} \\ &= 1 - e^{-4}\left[1 + \frac{4}{1} + \frac{4^2}{2} + \frac{4^3}{6} + \frac{4^4}{24} + \frac{4^5}{120} + \frac{4^6}{720} + \frac{4^7}{5040}\right] = 0.051\,. \end{aligned}$$

Part (b): If we now assume that $P\{N \geq 8\}$ computed above is independent from month to month we are looking for the probability this event happens at least twice or

$$\begin{aligned} &= 1 - \left(\begin{array}{c} 12 \\ 0 \end{array}\right)(1 - P\{N \geq 8\})^{12} - \left(\begin{array}{c} 12 \\ 1 \end{array}\right)P\{N \geq 8\}(1 - P\{N \geq 8\})^{11} \\ &= 1 - (1 - 0.511)^{12} - 12(0.0511)(1 - 0.0511)^{11} = 0.122\,. \end{aligned}$$

Part (c): I would assume that month to month, the event eight or more suicides would be independent and thus the probability that in the first month when eight suicides occurs would be given by a geometric random variable with parameter $p = P\{N \geq 8\} = 0.0511$. A geometric random variable represents the probability of repeating an experiment until a success occurs and is given by

$$P\{X = n\} = (1 - p)^{n-1}p \quad \text{for} \quad n = 1, 2, 3, \cdots$$

Problem 65 (the diseased)

Part (a): Since the probability that the number of soldiers with the given disease is a binomial distribution with parameters $(n, p) = (500, \frac{1}{10^3})$, we can approximate this distribution with a Poisson distribution with rate $\lambda = 500\frac{1}{10^3} = 0.5$. Then the required probability is given by

$$P\{N \geq 1\} = 1 - P\{N = 0\} = 1 - e^{-0.5} \approx 0.3934\,.$$

Part (b): We are now looking for

$$\begin{aligned}
P\{N \geq 2 | N > 0\} &= \frac{P\{N \geq 2\,, N > 0\}}{P\{N > 0\}} \\
&= \frac{1 - P\{N < 2\}}{P\{N > 0\}} \\
&\approx \frac{1 - e^{-0.5}(1 + 0.5)}{0.3934} \\
&= 0.2293\,.
\end{aligned}$$

Part (c): If Jones knows that he has the disease then the news that the test result comes back positive is not informative to him. Therefore he believes that the distribution of the number of men with the disease is binomial with parameters $(n, p) = (499, \frac{1}{10^3})$. As such, it can be approximated with a Poisson distribution with parameter $\lambda = np = \frac{499}{10^3} = 0.499$. Then to him the probability that more than one person has the disease is given by

$$P\{N \geq 2 | N > 0\} = 1 - P\{N < 1\} = 1 - e^{-0.499} \approx 0.3928\,.$$

Part (d): We desire to compute the probability that any of the $500 - i$ remaining people have the disease that is (with the number N the total number of people with the disease) let E be the event that the people $1, 2, 3, \cdots i - 1$ do not have the disease while i does the probability we desire is then

$$P\{N \geq 2 | E\} = \frac{P\{N \geq 2\,, E\}}{P\{E\}}\,.$$

Now the probability $P\{E\} = (1 - p)^i p$, since E is a geometric random variable. Now $P\{N \geq 2\,, E\}$ is the probability that since person i has the disease that at least one more person has the disease in the $M - i$ additional people (here $M = 500$) and is given by

$$\sum_{k=1}^{M-i} \binom{M-i}{k} p^k (1-p)^{M-i-k}$$

so this probability (the entire conditional probability) is then

$$P\{N \geq 2 | E\} = \frac{\sum_{k=1}^{M-i} \binom{M-i}{k} p^k (1-p)^{M-i-k}}{(1-p)^i p}\,,$$

which becomes (when we put the numbers for this problem in the expression above) the following

$$P\{N \geq 2|E\} = \frac{\sum_{k=1}^{500-i} \left(\begin{array}{c} 500-i \\ k \end{array} \right) \left(\frac{1}{10^3}\right)^k \left(1 - \frac{1}{10^3}\right)^{500-i-k}}{\left(1 - \frac{1}{10^3}\right)^i \left(\frac{1}{10^3}\right)}.$$

Problem 66 (seating couples next to each other)

Part (a): There are $(2n-1)!$ different possible seating orders around a circular table when each person is considered unique. For couple i to be seated next to each other, consider this couple as one unit, then we have in total now

$$2n - 2 + 1 = 2n - 1\,,$$

unique "items" to place around our table. Here an item can be an individual person or the ith couple considered as one unit. Specifically we have taken the total $2n$ people and subtracted the specific ith couple (of two people) and put back the couple considered as one unit (the plus one). Thus there are $(2n-1-1)! = (2n-2)!$ rotational orderings of the remaining $2n-2$ people and the "fused" couple. Since there are an additional ordering of the individual people in the pair, we have a total of $2(2n-2)!$ orderings where couple i is together. Thus our probability is given by

$$P(C_i) = \frac{2(2n-2)!}{(2n-1)!} = \frac{2}{2n-1}\,.$$

Part (b): To compute $P(C_j|C_i)$ when $j \neq i$ we note that it is equal to

$$\frac{P(C_j, C_i)}{P(C_i)}\,.$$

Here $P(C_j, C_i)$ is the joint probability where both couple i and couple j are together. Since we have evaluated $P(C_i)$ in Part a of this problem we will now evaluate $P(C_j, C_i)$ in the same way as earlier. With couple i and j considered as individual units, the number of "items" we have to distribute around our table is given by

$$2n - 2 + 1 - 2 + 1 = 2n - 2\,.$$

Here as before we subtract the individual people in the couple and then add back in a "fused" couple considered as one unit. Thus the number of unique permutations of these items around our table is given by $4(2n-2-1)! = 4(2n-3)!$. The factor of four is for the different orderings of the husband and wife in each fused pair. Thus our our joint probability is then given by

$$P(C_j, C_i) = \frac{4(2n-3)!}{(2n-1)!} = \frac{2}{(2n-1)(n-1)}\,,$$

so that our conditional probability $P(C_j|C_i)$ is given by

$$P(C_j|C_i) = \frac{2/(2n-1)(n-1)}{2/(2n-1)} = \frac{1}{n-1}\,.$$

Part (c): When n is large we want to approximate $1 - P(C_1 \cup C_2 \cup \ldots \cup C_n)$, which is given by

$$\begin{aligned} 1 - P(C_1 \cup C_2 \cup \ldots \cup C_n) &= 1 - \left(\sum_{i=1}^{n} P(C_i) - \sum_{i<j} P(C_i, C_j) + \cdots \right) \\ &= 1 - \left(\frac{2n}{2n-1} - \sum_{i<j} P(C_j|C_i) P(C_i) + \cdots \right) \\ &= 1 - \left(\frac{2n}{2n-1} - \binom{n}{2} \frac{2}{(2n-1)(n-1)} + \cdots \right) \end{aligned}$$

But since $P(C_j|C_i) = \frac{1}{n-1} \approx \frac{1}{n-1/2} = P(C_j)$, when n is very large. Thus while the events C_i and C_j are not independent, their dependence is weak for large n. Thus by the Poisson paradigm we can expect the number of couples sitting together to have a Poisson approximation with rate $\lambda = n\left(\frac{2}{2n-1}\right) \approx 1$. Thus the probability that no married couple sits next to each other is $P\{N = 0\} = e^{-1}$.

Problem 72 (playing a series of games)

Let team A have the larger probability of winning of 0.6 and let E_i be the event that team A wins the tournament in i games. We then calculate

$$P(E_4) = (0.6)^4 = 0.1296 .$$

To calculate $P(E_i)$ for $i \geq 5$ note that since A must win the tournament that team must win the last game played. Thus

$$\begin{aligned} P(E_5) &= P(A \text{ wins the last game}) P(A \text{ wins 3 of the previous 4 games}) \\ &= 0.6 \left(\binom{4}{3} 0.6^3 \, 0.4 \right) = 0.2074 . \end{aligned}$$

In the same way we have

$$\begin{aligned} P(E_6) &= P(A \text{ wins the last game}) P(A \text{ wins 3 of the previous 5 games}) \\ &= 0.6 \left(\binom{5}{3} 0.6^3 \, 0.4^2 \right) = 0.2074 \\ P(E_7) &= P(A \text{ wins the last game}) P(A \text{ wins 3 of the previous 6 games}) \\ &= 0.6 \left(\binom{6}{3} 0.6^3 \, 0.4^3 \right) = 0.1659 . \end{aligned}$$

The total probability that team A wins the tournament is then given by the sum of these probabilities or

$$P(\text{first to win 4 games}) = 0.1296 + 0.2074 + 0.2074 + 0.1659 = 0.7102 .$$

The probability of winning 2-out-of-3 tournament is given by

$$P(\text{win 2-out-of 3 games}) = (0.6)^2 + (0.6)\left(\begin{array}{c} 2 \\ 1 \end{array}\right) 0.6^1\, 0.4^1 = 0.648\,.$$

Thus from the numbers above we see that

$$P(\text{win 2-out-of 3 games}) < P(\text{first to win 4 games})\,,$$

thus the stronger team would prefer the tournament where four games are played.

Problem 85 (k types of coupons)

The probability of collecting a coupon of type i is p_i. Let X be the total number of distinct coupons collected and let X_i be an indicator random variable such that $X_i = 1$ if after collecting our n coupons we have a coupon of type i (X_i is 0 otherwise). Then

$$\begin{aligned} P\{X_i = 1\} &= 1 - P\{X_i = 0\} = 1 - P(\text{we never collect a coupon of type } i) \\ &= 1 - (1 - p_i)^n\,. \end{aligned}$$

Then we can express X as $X = \sum_{i=1}^{k} X_i$ which gives the expectation of

$$E[X] = \sum_{i=1}^{k} E[X_i] = \sum_{i=1}^{k} (1 - (1 - p_i)^n) = k - \sum_{i=1}^{k} (1 - p_i)^n\,.$$

Chapter 4: Theoretical Exercises

Problem 6 (the sum of cumulative probabilities)

Consider the claimed expression for $E[N]$ that is $\sum_{i=1}^{\infty} P\{N \geq i\}$. Now $P\{N \geq i\} = \sum_{k=i}^{\infty} P\{N = k\}$ and inserting this into the above summation gives

$$\sum_{i=1}^{\infty} \sum_{k=i}^{\infty} P\{N = k\}\,.$$

We can graphically represent this summation in the (i, k) plane as where the summation is done along the i axis first and then along the k axis i.e. we sum by columns upward first. Now changing the order of the the summation to sum along rows first row i.e. k is the outer index and i is the inner index we have that the above is equivalent to

$$\sum_{k=1}^{\infty} \sum_{i=1}^{k} P\{N = k\}\,.$$

Which can be written (since $P\{N = k\}$ does not depend on the index i) as

$$\sum_{k=1}^{\infty} k P\{N = k\} = E[N]\,,$$

as expected.

Problem 7 (the first moments of cumulative probabilities)

Following the hint we have

$$\sum_{i=0}^{\infty} iP\{N > i\} = \sum_{i=0}^{\infty} i \sum_{k=i+1}^{\infty} P\{N = k\} = \sum_{i=0}^{\infty} \sum_{k=i+1}^{\infty} iP\{N = k\}\,.$$

Since $P\{N > i\} = \sum_{k=i+1}^{\infty} P\{N = k\}$ when N is an integral valued random variable. Now to proceed we will change the order of summation. This can be explained by graphically denoting the summation points in the (i, k) plane. Now in the formulation given above the summation is done in columns moving from left to right in the triangle of points above. Equivalently we will instead perform our summation over rows in the triangle of points. Doing this we would have

$$\sum_{k=1}^{\infty} \sum_{i=0}^{k-1} iP\{N = k\}\,.$$

Where the outer sum represents selecting the individual rows and the inner sum the summation across that row. This sum simplifies to

$$\begin{aligned}
\sum_{k=1}^{\infty} P\{N = k\} \sum_{i=0}^{k-1} i &= \sum_{k=1}^{\infty} P\{N = k\} \left(\frac{k(k-1)}{2} \right) \\
&= \frac{1}{2} \left(\sum_{k=1}^{\infty} k^2 P\{N = k\} - \sum_{k=1}^{\infty} kP\{N = k\} \right) \\
&= \frac{1}{2} \left(E[N^2] - E[N] \right)\,,
\end{aligned}$$

as requested.

Problem 8 (an exponential expectation)

If X is a random variable such that $P\{X = 1\} = p = 1 - P\{X = -1\}$, then

$$E[e^X] = P\{X = 1\}c + P\{X = -1\}c^{-1} = 1\,.$$

With the above information we have that this becomes $pc + (1 - p)c^{-1} = 1$. On solving for c (using the quadratic equation) we get that

$$c = \frac{1 \pm |1 - 2p|}{2p}\,.$$

Thus we have (taking the two possible signs) that

$$c = \begin{cases} \frac{1+(1-2p)}{2p} = \frac{1-p}{p} \\ \frac{1-(1-2p)}{2p} = 1 \end{cases}\,.$$

These are the two possible values for c. Since $c \neq 1$ then we must have $c = \frac{1-p}{p}$.

Problem 9 (the expected value of standardized variables)

Define our random variable Y by $Y = \frac{X-\mu}{\sigma}$ then

$$\begin{aligned} E[Y] &= \sum_i \left(\frac{x_i - \mu}{\sigma}\right) p(x_i) = \frac{1}{\sigma}\sum_i (x_i - \mu) p(x_i) \\ &= \frac{1}{\sigma}\left(\sum_i x_i p(x_i) - \mu \sum_i p(x_i)\right) = \frac{1}{\sigma}(E[X] - \mu) = 0\,. \end{aligned}$$

And also

$$\begin{aligned} E[Y^2] &= \sum_i \left(\frac{x_i - \mu}{\sigma}\right)^2 p(x_i) = \sum_i \left(\frac{x_i^2 - 2x_i\mu + \mu^2}{\sigma^2}\right) p(x_i) \\ &= \frac{1}{\sigma^2}\left(\sum_i x_i^2 p(x_i) - 2\mu \sum_i x_i p(x_i) + \mu^2 \sum_i p(x_i)\right)\,. \end{aligned}$$

Now since $E[X^2] = \sum_i x_i^2 p(x_i)$, $E[X] = \sum_i x_i p(x_i)$, and $\sum_i p(x_i) = 1$, we see that

$$E[Y^2] = \frac{1}{\sigma^2}\left(E[X^2] - 2\mu E[X] + \mu^2\right)\,.$$

Since $\mu = E[X]$ the above becomes

$$\text{Var}(Y) = \frac{1}{\sigma^2}\left(E[X^2] - 2\mu^2 + \mu^2\right) = \frac{1}{\sigma^2}\left(E[X^2] - \mu^2\right) = \frac{1}{\sigma^2}\sigma^2 = 1\,.$$

Where we have used the fact that $\text{Var}(X) = E[X^2] - E[X]^2 = \sigma^2$.

Another way to solve this problem that might be simpler to understand is to just use the some of the basic properties of expectation and variance. For example we have

$$E[Y] = E\left[\frac{X - \mu}{\sigma}\right] = \frac{1}{\sigma}(E[X] - \mu) = 0\,.$$

While for the variance

$$\text{Var}(Y) = \text{Var}\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma^2}\text{Var}(X - \mu) = \frac{1}{\sigma^2}\text{Var}(X) = \frac{\sigma^2}{\sigma^2} = 1\,.$$

Problem 10 (an expectation with a binomial random variable)

If X is a binomial random variable with parameters (n, p) then

$$\begin{aligned} E\left[\frac{1}{X+1}\right] &= \sum_{k=0}^{n} \left(\frac{1}{k+1}\right) P\{X = k\} \\ &= \sum_{k=0}^{n} \left(\frac{1}{k+1}\right) \binom{n}{k} p^k (1-p)^{n-k}\,. \end{aligned}$$

Factoring out $1/(n+1)$ we obtain

$$E\left[\frac{1}{X+1}\right] = \frac{1}{n+1}\sum_{k=0}^{n}\left(\frac{n+1}{k+1}\right)\left(\begin{array}{c} n \\ k \end{array}\right) p^k(1-p)^{n-k}.$$

This result is beneficial since if we now consider the fraction and the n choose k term we see that

$$\left(\frac{n+1}{k+1}\right)\left(\begin{array}{c} n \\ k \end{array}\right) = \left(\frac{n+1}{k+1}\right)\frac{n!}{k!(n-k)!} = \frac{(n+1)!}{(k+1)!(n-k)!} = \left(\begin{array}{c} n+1 \\ k+1 \end{array}\right).$$

This substitution turns our summation into the following

$$E\left[\frac{1}{X+1}\right] = \frac{1}{n+1}\sum_{k=0}^{n}\left(\begin{array}{c} n+1 \\ k+1 \end{array}\right) p^k(1-p)^{n-k}.$$

the following manipulations allow us to evaluate this summation. We have

$$\begin{aligned}
E\left[\frac{1}{X+1}\right] &= \frac{1}{p(n+1)}\sum_{k=0}^{n}\left(\begin{array}{c} n+1 \\ k+1 \end{array}\right) p^{k+1}(1-p)^{n+1-(k+1)} \\
&= \frac{1}{p(n+1)}\sum_{k=1}^{n+1}\left(\begin{array}{c} n+1 \\ k \end{array}\right) p^{k}(1-p)^{n+1-k} \\
&= \frac{1}{p(n+1)}\left[\sum_{k=0}^{n+1}\left(\begin{array}{c} n+1 \\ k \end{array}\right) p^{k}(1-p)^{n+1-k} - (1-p)^{n+1}\right] \\
&= \frac{1}{p(n+1)}(1-(1-p)^{n+1}) \\
&= \frac{1-(1-p)^{n+1}}{p(n+1)},
\end{aligned}$$

as we were to show.

Problem 11 (each sequence of k successes is equally likely)

Each specific instance of k success and $n-k$ failures has probability $p^k(1-p)^{n-k}$. Since each success occurs with probability p each failure occurs with probability $1-p$. As each arraignment has the same number of p's and $1-p$'s each has the *same* probability.

Problem 12

Warning: Here are some notes that I had lying around on this problem. I should state that I've not had the time I would like to fully verify this solution. Caveat emptor.

n antennas with m defective, and $n-m$ functioning. Then we have $\left(\begin{array}{c} n-m+1 \\ m \end{array}\right)$ is the number of orderings with no two defective antenna consecutive from a total of n antennas.

Therefore the probability of finding m defective antennas from n with no two defective antennas consecutive is

$$P\{M=m\} = \frac{\left(\begin{array}{c} n-m+1 \\ m \end{array}\right)}{n-m+1}.$$

If we don't worry about consecutive defective antennas. The number of orderings of consecutive defective antennas is given by $(n-m+1)^m$ so that

$$P\{M=m\} = \frac{\left(\begin{array}{c} n-m+1 \\ m \end{array}\right)}{(n-m+1)^m}.$$

Thus the probability that no two neighboring components are non functional can be obtained by conditioning on the number of defective components

$$\begin{aligned} P(A) &= \sum_{m=0}^{n} P(A|M=m)P(M=m) \\ &= \sum_{m=0}^{n} \frac{\left(\begin{array}{c} n-m+1 \\ m \end{array}\right)}{(n-m+1)^m} \left(\begin{array}{c} n \\ m \end{array}\right) (1-p)^m p^{n-m}. \end{aligned}$$

Problem 13 (maximum likelihood estimation with a binomial random variable)

Since X is a binomial random variable with parameters (n, p) we have that

$$P\{X=k\} = \left(\begin{array}{c} n \\ k \end{array}\right) p^k (1-p)^{n-k}.$$

Then the p that maximizes this expression is given by taking the derivative of the above (with respect to p) setting the resulting expression equal to zero and solving for p. We find that this derivative is given by

$$\frac{d}{dp} P\{X=k\} = \left(\begin{array}{c} n \\ k \end{array}\right) k p^{k-1} (1-p)^{n-k} + \left(\begin{array}{c} n \\ k \end{array}\right) p^k (1-p)^{n-k-1} (n-k)(-1).$$

Which when set equal to zero and solve for p we find that $p = \frac{k}{n}$, or the empirical counting estimate of the probability of success.

Problem 14 (having children)

Now $P\{X=n\} = \alpha p^n$ so imposing $\sum_{n=0}^{\infty} p(n) = 1$ requires that

$$\alpha \sum_{n \geq 0} p^n = 1 \Rightarrow \alpha \frac{1}{1-p} = 1 \Rightarrow \alpha = 1-p.$$

so that $P\{X = n\} = (1-p)p^n$.

Part (a): The proportions of families with no children is $P\{X = 0\} = 1 - p$

Part (b): We have that

$$P\{B = k\} = \sum_{i \geq k} P\{B = k|X = i\}P\{X = i\} \, ,$$

where in computing $P\{B = k\}$ we have conditioned on the number of children a given family has. Now we know $P\{X = i\} = (1-p)p^i$. In addition, $P\{B = k|X = i\} = \left(\begin{array}{c} i \\ k \end{array} \right) \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{i-k} = \left(\begin{array}{c} i \\ k \end{array} \right) \left(\frac{1}{2}\right)^i$. This later probability is because the probability that we have k boys (given that we have i children) is a binomial random variable with probability of success 1/2. Combining these two results we find

$$\begin{aligned} P\{B = k\} &= \sum_{i \geq k} \left(\begin{array}{c} i \\ k \end{array} \right) \left(\frac{1}{2}\right)^i (1-p)p^i = (1-p) \sum_{i \geq k} \left(\begin{array}{c} i \\ k \end{array} \right) \left(\frac{p}{2}\right)^i \\ &= (1-p) \left[\left(\frac{p}{2}\right)^k + \left(\begin{array}{c} k+1 \\ k \end{array} \right) \left(\frac{p}{2}\right)^{k+1} + \left(\begin{array}{c} k+2 \\ k \end{array} \right) \left(\frac{p}{2}\right)^{k+2} + \cdots \right] . \end{aligned}$$

Problem 15

Warning: Here are some notes that I had lying around on this problem. I should state that I've not had the time I would like to fully verify this solution. Caveat emptor.

Let P_n be the probability that we we obtain an even number heads in n flips. Now conditioning on the results of the first flip we find that

$$P_n = p(1 - P_{n-1}) + (1-p)P_{n-1} \, .$$

To explain this, the first term $p(1 - P_{n-1})$ is the probability we get a head (p) times the probability that we have an odd number of heads in $n-1$ flips. The second term $(1-p)P_{n-1}$ is the probability we have a tail times the probability of an even number of heads in $n-1$ tosses. The above simplifies to

$$p + (1-2p)P_{n-1} \, .$$

We can check that the suggested expression satisfies this recurrence relationship. That is we ask if

$$\begin{aligned} \frac{1}{2}(1 + (1-2p)^n) &= p + (1-2p)(\frac{1}{2}(1 + (1-2p)^{n-1}) \\ &= p + \frac{1-2p}{p} + \frac{1}{2}(1-2p)^n = \frac{1}{2}(1 + (1-2p)^n) \, , \end{aligned}$$

giving a true identity. This result should also be able to be shown by explicitly enumerated all n tosses with an even number if heads as done in the book.

Problem 16 (the location of the maximum of the Poisson distribution)

Since X is a Poisson random variable the probability mass function for X is given by

$$P\{X = i\} = \frac{e^{-\lambda}\lambda^i}{i!}.$$

Following the hint we compute the requested fraction. We find that

$$\frac{P\{X = i\}}{P\{X = i-1\}} = \left(\frac{e^{-\lambda}\lambda^i}{i!}\right)\left(\frac{(i-1)!}{e^{-\lambda}\lambda^{i-1}}\right) = \frac{\lambda}{i}.$$

Now from the above expression if $i < \lambda$ then the "lambda" fraction $\frac{\lambda}{i} > 1$, meaning that the probabilities satisfy $P\{X = i\} > P\{X = i-1\}$ which implies that $P\{X = i\}$ is increasing for these values of i. On the other hand if $i > \lambda$ then $\frac{\lambda}{i} < 1$ we $P\{X = i\} < P\{X = i-1\}$ and $P\{X = i\}$ is decreasing for these values of i. Thus when $i < \lambda$, our probability $P\{X = i\}$ is increasing, while when $i > \lambda$, our probability $P\{X = i\}$ is decreasing. From this we see that the maximum of $P\{X = i\}$ is then when i is the largest integer still less than or equal to λ.

Problem 17 (the probability of an even Poisson sample)

Since X is a Poisson random variable the probability mass function for X is given by

$$P\{X = i\} = \frac{e^{-\lambda}\lambda^i}{i!}.$$

To help solve this problem it is helpful to recall that a binomial random variable with parameters (n, p) can be approximated by a Poisson random variable with $\lambda = np$, and that this approximation improves as $n \to \infty$. To begin then, let E denote the event that X is even. Then to evaluate the expression $P\{E\}$ we will use the fact that a binomial random variable can be approximated by a Poisson random variable. When we consider X to be a binomial random variable we have from theoretical Exercise 15 in this chapter that

$$P\{E\} = \frac{1}{2}(1 + (q-p)^n).$$

Using the Poisson approximation to the binomial we will have that $p = \lambda/n$ and $q = 1 - p = 1 - \lambda/n$, so the above expression becomes

$$P\{E\} = \frac{1}{2}\left(1 + \left(1 - \frac{2\lambda}{n}\right)^n\right).$$

Taking n to infinity (as required to make the binomial approximation by the Poisson distribution exact) and remembering that

$$\lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n = e^x,$$

the probability $P\{E\}$ above goes to

$$P\{E\} = \frac{1}{2}\left(1 + e^{-2\lambda}\right),$$

as we were to show.

Part (b): To directly evaluate this probability consider the summation representation of the requested probability, i.e.

$$\begin{aligned} P\{E\} &= \sum_{i=0,2,4,\cdots}^{\infty} \frac{e^{-\lambda}\lambda^i}{i!} \\ &= e^{-\lambda}\sum_{i=0}^{\infty} \frac{\lambda^{2i}}{(2i)!} \,. \end{aligned}$$

When we look at this it looks like the Taylor expansion of $\cos(\lambda)$ but without the required alternating $(-1)^i$ factor. This observation might trigger the recollection that the above series is in fact the Taylor expansion of the $\cosh(\lambda)$ function. This can be seen from the definition of the cosh function which is

$$\cosh(\lambda) = \frac{e^{\lambda} + e^{-\lambda}}{2} \,,$$

when one Taylor expands the exponentials on the right hand side of the above expression. Thus the above probability for $P\{E\}$ is given by

$$e^{-\lambda}\left(\frac{1}{2}(e^{\lambda} + e^{-\lambda})\right) = \frac{1}{2}(1 + e^{-2\lambda}) \,,$$

as claimed.

Problem 18 (maximizing λ in a Poisson distribution)

If X is a Poisson random variable then $P\{X = k\} = \frac{e^{-\lambda}\lambda^k}{k!}$. Now to determine the value of λ that maximizes this expression we differentiate $P\{X = k\}$ with respect to λ and set the resulting expression equal to zero. We have the derivative (equated equal to zero) given by

$$-\frac{e^{-\lambda}\lambda^k}{k!} + \frac{e^{-\lambda}k\lambda^{k-1}}{k!} = 0 \,.$$

or

$$-\lambda^k + k\lambda^{k-1} = 0 \,.$$

Since $\lambda \neq 0$, we have $\lambda = k$. We should check this value is indeed a maximum by computing the second derivative of $P\{X = k\}$ and showing that when $\lambda = k$ it is negative.

Problem 19 (the Poisson expectation of powers)

If X is a Poisson random variable then from the definition of the expectation we have that

$$E[X^n] = \sum_{i=0}^{\infty} i^n e^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda}\sum_{i=0}^{\infty}\frac{i^n\lambda^n}{i!}e^{-\lambda} = \sum_{i=1}^{\infty}\frac{i^n\lambda^i}{i!} \,,$$

since (assuming $n \neq 0$) when $i = 0$ the first term vanishes. Continuing our calculation we can cancel a factor of i and find that

$$\begin{aligned} E[X^n] &= e^{-\lambda} \sum_{i=1}^{\infty} \frac{i^{n-1}\lambda^i}{(i-1)!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{(i+1)^{n-1}\lambda^{i+1}}{i!} \\ &= \lambda \sum_{i=0}^{\infty} \frac{(i+1)^{n-1}e^{-\lambda}\lambda^i}{i!} . \end{aligned}$$

Now this sum can be recognized as the expectation of the variable $(X+1)^{n-1}$ so we see that

$$E[X^n] = \lambda E[(X+1)^{n-1}] .$$

From the result we have

$$E[X^3] = \lambda E[(X+1)^2] = \lambda^2 E[(X+2)] = \lambda^2(\lambda+2) = \lambda^3 + 2\lambda .$$

Problem 20 (flipping many coins)

Part (a): The total event of tossing all n coins is equivalent to to performing n, Bernoulli trials with a probability of success equal on each trial of p. Thus the the total number of successes is a Binomial random variable which we know can be approximated well as a Poisson random variable i.e.

$$P\{X = i\} \approx \frac{e^{-\lambda}\lambda^i}{i!} ,$$

so that $P\{X = 1\} \approx e^{-\lambda}\lambda$, thus the reasoning is correct.

Part (b): This is false since $P\{X = 1\}$ is the probability that only one head appears, when we have no other information about the number of heads. The expression $P\{Y = 1|Y > 0\}$ is the probability we have one head given that we know at least one head appears. Since before the experiment begins we don't know that we will have at least one head we can't condition on that fact.

Part (c): This is not true since $P\{X = 1\}$ is the probability that any one but one of the n trials result in a one, while $P\{Y = 0\}$ is the probability that a fixed set of $n-1$ flips results in no heads. That is we don't allow the set of $n-1$ of flips chosen to change. If we did, then we have n choices for which flip lands heads giving $ne^{-\lambda}$ and a probability p that the chosen position does indeed give a head, giving a combined probability $pne^{-\lambda} = \lambda e^{-\lambda}$ which is the correct answers.

Problem 21 (the birthdays of i and j)

Part (a): The events $E_{3,4}$ and $E_{1,2}$ would be independent since they consider different people so

$$P(E_{3,4}|E_{1,2}) = P(E_{3,4}) = \frac{1}{365} .$$

Part (b): Now $E_{1,3}$ and $E_{1,2}$ are still independent since if persons one and two have the same birthday, this information tells us nothing about the coincidence of the birthdays of persons one and three. Now since $E_{1,2}$ means that person one and two share the same birthday (one of the 365 days) then since person three must have this exact same day as his birthday we see that $P(E_{1,3}|E_{1,2}) = \frac{1}{365}$.

Part (c): Now $E_{2,3}$ and $E_{1,2} \cap E_{1,3}$ are not independent since the sets depend on all the same people. Then since person one and two have the same birthday as persons one and three, it follows that two and three have the same birthday. This means

$$P(E_{2,3}|E_{1,2} \cap E_{1,3}) = 1 .$$

Since the probabilities E_{ij} given other pairings can jump from $\frac{1}{365}$ to 1 we can conclude that these events are not independent. To be independent would require

$$P(E_{2,3}|E_{1,2} \cap E_{1,3}) = \frac{P(E_{2,3} \cap E_{1,2} \cap E_{1,3})}{P(E_{1,2} \cap E_{1,3})} = \frac{P(E_{2,3})P(E_{1,2})P(E_{1,3})}{P(E_{1,2})P(E_{1,3})} = P(E_{2,3}) .$$

But the left hand side of the above expression is equal to 1 while the right hand side is equal to $\frac{1}{365}$. As these two are not equal the events E_{ij} are not independent.

Problem 25 (events registered with probability p)

We can solve this problem by conditioning on the number of true events (from the original Poisson random variable N) that occur. We begin by letting M be the number of events counted by our "filtered" Poisson random variable. Then we want to show that M is another Poisson random variable with parameter λp. To do so consider the probability that M has counted j "filtered events", by conditioning on the number of observed events from the original Poisson random variable. We find

$$P\{M = j\} = \sum_{n=0}^{\infty} P\{M = j|N = n\} \left(\frac{e^{-\lambda}\lambda^n}{n!} \right)$$

The conditional probability in this sum can be computed using the acceptance rule defined above. For if we have n original events the number of derived events is a binomial random variable with parameters (n, p). Specifically then we have

$$P\{M = j|N = n\} = \begin{cases} \binom{n}{j} p^j (1-p)^{n-j} & j \le n \\ 0 & j > n . \end{cases}$$

Putting this result into the original expression for $P\{M = j\}$ we find that

$$P\{M = j\} = \sum_{n=j}^{\infty} \binom{n}{j} p^j (1-p)^{n-j} \left(\frac{e^{-\lambda}\lambda^n}{n!} \right)$$

To evaluate this we note that $\left(\begin{array}{c} n \\ j \end{array} \right) \frac{1}{n!} = \frac{1}{j!(n-j)!}$, so that the above simplifies as following

$$
\begin{aligned}
P\{M = j\} &= \frac{e^{-\lambda} p^j}{j!} \sum_{n=j}^{\infty} \frac{1}{(n-j)!} (1-p)^{n-j} \lambda^n \\
&= \frac{e^{-\lambda} p^j}{j!} \sum_{n=j}^{\infty} \frac{1}{(n-j)!} (1-p)^{n-j} (\lambda)^j \lambda^{n-j} \\
&= \frac{e^{-\lambda} (p\lambda)^j}{j!} \sum_{n=j}^{\infty} \frac{((1-p)\lambda)^{n-j}}{(n-j)!} \\
&= \frac{e^{-\lambda} (p\lambda)^j}{j!} \sum_{n=0}^{\infty} \frac{((1-p)\lambda)^{n}}{n!} \\
&= \frac{e^{-\lambda} (p\lambda)^j}{j!} e^{(1-p)\lambda} = e^{-p\lambda} \frac{(p\lambda)^j}{j!} \,,
\end{aligned}
$$

from which we can see M is a Poisson random variable with parameter λp as claimed.

Problem 26 (an integral expression for the CDF of a Poisson random variable)

We will begin by evaluating $\int_{\lambda}^{\infty} e^{-x} x^n dx$. To perform repeated integration by parts we remember the integration by parts "formula" $udv = uv - vdu$, and in the following we will let u be the polynomial in x and dv the exponential. To start this translates into letting $u = x^n$ and $dv = e^{-x}$, and we have

$$
\begin{aligned}
\int_{\lambda}^{\infty} e^{-x} x^n dx &= -x^n e^{-x} \Big|_{\lambda}^{\infty} + \int_{\lambda}^{\infty} n x^{n-1} e^{-x} dx \\
&= \lambda^n e^{-\lambda} + n \int_{\lambda}^{\infty} x^{n-1} e^{-x} dx \\
&= \lambda^n e^{-\lambda} + n \left[-x^{n-1} e^{-x} \Big|_{\lambda}^{\infty} + \int_{\lambda}^{\infty} (n-1) x^{n-2} e^{-x} dx \right] \\
&= \lambda^n e^{-\lambda} + n \lambda^{n-1} e^{-\lambda} + n(n-1) \int_{\lambda}^{\infty} x^{n-2} e^{-x} dx \,.
\end{aligned}
$$

Continuing to perform one more integration by parts (so that we can fully see the pattern) we have that this last integral given by

$$
\begin{aligned}
\int_{\lambda}^{\infty} x^{n-2} e^{-x} dx &= -x^{n-2} e^{-x} \Big|_{\lambda}^{\infty} + \int_{\lambda}^{\infty} (n-2) x^{n-3} e^{-x} dx \\
&= \lambda^{n-2} e^{-\lambda} + (n-2) \int_{\lambda}^{\infty} x^{n-3} e^{-x} dx \,.
\end{aligned}
$$

Then we have for our total integral the following

$$
\begin{aligned}
\int_{\lambda}^{\infty} e^{-x} x^n dx &= \lambda^n e^{-\lambda} + n \lambda^{n-1} e^{-\lambda} + n(n-1) \lambda^{n-2} e^{-\lambda} \\
&+ n(n-1)(n-2) \int_{\lambda}^{\infty} x^{n-3} e^{-x} dx \,.
\end{aligned}
$$

Using mathematical induction the total pattern can be seen as

$$
\begin{aligned}
\int_{\lambda}^{\infty} e^{-x}x^n dx &= \lambda^n e^{-\lambda} + n\lambda^{n-1}e^{-\lambda} + n(n-1)\lambda^{n-2}e^{-\lambda} + \cdots \\
&+ n(n-1)(n-2)\cdots(n-k)\int_{\lambda}^{\infty} x^{n-k}e^{-x}dx \\
&= \lambda^n e^{-\lambda} + n\lambda^{n-1}e^{-\lambda} + n(n-1)\lambda^{n-2}e^{-\lambda} + \cdots + n!\int_{\lambda}^{\infty} e^{-x}dx \\
&= \lambda^n e^{-\lambda} + n\lambda^{n-1}e^{-\lambda} + n(n-1)\lambda^{n-2}e^{-\lambda} + \cdots + n!e^{-x}\,.
\end{aligned}
$$

When we divide this sum by $n!$ we find it is given by

$$
\frac{\lambda^n}{n!}e^{-\lambda} + \frac{\lambda^{n-1}}{(n-1)!}e^{-\lambda} + \frac{\lambda^{n-2}}{(n-2)!}e^{-\lambda} + \cdots + \lambda e^{-\lambda} + e^{-\lambda}
$$

or the left hand side of the expression given in the problem statement i.e.

$$
\sum_{i=0}^{n} \frac{e^{-\lambda}\lambda^i}{i!}\,,
$$

as we were to show.

Problem 29 (ratios of hypergeometric probabilities)

For a Hypergeometric random variable we have

$$
P(i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \quad \text{for} \quad i = 0, 1, \cdots, m\,.
$$

So that the requested ratio is given by

$$
\begin{aligned}
\frac{P(k+1)}{P(k)} &= \frac{\binom{m}{k+1}\binom{N-m}{n-k-1}}{\binom{N}{n}} \cdot \frac{\binom{N}{n}}{\binom{m}{k}\binom{N-m}{n-k}} \\
&= \frac{\binom{m}{k+1}}{\binom{m}{k}} \frac{\binom{N-m}{n-k-1}}{\binom{N-m}{n-k}} \\
&= \frac{\frac{m!}{(k+1)!(m-k-1)!} \cdot \frac{(N-m)!}{(n-k-1)!(N-m-n+k+1)!}}{\frac{m!}{k!(m-k)!} \cdot \frac{(N-m)!}{(n-k)!(N-m-n+k)!}} \\
&= \frac{k!(m-k)!}{(k+1)!(m-k-1)!} \cdot \frac{(n-k)!(N-m-n+k)!}{(n-k-1)!(N-m-n+k+1)!} \\
&= \frac{(m-k)(n-k)}{(k+1)(N-m-n+k+1)}\,.
\end{aligned}
$$

Problem 32 (repeated draws from a jar)

At each draw, the boy has a memory bank of numbers he has seen and X is the random variable that determines the number of draws before he sees a number twice (once to fill his memory and then viewed a second time). Then $F(x) = P\{X \leq x\}$, now

$$\begin{aligned} F(1) &= P\{X \leq 1\} = 0 \\ F(2) &= P\{X \leq 2\} = \frac{1}{n}, \end{aligned}$$

since he has seen only one chip and therefore has a $1/n$ chance of redrawing this chip. Now

$$F(3) = P\{X \leq 2\} + P\{X = 3 | X > 2\},$$

and $P\{X = 3|X > 2\}$ is the probability that the boy draws two chips and then his third chip is a duplicate of one of the first two draws. We are assuming that X is not less than or equal to two i.e. the first two draws result in unseen numbers. Thus $P\{X = 3|X > 2\} = \frac{2}{n}$. In the same way

$$P\{X = i|X > i-1\} = \frac{i-1}{n} \quad \text{for} \quad 1 \leq i \leq n+1.$$

Therefore for $1 \leq i \leq n+1$ we have

$$\begin{aligned} F(i) &= \sum_{k=1}^{i} \frac{k-1}{n} = \frac{1}{n}\sum_{k=1}^{i}(k-1) = \frac{1}{n}\left[\sum_{k=1}^{i} k - i\right] \\ &= \frac{1}{n}\left[\frac{i(i+1)}{2} - i\right] = \frac{i(i-1)}{2n}. \end{aligned}$$

Problem 35 (double replacement of balls in our urn)

Let X be the selection number for the first selection of a blue ball. Recall that $P\{X > i\} = 1 - P\{X \leq i\}$. Now we can compute $P\{X = i\}$ by induction. First we see that

$$P\{X = 1\} = \frac{1}{2}.$$

Next we have

$$P\{X = 2\} = P\{X = 2|B_1 = R\}P\{B_1 = R\} = \left(\frac{1}{3}\right)\left(\frac{1}{2}\right) = \frac{1}{2 \cdot 3},$$

where we have conditioned on the fact that the first ball must be red. Continuing we see that

$$\begin{aligned} P\{X = 3\} &= P\{X = 3|B_1, B_2 = R, R\}P\{RR\} = \left(\frac{1}{4}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right) = \frac{1}{3 \cdot 4} \\ P\{X = 4\} &= P\{X = 4|B_1, B_2, B_3 = R, R, R\}P\{RRR\} = \left(\frac{1}{5}\right)\left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{3}{4}\right) = \frac{1}{4 \cdot 5}. \end{aligned}$$

By induction we conclude that

$$P\{X = i\} = \frac{1}{i(i+1)}\,.$$

So that

$$P\{X \le i\} = \sum_{k=1}^{i} P\{X = k\} = \sum_{k=1}^{i} \frac{1}{k(k+1)}\,.$$

Now using partial fractions we see that the fraction we are summing is given by

$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}\,.$$

So that we see that our sum above is of the "telescoping" type and simplifies as

$$\begin{aligned}
\sum_{k=1}^{i} \frac{1}{k(k+1)} &= \sum_{k=1}^{i} \left(\frac{1}{k} - \frac{1}{k+1}\right) \\
&= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{i-1} - \frac{1}{i}\right) + \left(\frac{1}{i} - \frac{1}{i+1}\right) \\
&= 1 - \frac{1}{i+1}\,.
\end{aligned}$$

Thus $P\{X > i\} = 1 - P\{X \le i\} = \frac{1}{i+1}$ for $i \ge 1$.

Part (b): From our expression $P\{X \le i\} = 1 - \frac{1}{i+1}$, we have that $P\{X \le \infty\} = 1 - 0 = 1$, thus with probability one X will be finite i.e. the blue ball is eventually chosen.

Part (c): Using the definition of expectation we have

$$E[X] = \sum_{i=1}^{\infty} iP\{X = i\} = \sum_{i=1}^{\infty} \frac{1}{i+1} = +\infty\,,$$

thus the blue ball is eventually chosen but on average it is not chosen until very late.

Chapter 4: Self-Test Problems and Exercises

Chapter 5 (Continuous Random Variables)

Chapter 5: Problems

Problem 1 (normalizing a continuous random variable)

Part (a): The integral of the f must evaluate to one, which requires

$$\begin{aligned}\int_{-1}^{1} c(1-x^2)dx &= 2c\int_0^1 (1-x^2)dx \\ &= 2c\left(x - \frac{x^3}{3}\right|_0^1 = 2c\left(1-\frac{1}{3}\right) = \frac{4c}{3}\,.\end{aligned}$$

For this to equal one, we must have $c = \frac{3}{4}$.

Part (b): The cumulative distribution is given by

$$\begin{aligned}F(x) &= \int_{-1}^{x} \frac{3}{4}(1-\xi^2)d\xi \\ &= \frac{3}{4}\left(\xi - \frac{\xi^3}{3}\right|_{-1}^{x} \\ &= \frac{3}{4}\left(x - \frac{x^3}{3}\right) + \frac{1}{2} \quad \text{for} \quad -1 \le x \le 1\,.\end{aligned}$$

Problem 2 (how long can our system function?)

We must first evaluate the constant in our distribution function. Specifically to be a probability density we must have

$$\int_0^\infty cxe^{-x/2}dx = 1\,.$$

Integrating by parts we find that

$$\begin{aligned}\int_0^\infty cxe^{-x/2}dx &= c\left[\frac{xe^{-x/2}}{(-1/2)}\right|_0^\infty - \frac{1}{(-1/2)}\int_0^\infty e^{-x/2}dx\right] \\ &= c\left[2\int_0^\infty e^{-x/2}dx\right] \\ &= 2c\,\frac{e^{-x/2}}{(-1/2)}\bigg|_0^\infty = -4c(0-1) = 4c\,.\end{aligned}$$

So for this to equal one we must have $c = 1/4$. Then the probability that our system last at least five months is given by

$$\begin{aligned}
\int_5^\infty \frac{1}{4} x e^{-x/2} dx &= \frac{1}{4}\left[\left.\frac{xe^{-x/2}}{(-1/2)}\right|_5^\infty - \int_5^\infty \frac{e^{-x/2}}{(-1/2)} dx \right] \\
&= \frac{1}{4}\left[0 + 10e^{-5/2} + 2\int_5^\infty e^{-x/2} dx \right] \\
&= \cdots = \frac{7}{2} e^{-5/2} .
\end{aligned}$$

Problem 3 (possible density functions)

Even with a value of C specified a problem with this function f is that it is negative for some values of x. Specifically f will be *zero* when $x(2 - x^2) = 0$, which happens when $x = 0$ or $x = \pm\sqrt{2} = \pm 1.4142$. With these zeros found we see that if x is less than $\sqrt{2}$ then $x(2 - x^2)$ is positive, however if x is greater than $\sqrt{2}$ (but still less than $5/2$) the expression $x(2-x^2)$ is negative. Thus whatever the sign of c, $f(x)$ will be negative for some region of the interval. Since f cannot be negative this functional form cannot be a probability density function.

For the second function this f is zero when $x(2 - x) = 0$, which happens when $x = 0$ and $x = 2$. Since $2 < 5/2 = 2.5$. This f will also change signs regardless of the constant C as x crosses the value 2. Since f takes on both positive and negative signed values it can't be a distribution function.

Problem 4 (the lifetime of electronics)

Part (a): The requested probability is given by

$$P\{X > 20\} = \int_{20}^\infty \frac{10}{x^2} dx = \frac{1}{2} .$$

Part (b): The requested cumulative distribution function is given by

$$F(x) = \int_{10}^\infty \frac{10}{\xi^2} d\xi = \left.\frac{10\xi^{-1}}{(-1)}\right|_{10}^x = 1 - \frac{10}{x} \quad \text{for} \quad 10 \le x .$$

Part (c): To function for at least fifteen hours will happen with probability $1 - F(15) = 1 - (1 - \frac{10}{15}) = \frac{2}{3}$. To have three of six such devices function for at least fifteen hours is given by sums of binomial probability density functions. Specifically we have this probability given by

$$\sum_{k=3}^{6} \binom{6}{k} \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^{6-k} ,$$

which we recognized as the "complement" of the binomial cumulative distribution function. To evaluate this we can use the Matlab command `binocdf(2,6,2/3)`. See the Matlab file `chap_5_prob_4.m` for these calculations and we find that the above equals 0.8999. In performing this analysis we are assuming independence of the devices.

Problem 11 (picking a point on a line)

An interpretation of this statement is that a point is picked randomly on a line segment of length L would be that the point "X" is selected from a uniform distribution over the interval $[0, L]$. Then the question asks us to find

$$P\left\{\frac{\min(X, L-X)}{\max(X, L-X)} < \frac{1}{4}\right\}.$$

This probability can be evaluated by integrating over the appropriate region. Formally we have the above equal to

$$\int_E p(x)dx$$

where $p(x)$ is the uniform probability density for our problem, i.e. $\frac{1}{L}$ and the set "E" is $x \in [0, L]$ and satisfying the inequality above, i.e.

$$\min(x, L-x) \leq \frac{1}{4}\max(x, L-x).$$

Plotting the functions $\max(x, L-x)$, and $\min(x, L-x)$ in Figure 2, we see that the regions of X where we should compute the integral above are restricted to the two ends of the segment. Specifically, the integral above becomes,

$$\int_0^{l_1} p(x)dx + \int_{l_2}^{L} p(x)dx.$$

since the region $\min(x, L-x) < \frac{1}{4}\max(x, L-x)$ in satisfied in the region $[0, l_1]$ and $[l_2, L]$ only. Here l_1 is the solution to

$$\min(x, L-x) = \frac{1}{4}\max(x, L-x) \quad \text{when} \quad x < L-x,$$

i.e. we need to solve

$$x = \frac{1}{4}(L-x)$$

which has as its solution $x = \frac{L}{5}$. For l_2 we must solve

$$\min(x, L-x) = \frac{1}{4}\max(x, L-x) \quad \text{when} \quad L-x < x,$$

i.e. we need to solve

$$L - x = \frac{1}{4}x,$$

which has as its solution $x = \frac{4}{5}L$. With these two limits we have for our probability

$$\int_0^{\frac{L}{5}} \frac{1}{L}dx + \int_{\frac{4}{5}L}^{L} \frac{1}{L}dx = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}.$$

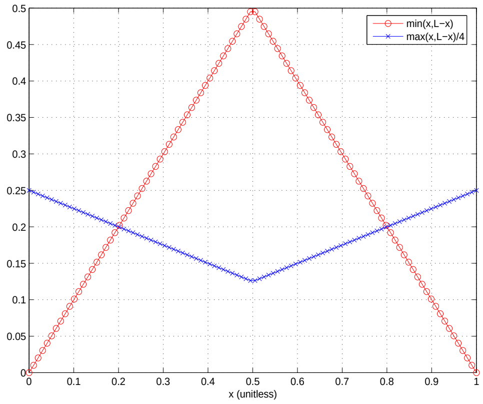

Figure 2: A graphical view of the region of x's over which the integral for this problem should be computed.

Problem 15 (some normal probabilities)

In general to solve this problem we will convert each probability to a corresponding one involving unit normal random variables and then compute this second probability using the cumulative distribution function $\Phi(\cdot)$.

Part (a): We find

$$\begin{aligned} P\{X > 5\} &= P\{\frac{X-10}{6} > \frac{5-10}{6}\} \\ &= P\{Z > -0.833\} = 1 - P\{Z < -0.833\} = \Phi(-0.833)\,. \end{aligned}$$

Part (b): $P\{4 < X < 16\} = P\{\frac{4-10}{6} < \frac{X-10}{6} < \frac{16-10}{6}\} = P\{-1 < Z < +1\} = \Phi(1) - \Phi(-1)$.

Part (c): $P\{X < 8\} = P\{\frac{X-10}{6} < \frac{8-10}{6}\} = P\{Z < -0.333\} = \Phi(-0.333)$.

Part (d): $P\{X < 20\} = \Phi(\frac{20-10}{6}) = \Phi(1.66)$, following the same steps as in Part (a).

Part (e): $P\{X > 16\} = P\{\frac{X-10}{6} > \frac{16-10}{6}\} = P\{Z > 1\} = 1 - P\{Z < 1\} = 1 - \Phi(1)$.

Problem 16 (annual rainfall)

The probability that we have over 50 inches of rain in one year is

$$P\{X > 50\} = P\{\frac{X-40}{4} > \frac{50-40}{4}\} = P\{Z > 2.5\} = 1 - P\{Z < 2.5\} = 1 - \Phi(2.5)\,.$$

If we assume that the probabilities each year are uncorrelated and also assuming that the probability this event happens at year n is a geometric random variable, the probability that it takes over ten years before this event happens is one minus the probability that it takes $1, 2, 3, \cdots, 10$ years for this event to occur. That is if E is the event that over ten years pass before this event happens, than

$$P(E) = 1 - \sum_{i=1}^{10} p(1-p)^i .$$

Since $\sum_{i=1}^{N} a^i = \frac{a - a^{N+1}}{1-a}$ we have that

$$P(E) = 1 - p\left(\frac{(1-p) - (1-p)^{11}}{1 - (1-p)}\right) = 1 - ((1-p) - (1-p)^{11}) = p + (1-p)^{11} .$$

When we put $p = 1 - \Phi(2.5)$ we get the desired probability.

Problem 17 (the expected number of points scored)

We desire to calculate $E[P(D)]$, where $P(D)$ is the points scored when the distance to the target is D. This becomes

$$\begin{aligned}
E[P(D)] &= \int_0^{10} P(D) f(D) dD \\
&= \frac{1}{10} \int_0^{10} P(D) dD \\
&= \frac{1}{10} \left(\int_0^1 10 dD + \int_1^3 5 dD + \int_3^5 3 dD + \int_5^{10} 0 dD \right) \\
&= \frac{1}{10} (10 + 5(2) + 3(2)) = \frac{26}{10} = 2.6 .
\end{aligned}$$

Problem 18 (variable limits on a normal random variable)

Since X is a normal random variable we can evaluate the given probability $P\{X > 9\}$ as

$$\begin{aligned}
P\{X > 9\} &= P\{\frac{X-5}{\sigma} > \frac{9-5}{\sigma}\} \\
&= P\left\{Z > \frac{4}{\sigma}\right\} \\
&= 1 - P\left\{Z < \frac{4}{\sigma}\right\} \\
&= 1 - \Phi(\frac{4}{\sigma}) = 0.2 ,
\end{aligned}$$

so solving for $\Phi(4/\sigma)$ we have that $\Phi(4/\sigma) = 0.8$, which can be inverted by using the Matlab command `norminv` and we calculate that

$$\frac{4}{\sigma} = \Phi^{-1}(0.8) = 0.8416 .$$

which then implies that $\sigma = 4.7527$, so $\text{Var}(X) = \sigma^2 \approx 22.58$.

Problem 19 (more limits on normal random variables)

Since X is a normal random variable we can evaluate the given probability $P\{X > c\}$ as

$$\begin{aligned} P\{X > c\} &= P\{\frac{X-12}{2} > \frac{c-12}{2}\} \\ &= P\left\{Z > \frac{c-12}{2}\right\} \\ &= 1 - P\left\{Z < \frac{c-12}{2}\right\} \\ &= 1 - \Phi(\frac{c-12}{2}) = 0.1\,, \end{aligned}$$

so solving for $\Phi(\frac{c-12}{2}$ we have that $\Phi(\frac{c-12}{2}) = 0.9$, which can be inverted by using the Matlab command `norminv` and we calculate that

$$\frac{c-12}{2} = \Phi^{-1}(0.9) = 1.28\,,$$

which then implies that $c = 14.56$.

Problem 20 (the expected number of people in favor of a proposition)

Now the number of people who favor the proposed rise in taxes is a binomial random variable with parameters $(n, p) = (100, 0.65)$. Using the normal approximation to the binomial, we have a normal with a mean of $np = 100(0.65) = 65$, and a variance of $\sigma^2 = np(1-p) = 100(0.65)(0.35) = 22.75$, so the probabilities desired are given as

Part (a):

$$\begin{aligned} P\{N \geq 50\} &= P\{N > 49.5\} \\ &= P\left\{\frac{N-65}{\sqrt{22.75}} > \frac{49.5-65}{4.76}\right\} \\ &= P\{Z > -3.249\} \\ &= 1 - \Phi(-3.249) = 1 - (1 - \Phi(3.249)) = \Phi(3.249)\,. \end{aligned}$$

Where in the first equality we have used the "continuity approximation". Using the Matlab command `normcdf(x)` to evaluate the function $\Phi(x)$ we have the above equal to ≈ 0.9994.

Part (b):

$$\begin{aligned} P\{60 \leq N \leq 70\} &= P\{59.5 < N < 70.5\} \\ &= P\left\{\frac{59.5-65}{\sqrt{22.75}} < Z < \frac{70.5-65}{\sqrt{22.75}}\right\} \\ &= P\{-1.155 < Z < 1.155\} \\ &= \Phi(1.155) - \Phi(-1.155) \approx 0.7519\,. \end{aligned}$$

Part (c):

$$
\begin{aligned}
P\{N < 75\} &= P\{N < 74.5\} \\
&= P\left\{Z < \frac{74.5 - 65}{4.76}\right\} \\
&= P\{Z < 1.99\} \\
&= \Phi(1.99) \approx 0.9767\,.
\end{aligned}
$$

Problem 21 (percentage of men with height greater than six feet two inches)

We desire to compute $P\{X > 6 \cdot 12 + 2\}$, where X is the random variable expressing height (measured in inches) of a 25-year old man. This probability can be computed by converting to the standard normal in the usual way. We have

$$
\begin{aligned}
P\{X > 6 \cdot 12 + 2\} &= P\left\{\frac{X - 71}{\sqrt{6.25}} > \frac{3}{\sqrt{6.25}}\right\} \\
&= P\left\{Z > \frac{3}{\sqrt{6.25}}\right\} \\
&= 1 - P\left\{Z < \frac{3}{\sqrt{6.25}}\right\} \\
&= 1 - \Phi(1.2) \approx 0.1151\,.
\end{aligned}
$$

For the second part of this problem we are looking for

$$
P\{X > 6 \cdot 12 + 5 | X > 6 \cdot 12\}\,.
$$

Again this can be computed by converting to a standard normal, after first considering the joint density. We have

$$
\begin{aligned}
P\{X > 6 \cdot 12 + 5 | X > 6 \cdot 12\} &= \frac{P\{X > 77, X > 72\}}{P\{X > 72\}} \\
&= \frac{P\{X > 77\}}{P\{X > 72\}} \\
&= \frac{1 - P\left\{Z < \frac{6}{\sqrt{6.25}}\right\}}{1 - P\left\{Z < \frac{1}{\sqrt{6.25}}\right\}} \\
&= \frac{1 - \Phi(\frac{6}{\sqrt{6.25}})}{1 - \Phi(\frac{1}{\sqrt{6.25}})} \approx 0.0238\,.
\end{aligned}
$$

Some of the calculations for this problem can be found in the file `chap_5_prob_21.m`.

Problem 22 (number of defective products)

Part (a): Lets calculate the percentage that are *acceptable* if we let the variable X be the width of our normally distributed slot this percentage is given by

$$P\{0.895 < X < 0.905\} = P\{X < 0.905\} - P\{X < 0.895\}.$$

Each of these individual cumulative probabilities can be calculated by transforming to the standard normal, in the usual way. We have that the above is equal to (since the population mean is 0.9 and the population standard deviation is 0.003)

$$\begin{aligned} & P\left\{\frac{X-0.9}{0.003} < \frac{0.905-0.9}{0.003}\right\} - P\left\{\frac{X-0.9}{0.003} < \frac{0.895-0.9}{0.003}\right\} \\ = & \ \Phi\left(1.667\right) - \Phi\left(-1.667\right) = 0.904 . \end{aligned}$$

So that the probability (or percentage) of defective forgings is one minus this number (times 100 to convert to percentages). This is $0.095 \times 100 = 9.5$.

Part (b): This question is asking to find the value of σ such that

$$P\{0.895 < X < 0.905\} = \frac{99}{100}.$$

Since these limits on X are symmetric about $X = 0.9$ we can simplify this probability by using

$$P\{0.895 < X < 0.905\} = 1 - 2P\{X < 0.895\} = 1 - 2P\left\{\frac{X-0.9}{\sigma} < \frac{0.905-0.9}{\sigma}\right\}$$

We thus have to solve for σ in

$$1 - 2\Phi(\frac{-0.005}{\sigma}) = 0.99$$

or inverting the Φ function and solving for σ we have

$$\sigma = \frac{-0.005}{\Phi^{-1}(0.005)}.$$

Using the Matlab command `norminv` to evaluate the above we have $\sigma = 0.0019$. See the Matlab file `chap_5_prob_22.m` for these calculations.

Problem 23 (probabilities on the number of 5's to appear)

The probability that one six occurs is $p = 1/6$ so the total number of sixes rolled is a binomial random variable. We can approximate this density as a Gaussian with a mean given by $np = \frac{1000}{6} \approx 166.6$ and a variance of $\sigma^2 = np(1-p) = 138.8$. Then the desired probabilities are

$$\begin{aligned} P\{150 \le N \le 200\} &= P\{149.5 < N < 200.5\} \\ &= P\{-1.45 < Z < 2.877\} \\ &= \Phi(2.87) - \Phi(-1.45) \approx 0.9253 . \end{aligned}$$

If we are told that a six appears two hundred times then the probability that a five will appear on the other rolls is 1/5 and it must appear on one of the $1000-200=800$ other rolls. Thus we can approximate the binomial random variable (with parameters $(n,p)=(800,1/5)$) with a normal with mean $np=\frac{800}{5}=160$ and variance $\sigma^2=np(1-p)=128$. So the requested probability is

$$\begin{aligned} P\{N<500\} &= P\{N<149.5\} \\ &= P\left\{Z<\frac{149.5-160}{\sqrt{128}}\right\} \\ &= P\{Z<-0.928\} \\ &= \Phi(-0.928)\approx 0.1767\,. \end{aligned}$$

Problem 24 (probability of enough long living batteries)

If each chips lifetime is denoted by the random variable X (assumed Gaussian with the given mean and variance), then each chip will have a lifetime less than $1.8\,10^6$ hours with probability given by

$$\begin{aligned} P\{X<1.8\,10^6\} &= P\left\{\frac{X-1.4\,10^6}{3\,10^5}<\frac{(1.8-1.4)\,10^6}{3\,10^5}\right\} \\ &= P\left\{Z<\frac{4}{3}\right\}=\Phi(4/3)\approx 0.9088\,. \end{aligned}$$

With this probability, the number N, in a batch of 100 that will have a lifetime less than $1.8\,10^6$ is a binomial random variable with parameters $(n,p)=(100,0.9088)$. Therefore, the probability that a batch will contain at least 20 is given by

$$P\{N\geq 20\}=\sum_{n=20}^{100}\left(\begin{array}{c}100\\ n\end{array}\right)(0.908)^n(1-0.908)^{100-n}\,.$$

Rather than evaluate this exactly we can approximate this binomial random variable N with a Gaussian random variable with a mean given by $\mu=np=100(0.908)=90.87$, and a variance given by $\sigma^2=np(1-p)=8.28$ (equivalently $\sigma=2.87$). Then the probability that a given batch of 100 has at least 20 that have lifetime less than $1.8\,10^6$ hours is given by

$$\begin{aligned} P\{N\geq 20\} &= P\{N\geq 19.5\} \\ &= P\left\{\frac{N-90.87}{2.87}\geq\frac{19.5-90.87}{2.87}\right\} \\ &\approx P\{Z\geq -24.9\} \\ &= 1-P\{Z\leq -24.9\} \\ &= 1-\Phi(-24.9)\approx 1\,. \end{aligned}$$

Where in the first line above we have used the continuity correction required when we approximate a discrete density by a continuous one, and in the third line above we use our Gaussian approximation to the binomial distribution.

Problem 25 (the probability of enough acceptable items)

The number N of acceptable items is a binomial random variable so we can approximate it with a Gaussian with mean $\mu = pn = 150(0.95) = 142.5$, and a variance of $\sigma^2 = np(1-p) = 7.125$. From the variance we have a standard deviation of $\sigma \approx 2.669$. Thus the desired probability is given by

$$\begin{aligned}
P\{N \geq 140\} &= P\{N \geq 139.5\} \\
&= P\left\{\frac{N - 142.5}{2.669} \geq \frac{139.5 - 142.5}{2.669}\right\} \\
&\approx P\{Z \geq -1.127\} \\
&= 1 - P\{Z \leq -1.127\} \\
&= 1 - \Phi(-1.127) \approx 0.8701\,.
\end{aligned}$$

Where in the first line above we have used the continuity correction required when we approximate a discrete density by a continuous one, and in the third line above we use our Gaussian approximation to the binomial distribution. We note that we solved this problem in terms of the number of items that are *acceptable*. An equivalent formulation could easily be done in terms of the number that are *unacceptable* by using the complementary probability $q \equiv 1 - p = 1 - 0.95 = 0.05$.

Problem 26 (calculating the probability of error)

Let N be the random variable that represents the number of heads that result when we flip our coin 1000 times. Then N is distributed as binomial random variable with a probability of success p that depends on whether we are considering the biased or unbiased (fair) coin. It the coin is actually *fair* we will make an error in our assessment of its type if N is greater than 525 according to the statement of this problem. Thus the probability that we reach a false conclusion is given by

$$P\{N \geq 525\}\,.$$

To compute this probability we will use the normal approximation to the binomial distribution. In this case the normal to use to approximate this binomial distribution has a mean given by $\mu = np = 1000(0.5) = 500$ and a variance given by $\sigma^2 = np(1-p) = 250$ since we know we are looking at the fair coin where $p = 0.5$. To evaluate this probability we have

$$\begin{aligned}
P\{N \geq 525\} &= P\{N \geq 524.5\} \\
&= P\left\{\frac{N - 500}{\sqrt{250}} \geq \frac{524.5 - 500}{\sqrt{250}}\right\} \\
&\approx P\{Z \geq 1.54\} \\
&= 1 - P\{Z \leq 1.54\} \\
&= 1 - \Phi(1.54) \approx 0.0606\,.
\end{aligned}$$

Where in the first line above we have used the continuity correction required when we approximate a discrete density by a continuous one, and in the third line above we use our Gaussian approximation to the binomial distribution. In the case where the coin is actually biased our probability of obtaining a head becomes $p = 0.55$ and we will reach a false conclusion in this case when

$$P\{N < 525\}\,.$$

To compute this probability we will use the normal approximation to the binomial distribution. In this case the normal to use to approximate this binomial distribution has a mean given by $\mu = np = 1000(0.55) = 550$ and a variance given by $\sigma^2 = np(1-p) = 247.5$. To evaluate this probability we have

$$\begin{aligned} P\{N < 525\} &= P\{N < 524.5\} \\ &= P\left\{\frac{N-550}{\sqrt{247.5}} < \frac{524.5-550}{\sqrt{247.5}}\right\} \\ &\approx P\{Z < -1.62\} \\ &= \Phi(-1.62) \approx 0.0525\,. \end{aligned}$$

Where in the first line above we have used the continuity correction required when we approximate a discrete density by a continuous one, and in the third line above we use our Gaussian approximation to the binomial distribution.

Problem 27 (fair coins)

Now $P\{N = 5800\} = P\{5799.5 \le N \le 5800.5\}$, by the continuity approximation. The second probability can be approximated by a normal with mean $np = 10^4\,(0.5) = 5000$ and variance $np(1-p) = 2500$, so that the above becomes

$$\begin{aligned} P\{\frac{5799.5-5000}{\sqrt{2500}} \le Z \le \frac{5800.5-5000}{\sqrt{2500}}\} &= \Phi\left(\frac{5800.5-5000}{\sqrt{2500}}\right) - \Phi\left(\frac{5799.5-5000}{\sqrt{2500}}\right) \\ &= 1 - 1 = 0\,, \end{aligned}$$

so there is effectively no probability this could have happened by chance.

Problem 28 (the number of left handed students)

The number of students that are left handed (denoted by N) is a Binomial random variable with parameters $(n,p) = (200, 0.12)$. From the normal approximation to the binomial we can approximate this distribution with a Gaussian with mean $\mu = pn = 200(0.12) = 24$, and a variance of $\sigma^2 = np(1-p) = 21.120$. From the variance we have a standard deviation of $\sigma \approx 4.59$. Thus the desired probability is given by

$$\begin{aligned} P\{N > 20\} &= P\{N > 19.5\} \\ &= P\left\{\frac{N-24}{4.59} > \frac{19.5-24}{4.59}\right\} \\ &\approx P\{Z > -0.9792\} \\ &= 1 - P\{Z \le -0.9792\} \\ &= 1 - \Phi(-0.9792) \approx 0.8363\,. \end{aligned}$$

Where in the second line above we have used the continuity correction that improves our accuracy when we approximate a discrete density by a continuous one, and in the third line above we use our Gaussian approximation to the binomial distribution. These calculations can be find in the file `chap_5_prob_28.m`.

Problem 29 (a simple model of stock movement)

If we count each time the stock rises in value as a "success", we see that the movement of the stock for one timestep is a Bernoulli random variable with parameter p. So after n timesteps the number of rises is a binomial random variable with parameters (n, p). The price of the security after n timesteps where we have k "successes" will then be given by $su^k d^{n-k}$. The probability we are looking for then is given by

$$
\begin{aligned}
P\{su^k d^{n-k} \geq 1.3\, s\} &= P\{u^k d^{n-k} \geq 1.3\} \\
&= P\left\{\left(\frac{u}{d}\right)^k \geq \frac{1.3}{d^n}\right\} \\
&= P\left\{k \geq \frac{\ln(\frac{1.3}{d^n})}{\ln(\frac{u}{d})}\right\} \\
&= P\left\{k \geq \frac{\ln(1.3) - n\ln(d)}{\ln(u) - \ln(d)}\right\} .
\end{aligned}
$$

Using the numbers given in the problem i.e. $d = 0.99$ $u = 1.012$, and $n = 1000$, we have that

$$
\frac{\ln(1.3) - n\ln(d)}{\ln(u) - \ln(d)} \approx 469.2 .
$$

To approximate the above probability we can use the Gaussian approximation to the binomial distribution, which would have a mean given by $np = 0.52(1000) = 520$ and a variance given by $np(1-p) = 249.6$, so using this approximation the above probability then becomes

$$
\begin{aligned}
P\{k \geq 469.2\} &= P\{k \geq 470\} \\
&= P\{k > 469.5\} \\
&= P\{Z > \frac{469.5 - 520}{15.7}\} \\
&= P\{Z > -3.21\} \\
&= 1 - P\{Z < -3.21\} \\
&= 1 - \Phi(-3.23) \approx 0.9994 .
\end{aligned}
$$

Problem 30 (priors on the type of region)

Let E be the event that we make an error in our classification of the given pixel. Then we can make an error in two symmetric ways. The first is that we classify the pixel as black when it should be classified as white. The second is where we classify the pixel as white when it should be black. Thus we can compute $P(E)$ by conditioning on the true type of the pixel i.e. whether it is B (for black) or W (for white). We have

$$P(E) = P(E|B)P(B) + P(E|W)P(W) .$$

Since we are told that the prior probability that the pixel is black is given by α, the prior probability that the pixel is W is then given by $1 - \alpha$ and the above becomes

$$P(E) = P(E|B)\alpha + P(E|W)(1 - \alpha) .$$

The problem then asks for the value of α such that the error in making each type of error is the same, we desire to pick α such that

$$P(E|B)\alpha = P(E|W)(1 - \alpha) ,$$

or upon solving for α we find that

$$\alpha = \frac{P(E|W)}{P(E|W) + P(E|B)} .$$

We now need to evaluate $P(E|W)$ and $P(E|B)$. Now $P(E|W)$ is the probability that we classify this pixel as *black* given that it is white. If we classify the pixel with a value of 5 as black, then all points with pixel value greater than 5 would also be classified as black and $P(E|W)$ is then given by

$$P(E|W) = \int_5^\infty \mathcal{N}(x; 4, 4)dx = \int_{(5-4)/2}^\infty \mathcal{N}(z; 0, 1)dz = 1 - \Phi(1/2) = 0.3085 .$$

Where $\mathcal{N}(x; \mu, \sigma^2)$ is an expression for the normal probability density function with mean μ and variance σ^2. In the same way we have that

$$P(E|B) = \int_{-\infty}^5 \mathcal{N}(x; 6, 9)dx = \int_{-\infty}^{(5-6)/3} \mathcal{N}(z; 0, 1)dz = \Phi(-1/3) = 0.3694 .$$

Thus with these two expressions α becomes

$$\alpha = \frac{1 - \Phi(1/2)}{(1 - \Phi(1/2)) + \Phi(-1/3)} = 0.4551 .$$

Problem 31 (the optimal location of a fire station)

Part (a): If x (the location of the fire) is uniformly distributed in $[0, A)$ then we would like to select a (the location of the fire station) such that

$$F(a) \equiv E[\,|X - a|\,] ,$$

is a minimum. We will compute this by breaking the integral involved in the definition of the expectation into regions where $x - a$ is negative and positive. We find that

$$\begin{aligned}
E[\,|X-a|\,] &= \int_0^A |x-a|\frac{1}{A}dx \\
&= -\frac{1}{A}\int_0^a (x-a)dx + \frac{1}{A}\int_a^A (x-a)dx \\
&= -\frac{1}{A}\left.\frac{(x-a)^2}{2}\right|_0^a + \frac{1}{A}\left.\frac{(x-a)^2}{2}\right|_a^A \\
&= -\frac{1}{A}\left(0 - \frac{a^2}{2}\right) + \frac{1}{A}\left(\frac{(A-a)^2}{2} - 0\right) \\
&= \frac{a^2}{2A} + \frac{(A-a)^2}{2A}\,.
\end{aligned}$$

To find the a that minimizes this we compute $F'(a)$ and set this equal to zero. Taking the derivative and setting this equal to zero we find that

$$F'(a) = \frac{a}{A} + \frac{2(A-a)(-1)}{2A} = 0\,.$$

Which gives a solution a^* given by $a^* = \frac{A}{2}$. A second derivative of our function F shows that $F''(a) = \frac{2}{A} > 0$ showing that the point $a^* = A/2$ is indeed a minimum.

Part (b): The problem formulation is the same as in part (a) but since the distribution of the location of fires is now an exponential we now want to minimize

$$F(a) \equiv E[\,|X-a|\,] = \int_0^\infty |x-a|\lambda e^{-\lambda x}dx\,.$$

We will compute this by breaking the integral involved in the definition of the expectation into regions where $x - a$ is negative and positive. We find that

$$\begin{aligned}
E[\,|X-a|\,] &= \int_0^\infty |x-a|\lambda e^{-\lambda x}dx \\
&= -\int_0^a (x-a)\lambda e^{-\lambda x}dx + \int_a^\infty (x-a)\lambda e^{-\lambda x}dx \\
&= -\lambda\left(\left.\frac{(x-a)}{-\lambda}e^{-\lambda x}\right|_0^a + \frac{1}{\lambda}\int_0^a e^{-\lambda x}dx\right) \\
&+ \lambda\left(\left.\frac{(x-a)}{-\lambda}e^{-\lambda x}\right|_a^\infty + \frac{1}{\lambda}\int_a^\infty e^{-\lambda x}dx\right) \\
&= -\lambda\left(\frac{-a}{\lambda} - \left.\frac{1}{\lambda^2}e^{-\lambda x}\right|_0^a\right) + \lambda\left(0 - \left.\frac{1}{\lambda^2}e^{-\lambda x}\right|_a^\infty\right) \\
&= a + \frac{1}{\lambda}(e^{-\lambda a} - 1) - \frac{1}{\lambda}(-e^{-\lambda a}) \\
&= a + \frac{-1 + 2e^{-\lambda a}}{\lambda}\,.
\end{aligned}$$

To find the a that minimizes this we compute $F'(a)$ and set this equal to zero. Taking the derivative we find that

$$F'(a) = 1 - 2e^{-\lambda a} = 0\,.$$

Which gives a solution a^* given by $a^* = \frac{\ln(2)}{\lambda}$. A second derivative of our function F shows that $F''(a) = 2\lambda e^{-\lambda a} > 0$ showing that the point $a^* = \frac{\ln(2)}{\lambda}$ is indeed a minimum.

Problem 32 (probability of repair times)

Part (a): We desire to compute $P\{T > 2\}$ which is given by

$$P\{T > 2\} = \int_2^\infty \frac{1}{2} e^{-1/2\,t} dt\,.$$

To evaluate this let $v = \frac{t}{2}$, giving $dv = \frac{dt}{2}$, from which the above becomes

$$P\{T > 2\} = \int_1^\infty e^{-v} dv = -e^{-v}\big|_1^\infty = e^{-1}\,.$$

Part (b): The probability we are after is given by $P\{T > 10|T > 9\}$ which equals $P\{T > 10 - 9\} = P\{T > 1\}$ by the memoryless property of the exponential distribution. This is given by

$$P\{T > 1\} = 1 - P\{T < 1\} = 1 - (1 - e^{-1/2}) = e^{-1/2}\,.$$

Problem 33 (a functioning radio?)

Because of the memoryless property of the exponential distribution the fact that the ratio is used is irrelevant. The probability requested is then

$$P\{T > 8 + t|T > t\} = P\{T > 8\} = 1 - P\{T < 8\} = 1 - (1 - e^{-\frac{1}{8}(8)}) = e^{-1}\,.$$

Problem 34 (getting additional miles from a car)

Since the exponential random variable has no memory, the fact that the car has been driven 100,000 miles makes no difference. The probability we are looking for is

$$P\{T > 20000\} = 1 - P\{T < 20000\} = 1 - (1 - e^{-\frac{1}{20}(20)}) = e^{-1}\,.$$

If the lifetime distribution is not exponential but is uniform over $(0, 40)$ then the desired probability is given by

$$P\{T_{\text{thous}} > 30|T_{\text{thous}} > 20\} = \frac{P\{T > 30\}}{P\{T > 10\}} = \frac{(1/4)}{(3/4)} = \frac{1}{3}\,.$$

Here T_{thous} is the distance driven in thousands of miles.

Problem 35 (lung cancer hazard rates)

Given a hazard rate of $\lambda(t)$ then from Example 5f we see that if E is the event that an "A" year old reaches age B is given by

$$P(E) = \exp\{-\int_A^B \lambda(t)dt\}\,,$$

so for this problem since our person is age forty we want

$$\exp\{-\int_{40}^{50} \lambda(t)dt\}\,.$$

First lets calculate

$$\int_{40}^{B} \lambda(t)dt = \int_{40}^{B} (0.027 + 0.00025(t-40)^2)dt = 0.027(B-40) + 0.00025\frac{(B-40)^3}{3}\,.$$

When $B = 50$ this number is 0.353, so our survival probability is $\exp(-0.353) = 0.702$. While if $B = 60$ this number is 1.2 so our survival probability is 0.299.

Problem 36 (surviving with a given hazard rate)

From Example 5f the probability that our object survives to age B (from zero) is given by

$$\exp\{-\int_0^B \lambda(t)dt\} = \exp\{-\int_0^B t^3 dt\} = \exp\{-\left.\frac{t^4}{4}\right|_0^B\} = \exp\{-\frac{B^4}{4}\}\,.$$

Part (a): The desired probability is when $B = 2$, which when put in the above gives 0.0183.

Part (b): The required probability is

$$\exp\{-\int_{0.4}^{1.4} \lambda(t)dt\} = \exp\{-\frac{1}{4}\left((1.4)^4 - (0.4)^4\right)\} = 0.3851\,.$$

Part (c): This probability is

$$\exp\{-\int_1^2 \lambda(t)dt\} = \exp\{-\frac{1}{4}\left(2^4 - 1\right)\} = 0.0235\,.$$

Problem 37 (uniform probabilities)

Part (a): The desired probability is

$$P\{|x| > \frac{1}{2}\} = \int_{|x|>\frac{1}{2}} f(x)dx = \frac{1}{2}\int_{|x|>\frac{1}{2}} dx = \frac{2}{2}\int_{1/2}^{1} dx = 1 - \frac{1}{2} = \frac{1}{2}\,.$$

Part (b): Define $Y = |x|$ and consider the distribution for Y i.e.

$$\begin{aligned} F_Y(a) &= P\{Y \leq a\} = P\{|x| \leq a\} = 2P\{0 \leq x \leq a\} \\ &= 2\int_0^a f(x)dx = 2\int_0^a \frac{1}{2}dx = a\,, \end{aligned}$$

for $0 \leq a \leq 1$ and is zero elsewhere. Then $f_Y(y) = \frac{dF_Y(a)}{da} = 1$ and Y (over a smaller range than X) is also a uniform distribution.

Problem 38 (the probability of roots)

The roots of the equation $4x^2 + 4Yx + Y + 2 = 0$ are given by

$$\begin{aligned} x &= \frac{-4Y \pm \sqrt{16Y^2 - 4(4)(y+2)}}{2(4)} \\ &= \frac{-Y \pm \sqrt{Y^2 - Y - 2}}{2}\,, \end{aligned}$$

which will be real if and only if $Y^2 - Y - 2 > 0$. Noticing that this expression factors into $(Y-2)(Y+1) > 0$, we see that this expression will be positive when $Y > 2$. Thus the probability we seek is given (if E is the event that x is real) by

$$P(E) = P\{Y > 2\} = \int_2^5 f(y)dy = \int_2^5 \frac{1}{5}dy = \frac{1}{5}(3) = \frac{3}{5}\,.$$

Problem 39 (the variable $Y = \log(X)$)

We begin by computing the cumulative distribution function for Y i.e. we have $F_Y(a)$

$$F_Y(a) = P\{Y \leq a\} = P\{\log(X) \leq a\} = P\{X \leq e^a\} = F_X(e^a)\,.$$

Now since X is an exponential random variable it has a cumulative distribution function given by $F_X(a) = 1 - e^{-\lambda a} = 1 - e^{-a}$, so that the above then becomes

$$F_Y(a) = 1 - e^{-e^a}\,.$$

The probability density function for Y is then the derivative of this expression with respect to a or

$$f_Y(a) = \frac{d}{da}F_Y(a) = \frac{d}{da}(1 - e^{-e^a}) = -e^{-e^a}(-e^a) = e^a e^{e^a} = e^{a+e^a}\,.$$

Problem 40 (the variable $Y = e^X$)

We begin by computing the cumulative distribution function for Y i.e.

$$F_Y(a) = P\{Y \leq a\} = P\{e^X \leq a\} = P\{X \leq \log(a)\} = F_X(\log(a))\,.$$

Now Since X is uniform distributed its cumulative distribution function is linear i.e. $F_X(a) = a$, so $F_Y(a) = \log(a)$ and the density function for Y is given by

$$f_Y(a) = \frac{d}{da} F_Y(a) = \frac{d}{da}(\log(a)) = \frac{1}{a} \quad \text{for} \quad 1 \leq a \leq e\,.$$

Problem 41 (the variable $R = A\sin(\theta)$)

We begin by computing the distribution function of the random variable R given by

$$F_R(a) \;=\; P\{R \leq a\} = P\{A\sin(\theta) \leq a\} = P\{\sin(\theta) \leq \frac{a}{A}\}\,.$$

To compute this we can plotting the function $\sin(\theta)$ for $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$ and see that the line $\frac{a}{A}$ crosses this function $\sin(\theta)$ at various points. These points determine the integration region of θ required to determined the probability above. We have that the probability above becomes

$$\int_{-\frac{\pi}{2}}^{\sin^{-1}(\frac{a}{A})} f_\theta(\theta) d\theta = \int_{-\frac{\pi}{2}}^{\sin^{-1}(\frac{a}{A})} \left(\frac{1}{\pi}\right) d\theta = \frac{1}{\pi}\left(\sin^{-1}(\frac{a}{A}) + \frac{\pi}{2}\right)\,.$$

From which we have a density function given by

$$\begin{aligned} f_R(a) &= \frac{d}{da} F_R(a) = \frac{d}{da}\left(\frac{1}{\pi}\sin^{-1}\left(\frac{a}{A}\right) + \frac{1}{2}\right) \\ &= \frac{1}{\pi} \frac{1}{\sqrt{1 - \left(\frac{a}{A}\right)^2}} \left(\frac{1}{A}\right) = \frac{A}{\pi\sqrt{A^2 - a^2}}\,. \end{aligned}$$

for $|a| \leq |A|$.

Chapter 5: Theoretical Exercises

Problem 1

Since $f(x)$ is a probability density if must integrate to one $\int_{-\infty}^{\infty} f(x)dx = 1$. In the case here using integration by parts this becomes

$$\begin{aligned} \int_0^\infty ax^2 e^{-bx^2} dx &= a \left.\frac{xe^{-bx^2}}{(-b)}\right|_0^\infty - a\int_0^\infty \left(\frac{e^{-bx^2}}{(-b)}\right) dx \\ &= 0 - 0 + \frac{a}{b}\int_0^\infty e^{-bx^2} dx\,. \end{aligned}$$

To evaluate this integral let $v = bx^2$ so that $dv = 2bxdx$, $x = \pm\sqrt{\frac{v}{b}}$, $dv = 2b\sqrt{\frac{v}{b}}dx$, which gives

$$dx = \left(\frac{b^{1/2}}{2b}\right) v^{-1/2} dv = \frac{v^{-1/2}}{2\sqrt{b}} dv\,,$$

and our integral above becomes

$$1 = \frac{a}{b}\frac{1}{2\sqrt{b}}\int_0^\infty v^{-1/2}e^{-v}dv\,.$$

Now the integral remaining can be seen to be

$$\int_0^\infty v^{-1/2}e^{-v}dv = \int_0^\infty v^{1/2-1}e^{-v}dv \equiv \Gamma(1/2) = \sqrt{\pi}\,.$$

Using this we have

$$1 = \frac{a}{2b^{3/2}}\sqrt{\pi}\,.$$

Thus $a = \frac{2b^{3/2}}{\sqrt{\pi}}$ is the relationship between a and b.

Problem 2

Consider the first integral

$$\int_0^\infty P\{Y < -y\}dy\,.$$

Using integration by parts with the substitutions $u(y) = P\{Y < -y\}$ and $dv(y) = dy$. Then the standard integration by parts differential formula becomes $u(y)dv(y) = u(y)v(y) - v(y)du(y)$, and the above becomes

$$\begin{aligned}\int_0^\infty P\{Y < -y\}dy &= P\{Y < -y\}y|_0^\infty - \int_0^\infty y\frac{d}{dy}P\{Y < -y\}dy \\ &= 0 - \int_0^\infty yf_Y(-y)(-1)dy = \int_0^\infty yf_Y(-y)dy\,.\end{aligned}$$

To evaluate this let $v = -y$ and the above integral becomes

$$\int_0^{-\infty} -vf_Y(v)(-dv) = -\int_{-\infty}^0 vf_Y(v)dv\,.$$

In the same way, for the second integral we have

$$\begin{aligned}\int_0^\infty P\{Y > y\}dy &= yP\{Y > y\}|_0^\infty - \int_0^\infty y\frac{d}{dy}P\{Y > y\}dy \\ &= -\int_0^\infty y\frac{d}{dy}P\{Y > y\}dy = -\int_0^\infty y\frac{d}{dy}(1 - P\{Y < y\})dy\,.\end{aligned}$$

Where we have used the fact that $P\{Y > y\} = 1 - P\{Y < y\}$. Taking the above derivative we get the above integral equal to

$$\int_0^\infty yf_Y(y)dy\,.$$

Now then

$$E[Y] = \int_{-\infty}^\infty yf_Y(y)dy = \int_{-\infty}^0 yf_Y(y)dy + \int_0^\infty yf_Y(y)dy\,.$$

Using the above two integral derived above we get this equal to

$$-\int_0^\infty P\{Y < -y\}dy + \int_0^\infty P\{Y > y\}dy\,.$$

Problem 3

From Problem 2 above we have that

$$\begin{aligned} E[g(X)] &= \int_0^\infty P\{g(X) > y\} dy - \int_0^\infty P\{g(X) < -y\} dy \\ &= \int_0^\infty \int_{x:g(x)>y} f(x) dx dy - \int_0^\infty \int_{x:g(x)<-y} f(x) dx dy \,. \end{aligned}$$

Now as in the proof of proposition 2.1 we can change the order of each integration with the following manipulations

$$\begin{aligned} E[g(X)] &= \int_{x:g(x)>0} \int_0^{g(x)} dy f(x) dx - \int_{x:g(x)<0} \int_0^{-g(x)} dy f(x) dx \\ &= \int_{x:g(x)>0} f(x) g(x) dx + \int_{x:g(x)<0} f(x) g(x) dx = \int_{-\infty}^{\infty} g(x) f(x) dx \,. \end{aligned}$$

Problem 4

Corollary 2.1 is $E[ax+b] = aE[x]+b$ and can be proven using the definition of the expectation

$$E[ax+b] = \int_{-\infty}^{\infty} (ax+b) f(x) dx = a \int_{-\infty}^{\infty} x f(x) dx + b \int_{-\infty}^{\infty} f(x) dx = aE[X] + b \,.$$

Problem 5

Compute $E[X^n]$ by using the given identity i.e.

$$E[X^n] = \int_0^\infty P\{X^n > t\} dt \,.$$

To evaluate this let $t = x^n$ then $dt = nx^{n-1}$ and we have

$$E[X^n] = \int_0^\infty P\{X^n > x^n\} n x^{n-1} dx = \int_0^\infty n x^{n-1} P\{X > x\} dx \,,$$

Using the fact that $P\{X^n > x^n\} = P\{X > x\}$ when X is a non-negative random variable.

Problem 6

We want a collection of events E_a with $0 < a < 1$ such that $P(E_a) = 1$ but $P(\cap_a E_a) = 0$. Let X be a uniform random variable over $(0, 1)$ and let the event E_a be the event that $X \neq a$. Since X is a continuous random variable $P(E_a) = 1$, since the probability $X = a$ (exactly) must be zero. Now the event $\cap_a E_a$ is the event that X is not any of the elements from $(0, 1)$. Since X must be at least one of these elements the probability of this intersection must be zero i.e. $P(\cap_a E_a) = 0$.

Problem 7

We have

$$\begin{aligned}\mathrm{Std}(aX+b) &= \sqrt{\mathrm{Var}(aX+b)} = \sqrt{\mathrm{Var}(aX)} \\ &= \sqrt{a^2\mathrm{Var}(X)} = |a|\sigma\,.\end{aligned}$$

Problem 8

We know that $P\{0 \le X \le c\} = 1$ and we want to show $\mathrm{Var}(X) \le \frac{c^2}{4}$. Following the hint we have

$$E[X^2] = \int_0^c x^2 f_X(x)dx \le c\int_0^c x f_X(x)dx = cE[X]\,,$$

since $x \le c$ for all $x \in [0, c]$. Now

$$\begin{aligned}\mathrm{Var}(X) &= E[X^2] - E[X]^2 \le cE[X] - E[X]^2 \\ &= c^2\left(\frac{E[X]}{c} - \frac{E[X]^2}{c^2}\right)\,.\end{aligned}$$

Define $\alpha = \frac{E[X]}{c}$ and we have

$$\mathrm{Var}(X) \le c^2(\alpha - \alpha^2) = c^2\alpha(1-\alpha)\,.$$

Now to select the value of α that maximizes the expression $\alpha(1-\alpha)$ for α in the range $0 \le \alpha \le 1$ we take the derivative with respect to α, set this expression equal to zero, and solve for α. The derivative gives

$$c^2(1-2\alpha) = 0\,,$$

which gives $\alpha = \frac{1}{2}$. A second derivative gives $-2c^2$ which is negative, showing that $\alpha = \frac{1}{2}$ is a maximum. The value at this maximum is

$$\frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}\,,$$

and so we have that $\mathrm{Var}(X) \le \frac{c^2}{4}$.

Problem 9

Part (a): $P\{Z > x\} = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$. Let $v = -z$ so that $dv = -dz$ and we have the above given by

$$\int_{-x}^{-\infty} \frac{1}{\sqrt{2\pi}} e^{-v^2/2}(-dv) = \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = P\{Z < -x\}\,.$$

Part (b): We find

$$
\begin{aligned}
P\{|Z| > x\} &= \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz + \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= \int_\infty^x \frac{1}{\sqrt{2\pi}} e^{-z^2/2} (-dz) + \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= 2 \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 2P\{Z > x\} .
\end{aligned}
$$

Part (c): We find

$$
\begin{aligned}
P\{|Z| \leq x\} &= \int_{-x}^x \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz - \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= P\{Z < x\} + \int_\infty^x \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= P\{Z < x\} - \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= P\{Z < x\} - \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz - \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz + \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= P\{Z < x\} + \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz - 1 \\
&= 2P\{Z < x\} - 1 .
\end{aligned}
$$

Problem 10 (points of inflection of the Gaussian

We are told that $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}$. And points of inflection are given by $f''(x) = 0$. To calculate $f''(x)$ we need $f'(x)$. We find

$$
f'(x) \approx -\left(\frac{x-\mu}{\sigma^2}\right) \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\} .
$$

So that the second derivative is given by

$$
f''(x) \approx -\frac{1}{\sigma^2} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\} + \left(\frac{(x-\mu)^2}{\sigma^2}\right) \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\} .
$$

Setting $f''(x)$ equal to zero we find that this requires x satisfy

$$
\exp\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\} \left[-1 + \frac{(x-\mu)^2}{\sigma^2}\right] = 0, .
$$

or $(x-\mu)^2 = \sigma^2$. Which has as solutions $x = \mu \pm \sigma$.

Problem 11 ($E[X^2]$ of an exponential random variable)

Theoretical Exercise number 5 states that

$$E[X^n] = \int_0^\infty n x^{n-1} P\{X > x\} dx \,.$$

For an exponential random variable we have our cumulative distribution function given by

$$P\{X \le x\} = 1 - e^{-\lambda x} \,.$$

so that $P\{X > x\} = e^{-\lambda x}$, and thus our expectation becomes

$$E[X^n] = \int_0^\infty n x^{n-1} e^{-\lambda x} dx$$

Now if $n = 2$ we find that this expression becomes in this case

$$\begin{aligned} E[X^2] &= \int_0^\infty 2x e^{-\lambda x} dx \\ &= 2 \int_0^\infty x e^{-\lambda x} dx \\ &= 2 \left[\left. \frac{x e^{-\lambda x}}{-\lambda} \right|_0^\infty + \frac{1}{\lambda} \int_0^\infty e^{-\lambda x} dx \right] \\ &= \frac{2}{\lambda} \left[\left. \frac{e^{-\lambda x}}{-\lambda} \right|_0^\infty \right] = \frac{2}{\lambda^2} \,, \end{aligned}$$

as expected.

Problem 12 (the median of a continuous random variable)

Part (a): When X is uniformly distributed over (a, b) the median is the value m that solves

$$\int_a^m \frac{dx}{b-a} = \int_m^b \frac{dx}{b-a} \,.$$

Integrating both sides gives that $m - a = b - m$, which has a solution of $m = \frac{a+b}{2}$.

Part (b): When X is a normal random variable with parameters (μ, σ^2) we find that m must satisfy

$$\int_{-\infty}^m \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\} dx = \int_m^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\} dx \,.$$

To evaluate the integral on both sides of this expression we let $v = \frac{x-\mu}{\sigma}$, so that $dv = \frac{dx}{\sigma}$ and each integral becomes

$$\begin{aligned} \int_{-\infty}^{\frac{m-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{v^2}{2}\} dv &= \int_{\frac{m-\mu}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}} \exp\{-\frac{v^2}{2}\} dv \\ &= 1 - \int_{-\infty}^{\frac{m-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{v^2}{2}\} dv \,. \end{aligned}$$

Remembering the definition of the cumulative distribution function $\Phi(\cdot)$ as

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy \,,$$

we see that the above can be written as $\Phi(\frac{m-\mu}{\sigma}) = 1 - \Phi(\frac{m-\mu}{\sigma})$, so that

$$2\Phi(\frac{m-\mu}{\sigma}) = 1 \quad \text{or} \quad \Phi(\frac{m-\mu}{\sigma}) = \frac{1}{2}$$

Thus we have $m = \mu + \sigma\Phi^{-1}(1/2)$, since we can compute Φ^{-1} using the Matlab function `norminv`, we find that $\Phi^{-1}(1/2) = 0$, which intern implies that $m = \mu$.

Part (c): If X is an exponential random variable with rate λ then m must satisfy

$$\int_0^m \lambda e^{-\lambda x} dx = \int_m^\infty \lambda e^{-\lambda x} dx = 1 - \int_0^m \lambda e^{-\lambda x} dx \,.$$

Introducing the cumulative distribution function for the exponential distribution (given by $F(x) = \int_0^m \lambda e^{-\lambda x} dx$) the above equation can be seen to be $F(m) = 1 - F(m)$ or $F(m) = \frac{1}{2}$. So in general the median m is given by $m = F^{-1}(1/2)$ where F is the cumulative distribution function. For the exponential random variable this expression gives

$$1 - e^{-\lambda m} = \frac{1}{2} \quad \text{or} \quad m = \frac{\ln(2)}{\lambda} \,.$$

Problem 14 (if X is an exponential random variable then cX is)

If X is an exponential random variable with parameter λ, then defining $Y = cX$ the distribution function for Y is given by

$$\begin{aligned} F_Y(a) &= P\{Y \le a\} \\ &= P\{cX \le a\} \\ &= P\left\{X \le \frac{a}{c}\right\} \\ &= F_X(\frac{a}{c}) \,. \end{aligned}$$

So, taking the derivative of the above expression, to obtain the density function for Y we see that

$$\begin{aligned} f_Y(a) &= \frac{dF_Y}{da} \\ &= \frac{d}{da} F_X(\frac{a}{c}) \\ &= F_X'(\frac{a}{c}) \frac{1}{c} \\ &= \frac{1}{c} f_X(\frac{a}{c}) \end{aligned}$$

But since X is an exponential random variable with parameters λ we have that

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

so that we have for $f_Y(y)$ the following

$$f_Y(y) = \frac{1}{c} \begin{cases} \lambda e^{-\lambda \frac{y}{c}} & \frac{y}{c} \geq 0 \\ 0 & \frac{y}{c} < 0 \end{cases}$$

or

$$f_Y(y) = \begin{cases} \frac{\lambda}{c} e^{-\frac{\lambda}{c} y} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

showing that Y is another exponential random variable with parameter $\frac{\lambda}{c}$.

Problem 15

The hazard rate function $\lambda(t)$ is defined by $\lambda(t) = \frac{f(t)}{\bar{F}(t)} = \frac{f(t)}{1-F(t)}$. For a uniform random variable distributed between $(0, a)$ we have

$$f(t) = \begin{cases} \frac{1}{a} & 0 \leq t \leq a \\ 0 & \text{otherwise} \end{cases}$$

and

$$F(t) = \int_0^t f(t')dt' = \int_0^t \frac{dt'}{a} = \frac{t}{a},$$

so the hazard rate function then is

$$\lambda(t) = \frac{(1/a)}{1 - \frac{t}{a}} = \frac{1}{a - t},$$

for $0 \leq t \leq a$.

Problem 16

For this problem if we are told that X has a hazard rate function $\lambda_X(t)$ we desire to compute the hazard rate function for $Y = aX$, with $a > 0$. When $Y = aX$ the probability density function of Y is given by $f_Y(y) = f_X(y/a)\left(\frac{1}{a}\right)$ and its distribution function is given by

$$F_y(c) = P\{Y \leq c\} = P\{aX \leq c\} = P\{X \leq \frac{c}{a}\} = F_X(\frac{c}{a}),$$

so the hazard rate for Y is given by

$$\lambda_Y(t) = \frac{f_Y(t)}{1 - F_Y(t)} = \frac{f_X(\frac{t}{a})\left(\frac{1}{a}\right)}{1 - F_X(t/a)} = \left(\frac{1}{a}\right)\left(\frac{f_X(t/a)}{1 - F_X(t/a)}\right) = \left(\frac{1}{a}\right)\lambda_X(t/a).$$

Problem 17

The Gamma density function is given by

$$f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \quad x \geq 0\,.$$

To check that it integrates to one we have

$$\int_0^\infty f(x)dx = \int_0^\infty \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} dx\,.$$

To evaluate this let $y = \lambda x$ so that $dy = \lambda dx$ to get

$$\int_0^\infty \frac{e^{-y} y^{\alpha-1}}{\Gamma(\alpha)} dy = \frac{1}{\Gamma(a)} \int_0^\infty e^{-y} y^{\alpha-1} dy\,.$$

We note that the integral on the right hand side $\int_0^\infty e^{-y} y^{\alpha-1} dy$ is the *definition* of the gamma function, $\Gamma(a)$, and the above becomes equal to one showing that the Gamma density integrates to one.

Problem 18 (the expectation of X^k when X is exponential)

If X is exponential with mean $1/\lambda$ then $f(x) = \lambda e^{-\lambda x}$ so that

$$E[X^k] = \int_0^\infty \lambda x^k e^{-\lambda x} dx = \lambda \int_0^\infty x^k e^{-\lambda x} dx\,.$$

To transform to the gamma integral, let $v = \lambda x$, so that $dv = \lambda dx$ and the above integral becomes

$$\lambda \int_0^\infty \frac{v^k}{\lambda^k} e^{-v} \frac{dv}{\lambda} = \lambda^{-k} \int_0^\infty v^k e^{-v} dv\,.$$

Remembering the definition of the Γ function as $\int_0^\infty v^k e^{-v} dv \equiv \Gamma(k+1)$ and that when k is an integer $\Gamma(k+1) = k!$, we see that the above integral is equal to $k!$ and we have that

$$E[X^k] = \frac{k!}{\lambda^k}\,,$$

as required.

Problem 19 (the variance of a gamma random variable)

If X is a gamma random variable then

$$f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}\,,$$

when $x \geq 0$ and is zero otherwise. To compute the variance we require $E[X^2]$ which is given by

$$\begin{aligned} E[X^2] &= \int_0^\infty x^2 f(x) dx \\ &= \int_0^\infty x^2 \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+1} e^{-\lambda x} dx \,. \end{aligned}$$

To evaluate the above integral, let $v = \lambda x$ so that $dv = \lambda dx$ then the above becomes

$$\frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \frac{v^{\alpha+1}}{\lambda^{\alpha+1}} e^{-v} \frac{dv}{\lambda} = \frac{\lambda^\alpha}{\lambda^{\alpha+2}\Gamma(\alpha)} \int_0^\infty v^{\alpha+1} e^{-v} dv = \frac{\Gamma(\alpha+2)}{\lambda^2 \Gamma(\alpha)} \,.$$

Where we have used the definition of the gamma function in the above. If we "factor" the gamma function as

$$\Gamma(\alpha+2) = (\alpha+1)\Gamma(\alpha+1) = (\alpha+1)\alpha\Gamma(\alpha) \,,$$

we see that

$$E[X^2] = \frac{\alpha(\alpha+1)}{\lambda^2} \,,$$

when X is a gamma random variable with parameters (α, λ). Since $E[X] = \frac{\alpha}{\lambda}$ we can compute $\text{Var}(X) = E[X^2] - E[X]^2$ as

$$\text{Var}(X) = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2} \,,$$

as claimed.

Problem 20 (the gamma function at $1/2$)

We want to consider $\Gamma(1/2)$ which is defined as

$$\Gamma(1/2) = \int_0^\infty x^{-1/2} e^{-x} dx \,.$$

Since the argument of the exponential is the square of the term $x^{1/2}$ this observation might motivate the substitution $y = \sqrt{x}$. Following the hint let $y = \sqrt{2x}$, so that

$$dy = \frac{1}{\sqrt{2x}} dx \,.$$

So that with this substitution $\Gamma(1/2)$ becomes

$$\Gamma(1/2) = \int_0^\infty \sqrt{2}\, dy\, e^{-y^2/2} = \sqrt{2} \int_0^\infty e^{-y^2/2} dy \,.$$

Now from the normalization of the standard Gaussian we know that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{y^2}{2}\} dy = 1\,,$$

which easily transforms (by integrating only over the positive real numbers) into

$$2\int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{y^2}{2}\} dy = 1\,.$$

Finally manipulating this into the specific integral required to evaluate $\Gamma(1/2)$ we find that

$$\sqrt{2}\int_{0}^{\infty} \exp\{-\frac{y^2}{2}\} dy = \sqrt{\pi}\,,$$

which shows that $\Gamma(1/2) = \sqrt{\pi}$ as requested.

Problem 21 (the hazard rate function for the gamma random variable)

The hazard rate function for a random variable T that has a density function $f(t)$ and distribution function $F(t)$ is given by

$$\lambda(t) = \frac{f(t)}{1 - F(t)}\,.$$

For a gamma distribution with parameters (α, λ) we know our $f(t)$ is given by

$$f(t) = \begin{cases} \frac{\lambda e^{-\lambda t}(\lambda t)^{\alpha-1}}{\Gamma(\alpha)} & t \geq 0 \\ 0 & t < 0\,. \end{cases}$$

Lets begin by calculating the cumulative density function for a gamma random variable with parameters (α, λ). We find that

$$F(t) = \int_0^t f(\xi) d\xi = \int_0^t \frac{\lambda e^{-\lambda\xi}(\lambda\xi)^{\alpha-1}}{\Gamma(\alpha)} d\xi\,,$$

which cannot be simplified further. We then have that

$$\begin{aligned} 1 - F(t) &= \int_0^\infty f(\xi)d\xi - \int_0^t f(\xi)d\xi \\ &= \int_t^\infty f(\xi)d\xi \\ &= \int_t^\infty \frac{\lambda e^{-\lambda\xi}(\lambda\xi)^{\alpha-1}}{\Gamma(\alpha)} d\xi\,, \end{aligned}$$

which also cannot be simplified further. Thus our hazard rate is given by

$$\begin{aligned} \lambda(t) &= \frac{\frac{\lambda e^{-\lambda t}(\lambda t)^{\alpha-1}}{\Gamma(\alpha)}}{\int_t^\infty \frac{\lambda e^{-\lambda\xi}(\lambda\xi)^{\alpha-1}}{\Gamma(\alpha)} d\xi} \\ &= \frac{t^{\alpha-1}e^{-\lambda t}}{\int_t^\infty \xi^{\alpha-1}e^{-\lambda\xi}d\xi} \\ &= \frac{1}{\int_t^\infty \left(\frac{\xi}{t}\right)^{\alpha-1} e^{-\lambda(\xi-t)}d\xi}\,. \end{aligned}$$

To try and simplify this further let $v = \frac{\xi}{t}$ so that $dv = \frac{d\xi}{t}$, and the above becomes

$$\lambda(t) = \frac{1}{\int_1^\infty v^{\alpha-1} e^{-\lambda t(v-1)} t dv} = \frac{1}{te^{\lambda t} \int_1^\infty v^{\alpha-1} e^{-\lambda t v} dv} .$$

Which is one expression for the hazard rate for a gamma random variable. We can try and reduce the integral in the bottom of the above fraction to that of the "upper incomplete gamma function" by making the substitution $y = \lambda t v$ so that $dy = \lambda t dv$ and obtaining

$$\begin{aligned}
\lambda(t) &= \frac{1}{te^{\lambda t} \int_{\lambda t}^\infty \frac{y^{\alpha-1}}{(\lambda t)^{\alpha-1}} e^{-y} \frac{dy}{\lambda t}} \\
&= \frac{(\lambda t)^\alpha}{te^{\lambda t}} \frac{1}{\int_{\lambda t}^\infty y^{\alpha-1} e^{-y} dy} \\
&= \frac{(\lambda t)^\alpha}{te^{\lambda t}} \frac{1}{\Gamma(\alpha, \lambda t)} .
\end{aligned}$$

Where we have introduced the **upper incomplete gamma function** whos definition is given by

$$\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt .$$

Problem 27 (modality of the beta distribution)

The beta distribution with parameters (a, b) has a probability density function given by

$$f(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} \quad \text{for} \quad 0 \le x \le 1 .$$

Part (a): Our mode of this distribution will equal either the endpoints of our interval i.e. $x = 0$ or $x = 1$ or the location where the first derivative of $f(x)$ vanishes. Computing this derivative the expression $\frac{df}{dx} = 0$ implies

$$\begin{aligned}
\frac{df}{dx}(x) &= (a-1)x^{a-2}(1-x)^{b-1} + (b-1)x^{a-1}(1-x)^{b-2}(-1) = 0 \\
&\Rightarrow x^{a-2}(1-x)^{b-2} \left[(2-a-b)x + (a-1) \right] = 0 ,
\end{aligned}$$

which can be solved for the x^* that makes this an equality and gives

$$x^* = \frac{a-1}{a+b-2} \quad \text{assuming} \quad a+b-2 \neq 0 .$$

In this case to guarantee that this is a *maximum* we should check that the second derivative of f at the value of $\frac{a-1}{a+b-2}$ is indeed *negative*. This second derivative is computed in the Mathematica file `chap_5_te_27.nb` where it is shown to be negative for the given domains of a and b. To guarantee that this value is *interior* to the interval $(0, 1)$ we should verify that

$$0 < \frac{a-1}{a+b-2} < 1$$

which since $a + b - 2 > 0$ is equivalent to

$$0 < a - 1 < a + b - 2 \, .$$

or from the first inequality we have that $a > 1$ and from the second inequality ($a-1 < a+b-2$) we have that $b > 1$ verifying that our point x^* is in the interior of this interval and our distribution is unimodal as was asked.

Part (b): Now the case when $a = b = 1$ is covered below, so lets consider $a = 1$. From the requirement $a + b < 2$ we must have $b < 1$ and our density function in this case is given by

$$f(x) = \frac{(1-x)^{b-1}}{B(1,b)} \, .$$

This has a derivative given by

$$f'(x) = \frac{(1-b)(1-x)^{b-2}}{B(1,b)} \, ,$$

and is *positive* over the entire interval since $b < 1$. Because the derivative is positive over the entire domain the distribution is unimodal and the single mode will occur at the right most limit i.e. $x = 1$. Now if $b = 1$ in the same way we have $a < 1$ and our density function is given by

$$f(x) = \frac{x^{a-1}}{B(a,1)} \, .$$

Which has a derivative given by

$$f'(x) = \frac{(a-1)x^{a-2}}{B(a,1)} \, ,$$

and is *negative* because $a < 1$. Because the derivative is negative over the entire domain the distribution is unimodal and the unique mode will occur at the left most limit of our domain i.e. $x = 0$. Finally, we consider the case where $a < 1$, $b < 1$ and neither equal to one. In this case from the derivative above our minimum or maximum is given by $\frac{a-1}{a+b-2}$ which for the domain of a and b given here is *positive* implying that the point x^* is a minimum. Thus we have two local maximums at the endpoints $x = 0$ and $x = 1$. One can also show (in the same way as above) that for this domain of a and b the point x^* is in the interior of the interval.

Part (c): If $a = b = 1$, then the density function for the beta distribution becomes (since $\text{Beta}(1,1) \equiv B(1,1) = 1$) is

$$f(x) = 1 \, ,$$

and we have the density of the uniform distribution, which is “flat” and has all points modes.

Problem 28 ($Y = F(X)$ is a uniform random variable)

If $Y = F(X)$ then the distribution function of Y is given by

$$\begin{aligned} F_Y(a) &= P\{Y \leq a\} \\ &= P\{F(X) \leq a\} \\ &= P\{X \leq F^{-1}(a)\} \\ &= F(F^{-1}(a)) = a\,. \end{aligned}$$

Thus $f_Y(a) = \frac{dF_Y}{da} = 1$, showing that Y is a uniform random variable.

Problem 29 (the probability density function for $Y = aX + b$)

We begin by computing the cumulative distribution function of the random variable Y as

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} \\ &= P\{aX + b \leq y\} \\ &= P\{X \leq \frac{y-b}{a}\} \\ &= F_X(\frac{y-b}{a})\,. \end{aligned}$$

Taking the derivative to obtain the distribution function for Y we find that

$$f_Y(y) = \frac{dF_Y}{dy} = F_X'(\frac{y-b}{a})\frac{1}{a} = \frac{1}{a}f_X(\frac{y-b}{a})\,.$$

Problem 30 (the probability density function for the lognormal distribution)

We begin by computing the cumulative distribution function of the random variable Y as

$$\begin{aligned} F_Y(a) &= P\{Y \leq a\} \\ &= P\{e^X \leq a\} \\ &= P\{X \leq \log(a)\} \\ &= F_X(\log(a))\,. \end{aligned}$$

Since X is a normal random variable with mean μ and variance σ^2 it has a cumulative distribution function given by

$$F_X(a) = \Phi\left(\frac{a-\mu}{\sigma}\right)$$

so that the cumulative distribution function for Y becomes

$$F_Y(a) = \Phi\left(\frac{\log(a)-\mu}{\sigma}\right)\,.$$

The density function for the random variable Y is given by the derivative of the cumulative distribution function thus we have

$$f_Y(a) = \frac{F_Y(a)}{da} = \Phi'\left(\frac{\log(a) - \mu}{\sigma}\right)\left(\frac{1}{\sigma}\right)\left(\frac{1}{a}\right).$$

Since $\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ we have for the probability density function for a lognormal random variable given by

$$f_Y(a) = \frac{1}{\sqrt{2\pi}\sigma a} \exp\left\{-\frac{1}{2}\frac{(\log(a) - \mu)^2}{\sigma}\right\}.$$

Problem 31 (Legendre's theorem on relatively primeness)

Part (a): If k is the greatest common divisor of *both* X and Y then k must divide the random variable X and the random variable Y. In addition, X/k and Y/k must be relatively prime i.e. have no common factors. Now to show the given probability we first argue that that k will divide X with probability $1/k$ (approximately) and divide Y with probability $1/k$ (approximately). This can be reasoned heuristically by considering the case where X and Y are drawn from say $1, 2, \ldots, 10$. Then if $k = 2$ the numbers five numbers $2, 4, 6, 8, 10$ are all divisible by 2 and so the probability 2 will divide a random number from this set is $5/10 = 1/2$. If $k = 3$ then the three numbers $3, 6, 9$ are all divisible by 3 and so the probability 3 will divide a random number from this set is $3/10 \approx 1/3$. In the same way when $k = 4$ the probability that 4 will divide one of the numbers in our set is $2/10 = 1/5 \approx 1/4$. These approximations become exact as N goes to infinity. Finally, X/k and Y/k will be relatively prime with probability Q_1. Letting $E_{X,k}$ to be event that X is divisible by k, $E_{Y,k}$ the event that Y is divisible by k, and $E_{X/k,Y/k}$ the event that X/k and Y/k are relatively prime we have that

$$\begin{aligned} Q_k &= P\{D = k\} \\ &= P\{E_{X,k}\}P\{E_{Y,k}\}P\{E_{X/k,Y/k}\} \\ &= \left(\frac{1}{k}\right)\left(\frac{1}{k}\right)Q_1. \end{aligned}$$

which is the desired results.

Part (b): From above we have that $Q_k = Q_1/(k^2)$, so summing both sides for $k = 1, 2, 3, \cdots$ gives (since $\sum_k Q_k = 1$, i.e. the greatest common divisor must be one of the numbers $1, 2, 3, \vdots$)

$$1 = Q_1 \sum_{k=1}^{\infty} \frac{1}{k^2},$$

which gives the desired result of

$$Q_1 = \frac{1}{\sum_{k=1}^{\infty} \frac{1}{k^2}}.$$

Since $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$ the above expression for Q_1 becomes

$$Q_1 = \frac{1}{\frac{\pi^2}{6}} = \frac{6}{\pi^2}.$$

Part (c): Now Q_1 is the probability that X and Y are relatively prime will be true if $P_1 = 2$ is not a divisor of X and Y. The probability that P_1 is not a divisor of X is $1/P_1$ and the same for Y. So the probability that P_1 is a divisor for *both* X and Y is $(1/P_1)^2$. The probability that P_1 is *not* a divisor of both will happen with probability $1 - (1/P_1)^2$. The same logic applies for P_2 giving that the probability that X and Y don't have P_2 as a factor is $1 - (1/P_2)^2$. Since for X and Y be be relatively prime they cannot have any P_i as a joint factor, and thus we are looking for the conjunction of each of the individual probabilities. This is that P_1 is not a divisor, that P_2 is not a divisor, etc. This requires the product of all of these terms giving for Q_1 that

$$Q_1 = \prod_{i=1}^{\infty} \left(1 - \frac{1}{P_i^2}\right) = \prod_{i=1}^{\infty} \left(\frac{P_i^2 - 1}{P_i^2}\right) .$$

Problem 32 (the P.D.F. for $Y = g(X)$, when g is decreasing)

Theorem 7.1 expresses how to obtain the probability density function for Y when $Y = g(X)$ and the probability density function for X is known. To prove this result in the case when $g(\cdot)$ is decreasing lets compute the cumulative distribution function for Y i.e.

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} \\ &= P\{g(X) \leq y\} \end{aligned}$$

By plotting a typical decreasing function $g(x)$ we see that the set above is given by the set of x values such that $x \geq g^{-1}(y)$ and the above expression becomes

$$F_Y(y) = \int_{g^{-1}(y)}^{\infty} f(x)dx .$$

Talking the derivative of this expression with respect to y we obtain

$$F_Y'(y) = f(g^{-1}(y))(-1)\frac{dg^{-1}(y)}{dy} .$$

Since $\frac{dg^{-1}(y)}{dy}$ is negative

$$(-1)\frac{dg^{-1}(y)}{dy} = \left|\frac{dg^{-1}(y)}{dy}\right| ,$$

and using this in the above the theorem in this case is proven.

Chapter 5: Self-Test Problems and Exercises

Problem 1 (playing times for basketball)

Part (a): The probability that the players plays over fifteen minute is given by

$$\begin{aligned} \int_{15}^{40} f(x)dx &= \int_{15}^{20} 0.025dx + \int_{20}^{30} 0.05dx + \int_{30}^{40} 0.025dx \\ &= 0.025 \cdot (5) + 0.05 \cdot (10) + 0.025 \cdot (10) = 0.875 . \end{aligned}$$

Part (b): The probability that the players plays between 20 and 35 minute is given by

$$\begin{aligned}\int_{20}^{35} f(x)dx &= \int_{20}^{30} 0.05dx + \int_{30}^{35} 0.025dx \\ &= 0.05 \cdot (10) + 0.025 \cdot (5) = 0.625\,.\end{aligned}$$

Part (c): The probability that the players plays less than 30 minutes is given by

$$\begin{aligned}\int_{10}^{30} f(x)dx &= \int_{10}^{20} 0.025dx + \int_{20}^{30} 0.05dx \\ &= 0.025 \cdot (10) + 0.05 \cdot (10) = 0.75\,.\end{aligned}$$

Part (d): The probability that the players plays more than 36 minutes is given by

$$\int_{36}^{40} f(x)dx = 0.025 \cdot (4) = 0.1\,.$$

Problem 2 (a power law probability density)

Part (a): Our random variable must normalize so that $\int f(x)dx = 1$, or

$$\int_0^1 cx^n dx = c\frac{x^{n+1}}{n+1}\bigg|_0^1 = \frac{c}{n+1}\,.$$

so that from the above we see that $c = n+1$. Our probability density function is then given by

$$f(x) = \begin{cases} (n+1)x^n & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Part (b): This expression is then given by

$$P\{X > x\} = \int_x^1 (n+1)\xi^n d\xi = \xi^{n+1}\big|_x^1 = 1 - x^{n+1} \quad \text{for} \quad 0 < x < 1\,.$$

Thus we have

$$P\{X > x\} = \begin{cases} 1 & x < 0 \\ 1 - x^{n+1} & 0 < x < 1 \\ 0 & x > 1 \end{cases}$$

Problem 3 (computing $E[X]$ and $\mathrm{Var}(X)$)

Given this $f(x)$ we compute the constant c by requiring $\int f(x)dx = 1$, which in this case is

$$\int_0^2 cx^4 dx = \frac{cx^5}{5}\bigg|_0^2 = \frac{c}{5}(2^5) = 1\,,$$

so $c = \frac{5}{32}$. Thus we can compute the expectation of X as

$$E[X] = \int_0^2 x\left(\frac{5}{32}\right)x^4 dx = \frac{5}{32}\int_0^2 x^5 dx = \frac{5}{32}\frac{x^6}{6}\Big|_0^2 = \frac{5}{3}.$$

And $\text{Var}(X) = E[X^2] - E[X]^2$, which means that we need

$$E[X^2] = \int_0^2 x^2\left(\frac{5}{32}\right)x^4 dx = \frac{5}{32}\int_0^2 x^6 dx = \frac{5}{2^5}\frac{x^7}{7}\Big|_0^2 = \frac{5\cdot 2^7}{2^5\cdot 7} = \frac{20}{7}.$$

Thus $\text{Var}(X) = \frac{20}{7} - \frac{25}{9} = \frac{5}{63}$.

Problem 4 (a continuous density)

Our $f(x)$ must integrate to one

$$\int_0^1 (ax + bx^2)dx = \frac{ax^2}{2} + \frac{bx^3}{3}\Big|_0^1 = \frac{a}{2} + \frac{b}{3} = 1.$$

We also told that $E[X] = 0.6$ so

$$E[X] = \int_0^1 (ax^2 + bx^3)dx = \frac{ax^3}{3} + \frac{bx^4}{4}\Big|_0^1 = \frac{a}{3} + \frac{b}{4} = 0.6.$$

Solving for the a and b we have the following system

$$\begin{bmatrix} 1/2 & 1/3 \\ 1/3 & 1/4 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0.6 \end{bmatrix}.$$

By Cramer's rule we have

$$a = \frac{\begin{vmatrix} 1 & 1/3 \\ 0.6 & 1/4 \end{vmatrix}}{\begin{vmatrix} 1/2 & 1/3 \\ 1/3 & 1/4 \end{vmatrix}} = \left(\frac{1}{20}\right)\left(\frac{72}{1}\right) = 3.6,$$

and

$$b = \frac{\begin{vmatrix} 1/2 & 1 \\ 1/3 & 0.6 \end{vmatrix}}{\frac{1}{8} - \frac{1}{9}} = -2.4.$$

Now for Part (a)

$$P\{X < 1/2\} = \int_0^{1/2} (3.6x - 2.4x^2)dx = \frac{3.6x^2}{2}\Big|_0^{1/2} - \frac{2.4x^3}{3}\Big|_0^{1/2} = 0.45 - 0.1 = 0.35.$$

For Part (b) we have $\text{Var}(X) = E[X^2] - E[X]^2$ so

$$\begin{aligned} E[X^2] &= \int_0^1 x^2(3.6x - 2.4x^2)dx = \int_0^1 (3.6x^3 - 2.4x^4)dx \\ &= \frac{3.6x^4}{4} - \frac{2.4x^5}{5}\Big|_0^1 = \frac{3.6}{4} - \frac{2.4}{5} = 0.42. \end{aligned}$$

So that $\text{Var}(X) = 0.42 - (0.6)^2 = 0.06$.

Problem 5 (a discrete uniform random variable)

We want to prove that $X = \text{int}(nU) + 1$ is a uniform random variable. To prove this first fix n, then $X = i$ is true if and only if

$$\text{Int}(nU) + 1 = i \quad \text{for} \quad i = 1, 2, 3, \cdots, n\,.$$

or

$$\text{Int}(nU) = i - 1\,.$$

or

$$\frac{i-1}{n} \le U < \frac{i}{n} \quad \text{for} \quad i = 1, 2, 3, \cdots, n$$

Thus the probability that $X = i$ is equal to

$$P\{X = i\} = \int_{\frac{i-1}{n}}^{\frac{i}{n}} 1 d\xi = \frac{i}{n} - \frac{i-1}{n} = \frac{1}{n} \quad \text{for} \quad i = 1, 2, 3, \cdots, n\,.$$

Problem 6 (bidding on a contract)

Assume we select a bid price b. Then our profit will be $b - 100$ if get the contract and zero if we don't get the contract. Thus our profit is a random variable that depends on the bid received by the competing company u. Our profit is then given by (here P is for *profit*)

$$P(b) = \begin{cases} 0 & b > u \\ b - 100 & b < u \end{cases}$$

Lets compute the expected profit

$$\begin{aligned} E[P(b)] &= \int_{70}^{b} 0 \cdot \frac{1}{140 - 70} d\xi + \int_{b}^{140} (b - 100) \cdot \frac{1}{140 - 70} d\xi \\ &= \frac{(b - 100)(140 - b)}{70} = \frac{240b - b^2 - 14000}{70}\,. \end{aligned}$$

Then to find the maximum of the expected profit we take the derivative of the above expression with respect to b, setting that expression equal to zero and solve for b. The derivative set equal to zero is given by

$$\frac{dE[P(b)]}{db} = \frac{1}{70}(240 - 2b) = 0\,.$$

Which has $b = 120$ as a solution. Since $\frac{d^2 E[P(b)]}{db^2} = -\frac{2}{70} < 0$, this value of b is indeed a maximum of the function $P(b)$. Using this value of b our expected profit is given by $\frac{400}{70} = \frac{40}{7}$.

Problem 7

Part (a): We want to compute $P\{U \geq 0.1\} = \int_{0.1}^{1} d\xi = 0.9$.

Part (b): We want to compute

$$\begin{aligned} P\{U \geq 0.2 | U \geq 0.1\} &= \frac{P\{U \geq 0.2, U \geq 0.1\}}{P\{U \geq 0.1\}} \\ &= \frac{P\{U \geq 0.2\}}{P\{U \geq 0.1\}} = \frac{1-0.2}{1-0.1} = \frac{0.8}{0.9} = \frac{8}{9}. \end{aligned}$$

Part (c): We want to compute

$$\begin{aligned} P\{U \geq 0.3 | U \geq 0.1, U \geq 0.2\} &= \frac{P\{U \geq 0.1, U \geq 0.2, U \geq 0.3\}}{P\{U \geq 0.1, U \geq 0.2\}} \\ &= \frac{P\{U \geq 0.3\}}{P\{U \geq 0.2\}} = \frac{0.7}{0.8} = \frac{7}{8}. \end{aligned}$$

Part (d): We have $P(\text{winner}) = P\{U \geq 0.3\} = 0.7$.

Problem 8 (IQ scores)

We can transform all questions to those involving a standard normal. We have with S the random variable denoting the score of our IQ test taker. Then we have

$$\begin{aligned} P\{S \geq 125\} &= 1 - P\{S \geq 125\} \\ &= 1 - P\{\frac{S-100}{15} \leq \frac{150-100}{15}\} \\ &= 1 - P\{Z \leq 1.66\} = 1 - \Phi(1.66). \end{aligned}$$

Part (b): We desire to compute

$$\begin{aligned} P\{90 \leq S \leq 110\} &= P\left\{\frac{90-100}{15} \leq Z \leq \frac{110-100}{15}\right\} \\ &= P\{-0.66 \leq Z \leq 0.66\} = \Phi(0.66) - \Phi(-0.66) \end{aligned}$$

Problem 9

Let 1:00$-T$ be the time we leave from the house. Then if X is the random variable denoting how long it takes us to go to work, we arrive at work at the time 1:00$-T+X$. To guarantee with 95% probability that we arrive on time we must require that 1:00$-T+X \leq$1:00 or

$$-T + X \leq 0 \quad \text{or} \quad X \leq T,$$

with 95% probability. Thus we require T such that $P\{X \leq T\} = 0.95$ or

$$P\{\frac{X-40}{7} \leq \frac{T-40}{7}\} = 0.95 \Rightarrow \Phi(\frac{T-40}{7}) = 0.95\,,$$
$$\frac{T-40}{7} = \Phi^{-1}(0.95) \Rightarrow T = 40 + 7\Phi^{-1}(0.95)\,.$$

Problem 10 (the lifetime of automobile tires)

Part (a): We want to compute $P\{X \geq 40000\}$, which we do by converting to a standard normal. We find

$$\begin{aligned} P\{X \geq 40000\} &= P\left\{\frac{X-34000}{4000} \geq 1.5\right\} \\ &= 1 - P\{Z < 1.5\} = 1 - \Phi(1.5) = 0.0668\,. \end{aligned}$$

Part (b): We want to compute $P\{30000 \leq X \leq 35000\}$, which we do by converting to a standard normal. We find

$$\begin{aligned} P\{30000 \leq X \leq 35000\} &= P\left\{\frac{30000-34000}{4000} \leq Z \leq \frac{35000-34000}{4000}\right\} \\ &= P\{-1 \leq Z \leq 0.25\} \approx 0.4401\,. \end{aligned}$$

Part (c): We want to compute

$$P\{X \geq 40000|X \geq 30000\} = \frac{P\{X \geq 40000, X \geq 30000\}}{P\{X \geq 30000\}} = \frac{P\{X \geq 40000\}}{P\{X \geq 30000\}}\,.$$

We again do this by converting to a standard normal. We find

$$\begin{aligned} \frac{P\{X \geq 40000\}}{P\{X \geq 30000\}} &= \frac{P\left\{Z \geq \frac{40000-34000}{4000}\right\}}{P\left\{Z \geq \frac{30000-34000}{4000}\right\}} \\ &= \frac{1-\Phi(1.5)}{1-\Phi(-1.0)} = 0.0794\,. \end{aligned}$$

All of these calculations can be found in the Matlab file `chap_5_st_10.m`.

Problem 11 (the annual rainfall in Cleveland)

Part (a): Let X be the random variable denting the annual rainfall in Cleveland. Then we want to evaluate $P\{X \geq 44\}$. Which we can do by converting to a standard normal. We find that

$$\begin{aligned} P\{X \geq 44\} &= P\left\{\frac{X-40.2}{8.4} \geq \frac{44-40.2}{8.4}\right\} \\ &= 1 - \Phi(0.452) = 0.3255\,. \end{aligned}$$

Part (b): Following the assumptions stated for this problem lets begin by calculating $P(A_i)$ for $i = 1, 2, \ldots, 7$. Assuming independence each is equal to the value calculated in part (a) of this problem. Lets denote that common value by p. Then the random variable representing the number of years where the rainfall exceeds 44 inches (in a seven year time frame) is a Binomial random variable with parameters $(n, p) = (7, 0.3255)$. Thus the probability that three of the next seven years will have more than 44 inches of rain is given by

$$\binom{7}{3} p^3 (1-p)^4 = 0.2498\,.$$

These calculations are performed in the Matlab file `chap_5_st_11.m`.

Problem 14 (hazard rates)

Part (a): We have

$$P\{X > 2\} = 1 - P\{X \le 2\} = 1 - (1 - e^{-2^2}) = e^{-2^2} = e^{-4}\,.$$

Part (b): We find

$$\begin{aligned} P\{1 < X < 3\} &= P\{X \le 3\} - P\{X < 1\} \\ &= (1 - e^{-9}) - (1 - e^{-1}) = e^{-1} - e^{-9}\,. \end{aligned}$$

Part (c): The hazard rate function is defined as

$$\lambda(x) = \frac{f(x)}{1 - F(x)}\,.$$

Where f is the density function and F is the distribution function. We find for this problem that

$$f(t) = \frac{dF}{dx} = \frac{d}{dx}(1 - e^{-x^2}) = 2xe^{-x^2}\,.$$

so $\lambda(x)$ is given by

$$\lambda(x) = \frac{2xe^{-x^2}}{1 - (1 - e^{-x^2})} = 2x\,.$$

Part (d): The expectation is given by (using integration by parts to evaluate the first integral)

$$\begin{aligned} E[X] &= \int_0^\infty x f(x) dx = \int_0^\infty x(2xe^{-x^2}) dx \\ &= 2\left(\left.\frac{xe^{-x^2}}{-2}\right|_0^\infty + \frac{1}{2}\int_0^\infty e^{-x^2} dx \right) \\ &= \int_0^\infty e^{-x^2} dx\,. \end{aligned}$$

From the unit normalization of the standard Gaussian $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-s^2/2}ds = 1$ we can compute the value of the above integral. Using this expression we find that $\int_0^\infty e^{-x^2}dx = \sqrt{\pi}/2$ thus

$$E[X] = \frac{\sqrt{\pi}}{2}.$$

Part (d): The variance is given by $\text{Var}(X) = E[X^2] - E[X]^2$ so first computing the expectation of X^2 we have that

$$\begin{aligned} E[X^2] &= \int_0^\infty x^2 f(x)dx = \int_0^\infty x^2(2xe^{-x^2})dx \\ &= 2\left(\left.\frac{x^2 e^{-x^2}}{-2}\right|_0^\infty + \frac{1}{2}\int_0^\infty 2xe^{-x^2}dx \right) \\ &= 2\int_0^\infty xe^{-x^2}dx = 2\left(\left.\frac{e^{-x^2}}{-2}\right|_0^\infty \right) = 1\,. \end{aligned}$$

Thus

$$\text{Var}(X) = 1 - \frac{\pi}{4} = \frac{4-\pi}{4}\,.$$

X_1, X_2	$P\{X_1, X_2\}$
0, 0	$\left(\frac{8}{13}\right)\left(\frac{7}{12}\right)$
0, 1	$\left(\frac{8}{13}\right)\left(\frac{5}{12}\right)$
1, 0	$\left(\frac{5}{13}\right)\left(\frac{8}{12}\right)$
1, 1	$\left(\frac{5}{13}\right)\left(\frac{4}{12}\right)$

Table 23: The joint probability distribution for Problem 2 in Chapter 6

$X_1, X_2\ X_3$	$P\{X_1, X_2, X_3\}$
0, 0, 0	$\left(\frac{8}{13}\right)\left(\frac{7}{12}\right)\left(\frac{6}{11}\right)$
0, 0, 1	$\left(\frac{8}{13}\right)\left(\frac{7}{12}\right)\left(\frac{5}{11}\right)$
0, 1, 0	$\left(\frac{8}{13}\right)\left(\frac{5}{12}\right)\left(\frac{7}{11}\right)$
0, 1, 1	$\left(\frac{8}{13}\right)\left(\frac{5}{12}\right)\left(\frac{4}{11}\right)$
1, 0, 0	$\left(\frac{5}{13}\right)\left(\frac{8}{12}\right)\left(\frac{7}{11}\right)$
1, 0, 1	$\left(\frac{5}{13}\right)\left(\frac{8}{12}\right)\left(\frac{4}{11}\right)$
1, 1, 0	$\left(\frac{5}{13}\right)\left(\frac{4}{12}\right)\left(\frac{8}{11}\right)$
1, 1, 1	$\left(\frac{5}{13}\right)\left(\frac{4}{12}\right)\left(\frac{3}{11}\right)$

Table 24: The joint probability distribution for Problem 3 in Chapter 6

Chapter 6 (Jointly Distributed Random Variables)

Chapter 6: Problems

Problem 2

We have five white and eight balls. Let $X_i = 1$ if the i ball selected is white, and equals zero otherwise.

Part (a): We want to compute $P\{X_1, X_2\}$. Producing the Table 23 we have We can check that the given numbers do sum to one i.e. that $\sum_{X_1,X_2} P\{X_1, X_2\} = 1$. We have

$$\frac{1}{12 \cdot 13}(56 + 40 + 40 + 20) = \frac{156}{12 \cdot 13} = 1\,.$$

Part (b): We want compute $P\{X_1, X_2, X_3\}$. Enoumerating these probabilities we compute our results in Table 23

Problem 3

We begin by defining $Y_i = 1$ if the i white ball (from five) is selected at step i and zero otherwise.

Part (a): Computing the joint probability by conditioning on the first ball drawn, we have

$$\begin{aligned} P\{Y_1 = 0, Y_2 = 0\} &= P\{Y_2 = 0 | B_1 \,\text{is}\, W_2\} P\{B_1 \,\text{is}\, W_2\} \\ &+ P\{Y_2 = 0 | B_1 \,\text{is not}\, W_2\} P\{B_1 \,\text{is not}\, W_2\} \\ &= 1 \left(\frac{1}{13}\right) + \left(\frac{11}{12}\right)\left(\frac{12}{13}\right) . \end{aligned}$$

Now the other probabilties are computed in the same way. We find

$$P\{Y_1 = 0, Y_2 = 1\} = \left(\frac{12}{13}\right)\left(\frac{1}{12}\right) .$$

$$P\{Y_1 = 1, Y_2 = 0\} = \left(\frac{1}{13}\right) P\{Y_2 = 0\} = \left(\frac{1}{13}\right) \frac{11}{12} .$$

and

$$P\{Y_1 = 1, Y_2 = 1\} = \left(\frac{1}{13}\right)\left(\frac{1}{12}\right) .$$

Problem 10

Part (a): We find (performing several manipulations on one line to save space) that

$$\begin{aligned} P\{X < Y\} &= \int\int_\Omega f(x,y) dx dy = \int_{x=0}^\infty \int_{y=x}^\infty f(x,y) dy dx \\ &= \int_{x=0}^\infty \int_{y=x}^\infty e^{-(x+y)} dy dx = \int_{x=0}^\infty e^{-x} \left. \frac{e^{-y}}{(-1)} \right|_x^\infty dx \\ &= \int_{x=0}^\infty e^{-x}(e^{-x}) dx = \int_{x=0}^\infty e^{-2x} dx = \left. \frac{e^{-2x}}{(-2)} \right|_0^\infty = \frac{1}{2} . \end{aligned}$$

Part (b): We compute that

$$\begin{aligned} P\{X < a\} &= \int\int_\Omega f(x,y) dx dy = \int_{x=0}^a \int_{y=0}^\infty e^{-(x+y)} dy dx \\ &= \int_{x=0}^a e^{-x} \left. \frac{e^{-y}}{(-1)} \right|_0^\infty dx = \int_{x=0}^a e^{-x} dx = \left. \frac{e^{-x}}{(-1)} \right|_0^a = 1 - e^{-a} . \end{aligned}$$

Problem 11 (shopping for a television)

We have $p_{TV} = 0.45$, $p_{PT} = 0.15$, $p_B = 0.4$, so this problem looks like a multinomial distribution. If we let N_{TV}, N_{PT}, and N_B be the number of ordinary televisions, plasma televisions, and people browsing. Then we desire to compute

$$
\begin{aligned}
P\{N_{TV}=2, N_{PT}=1, N_B=2\} &= \binom{N}{N_{TV}, N_{PT}, N_B} p_{TV}^{N_{TV}} p_{PT}^{N_{PT}} p_B^{N_B} \\
&= \left(\frac{5!}{2!1!2!}\right)(0.45)^2(0.15)^1(0.4)^2 \\
&= 0.1458\,.
\end{aligned}
$$

Problem 12

We begin by recalling the example 2b from this from this chapter. There the book shows that if the total number of people entering a store is given by a Poisson random variable with rate λ and each person is of one type (male) with probability p and another type (female) with probability $q = 1 - p$, then the number of males and females entering are given by Poisson random variables with rates λp and $\lambda(1-p)$ respectively. In addition, and more surprisingly, the random variables X and Y are independent. Thus the desired probability for this problem is

$$
P\{X \le 3|Y=10\} = P\{X \le 3\} = \sum_{i=0}^{3} \frac{e^{-\lambda p}(\lambda p)^i}{i!} = e^{-\lambda p} \sum_{i=0}^{3} \frac{(\lambda p)^i}{i!}\,.
$$

If we assume that men and women are equally likely to enter so that $p = 1/2$ then the above becomes with $\lambda = 10$

$$
e^{-5} \sum_{i=0}^{3} \frac{5^i}{i!} = e^{-5}\left(1 + 5 + \frac{5^2}{2} + \frac{5^3}{3!}\right) = e^{-5}(39.33) = 0.265\,.
$$

Problem 13

Let X be the random variable denoting the arrival time of the men and Y be the random variable denoting the arrival time of the women. Then X is a uniform random variable with $X \in$ [12:15, 12:45], and Y is a uniform random variable with $Y \in$ [12:00, 1:00]. To simplify our calculations, we will let the time 12:30 be denote zero and measure time in minutes. Under this convention X is a uniform random variable taking values in $[-15, +15]$ and Y is a uniform random variable taking values in $[-30, +30]$. Then the question asks us to compute $P\{|X - Y| \le 5\}$. To compute this, condition on the case when the men arrives first or second. This expression then becomes

$$
\begin{aligned}
P\{|X-Y| \le 5\} &= P\{Y - X \le 5|X < Y\}P\{X < Y\} \\
&+ P\{X - Y \le 5|X > Y\}P\{X > Y\}\,.
\end{aligned}
$$

We will first compute $P\{X < Y\}$ which can easily be computed from the joint density $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ since the arrival of X and Y are independent. From our chosen coordinate system we have a valid domain for X and Y that looks like

WWX draw figure ...

In that figure form the line $X = Y$ we see clearly the domain where $X < Y$. Thus

$$P\{X < Y\} = \int\int_{\Omega_{X<Y}} f_{X,Y}(x,y)dxdy\,.$$

Since if we integrate over all of Ω we must obtain one, the joint density is given by

$$f_{X,Y}(x,y) = \frac{1}{30 \cdot 60} = \frac{1}{1800}\,.$$

To evaluate this integral we regonnize it as the area of the suggested trapezoid. Using the result from elementary geometry that area of a trapezoid equals the height times the average of its two bases. We have

$$\begin{aligned} P\{X < Y\} &= \frac{1}{1800}\int\int_{\Omega_{X<Y}} dxdy \\ &= \frac{1}{1800}\left(30 \cdot \left(\frac{45+15}{2}\right)\right) = \frac{1}{2}\,. \end{aligned}$$

Thus the other conditioning probability $P\{X > Y\}$ is easy to compute

$$P\{X > Y\} = 1 - P\{X < Y\} = \frac{1}{2}\,.$$

Now lets compute $P\{Y - X \leq 5 | X < Y\}$ which from the definition of conditional probability is given by

$$\frac{P\{Y - X \leq 5, X < Y\}}{P\{X < Y\}}\,.$$

The top probability is then given by the following integration region.

WWX: draw region

$$\begin{aligned} P\{Y - X \leq 5, X < Y\} &= \int_{x=-15}^{15}\int_{y=-x}^{y=x+5} f_{X,Y}(x,y)dydx \\ &= \int_{x=-15}^{15} \frac{1}{1800}(x+5+x)dx \\ &= \frac{1}{1800}\int_{x=-15}^{15}(2x+5)dx \\ &= \frac{1}{1800}\left(x^2 + 5x\right|_{-15}^{15} = \frac{1}{12}\,. \end{aligned}$$

Thus $P\{Y - X \leq 5 | X < Y\} = \frac{1/12}{1/2} = \frac{1}{6}$. Now additional probabilities can be computed in using these same methods but a direct solution to the problem is to compute $P\{|X - Y| \leq 5\}$ directly. For example, to compute $P\{|X - Y| \leq 5\}$ we integrate over the region

WWX put plot here!!!

and is represented by the following integral

$$\begin{aligned}\int\int_{\Omega_{\{|X-Y|\leq 5\}}} f_{X,Y}(x,y)dxdy &= \frac{1}{1800}\int_{x=-15}^{+15}\int_{y=x-5}^{x+5} dydx \\ &= \frac{1}{1800}\int_{x=-15}^{15}(x+5-(x-5))dx = \frac{1}{6}\,.\end{aligned}$$

Now the probability that the man arrives first is given by $P\{X \leq Y\}$ and was computed before to be 1/2.

Problem 14

Let X denote the random variable providing the location of the accident along the road. The according to the problem specification let X be uniformly distributed between $[0, L]$. Let Y denote the random variable the location of the ambulance. Then we define $D = |X - Y|$ the random variable representing the distance between the accident and the ambulance. We want to compute

$$P\{D \leq d\} = \int\int_{X,Y\in\Omega} f(x,y)dxdy\,,$$

with Ω the set of points where $|X - Y| \leq d$. In the X, Y plane the set $|x - y| \leq d$ can be graphically represented by

WWX put drawing!!!

Using the diagram we see how to analytically integrate over the domain Ω. Specifically we have

$$\begin{aligned}P\{D \leq d\} &= \int_{x=0}^{d}\int_{y=0}^{x+d} f(x,y)dydx + \int_{x=d}^{L-d}\int_{y=x-d}^{x+d} f(x,y)dydx \\ &+ \int_{x=L-d}^{L}\int_{x-d}^{L} f(x,y)dydx\,,\end{aligned}$$

Since X and Y are independent $f_{X,Y}(x,y) = \frac{1}{L^2}$ so the expression above becomes

$$\begin{aligned}P\{D \leq d\} &= \frac{1}{L^2}\int_0^d (x+d)dx + \frac{1}{L^2}\int_d^{L-d}(x+d-(x-d))dx \\ &+ \frac{1}{L^2}\int_{L-d}^{L}(L-x+d)dx \\ &= \frac{1}{L^2}\left(\frac{x^2}{2}+dx\right|_0^d + \frac{1}{L^2}(2d(L-2d)) + \frac{1}{L^2}\left((L+d)x - \frac{x^2}{2}\right|_{L-d}^{L} \\ &= \frac{(2L-d)d}{L^2}\,.\end{aligned}$$

Then $f_D(d) = \frac{dF_D(d)}{d(d)}$ (meaning we take the derivative with resepct to the variable d) and we find $f_D(d) = \frac{2(L-d)}{L^2}$.

Problem 32

Part (a): Since each week the gross sales is a draw from an independent normal random variable the sum of n of these normal random variables is another normal with mean the sum of the n means and variance the sum of the n variances. For the case given in this problem we have two normals over the two weeks and the mean gross sales is

$$\mu = 2200 + 2200 = 4400 .$$

and variance is

$$\sigma^2 = 230^2 + 230^2 = 105800 .$$

The we desire to compute

$$\begin{aligned} P\{\text{Sales} \geq 5000\} &= 1 - P\{\text{Sales} \leq 5500\} \\ &= 1 - P\left\{\frac{S - 4400}{325.2} \leq \frac{5000 - 4400}{325.2}\right\} \\ &= 1 - \Phi(1.8446) = 0.0325 . \end{aligned}$$

Part (b): In this part of the problem I'll compute the probability that the weekly sales will exceed 2000 by treating this probability as the probability of success under a Binomial random variable model. I'll use the Binomial mass function to compute the probability that our weekly sales exceeds 2000 in two of the next three weeks. Defining p to to be

$$\begin{aligned} p &= P\{\text{Sales} \geq 2000\} \\ &= 1 - P\left\{\frac{\text{Sales} - 2200}{230} \leq \frac{2000 - 2200}{230}\right\} . \end{aligned}$$

Then with this probability we are after is

$$P_s = \binom{3}{2} p^2 (1-p)^1 + \binom{3}{3} p^3 = 0.9033 .$$

Problem 44

This is a problem in so called Bayesian inference. By this what I mean is that given information about the number of accidents that have occurred, we want to find the *density* (given this number of accidents) of the accident rate λ. Before observing the number of accidents in a year, the unknown λ is governed by a gamma distribution with parameters s and α. Specifically,

$$f(\lambda) = \begin{cases} \frac{s e^{-s\lambda}(s\lambda)^{\alpha-1}}{\Gamma(\alpha)} & \lambda \geq 0 \\ 0 & \lambda < 0 \end{cases} .$$

We wish to evaluate $p(\lambda|N = n)$ where $p(\lambda|N = n)$ is the probability that λ takes a given value after the observation that n accident occurred last year. From Bayes' rule we have that

$$P(\lambda|N = n) = \frac{p(N = n|\lambda) f(\lambda)}{\int_\Lambda p(N = n|\lambda) f(\lambda) d\lambda} \propto p(N = n|\lambda) f(\lambda) .$$

Now $p(N = n|\lambda)$ is a Poisson random variable with mean λ and therefore has a density function by

$$p(N = n|\lambda) = \frac{e^{-\lambda}\lambda^n}{n!} = \frac{e^{-\lambda}\lambda^n}{\Gamma(n+1)} \quad n = 0, 1, 2, \cdots$$

so the above expression above becomes

$$p(\lambda|N = n) \propto \left(\frac{e^{-\lambda}\lambda^n}{\Gamma(n+1)}\right)\left(\frac{se^{-s\lambda}(s\lambda)^{\alpha-1}}{\Gamma(\alpha)}\right) = \frac{e^{-\lambda(1+s)}s^{\alpha}\lambda^{n+\alpha+1}}{\Gamma(\alpha)\Gamma(n+1)},$$

which is proportional to $e^{-(1+s)\lambda}\lambda^{n+\alpha+1}$. The density that is proportional to an expression like this is another gamma distribution with parameters $s+1$ and $n+\alpha$. This is seen from the functional form of the gamma distribution (presented above) and from this we can easily compute the required normalizing factor. Specifically we have

$$p(\lambda|N = n) = \begin{cases} \frac{(s+1)e^{-(s+1)\lambda}((s+1)\lambda)^{n+\alpha-1}}{\Gamma(n+\alpha)} & \lambda \geq 0 \\ 0 & \lambda < 0 \end{cases}.$$

This is the conditional density of the accident parameter λ.

To determine the expected number of accidents the *following* year (denoted N_2) we will first compute the probability that we observe m accidents in that year i.e. $P\{N_2 = m\}$. This is given by conditioning on λ as follows

$$P\{N_2 = m\} = \int_{\Lambda} P\{N_2 = m|\lambda\}p(\lambda)d\lambda\,,$$

where in this expression everything is implicitly conditioned on $N_1 = n$ i.e. that in the first year we observed n accidents. The above integral becomes

$$\begin{aligned} P\{N_2 = m\} &= \int_{\lambda=0}^{\infty}\left(\frac{e^{-\lambda}\lambda^m}{m!}\right)\frac{(s+1)^{n+\alpha}}{\Gamma(n+\alpha)}e^{-(s+1)\lambda}\lambda^{n+\alpha-1}d\lambda \\ &= \frac{(s+1)^{n+\alpha}}{m!\Gamma(n+\alpha)}\int_0^{\infty} e^{-(s+2)\lambda}\lambda^{n+m+\alpha-1}d\lambda\,. \end{aligned}$$

To evaluate this integral let $v = (s+2)\lambda$ then $dv = (s+2)d\lambda$ and the above becomes

$$\begin{aligned} P\{N_2 = m\} &= \frac{(s+1)^{n+\alpha}}{m!\Gamma(n+\alpha)}\int_0^{\infty} e^{-v}\frac{v^{n+m+\alpha-1}}{(s+2)^{n+m+\alpha-1+1}}dv \\ &= \frac{(s+1)^{n+\alpha}}{m!\Gamma(n+\alpha)(s+2)^{n+m+\alpha}}\Gamma(n+m+\alpha) \\ &= \frac{\Gamma(n+m+\alpha)}{\Gamma(m+1)\Gamma(n+\alpha)}\frac{(s+1)^{n+\alpha}}{(s+2)^{n+m+\alpha}} \quad \text{for} \quad m = 0, 1, 2, 3, \cdots \end{aligned}$$

Now a generalization expression for the binomial coefficient $\binom{n}{k}$ is given by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)},$$

which from the definition in terms of the gamma function is valid for non integer n and k. Thus

$$\frac{\Gamma(n+m+\alpha)}{\Gamma(m+1)\Gamma(n+\alpha)} = \binom{n+m+\alpha-1}{m} = \binom{n+m+\alpha-1}{n+\alpha-1},$$

so the distribution for $P\{N_2 = m\}$ becomes (using the second expression for the ratio of gamma functions)

$$P\{N_2 = m\} = \binom{n+m+\alpha-1}{n+\alpha-1} \frac{(s+1)^{n+\alpha}}{(s+2)^{n+m+\alpha}} \quad \text{for} \quad m = 0, 1, 2, \cdots$$

We can shift this index N_2 to be "offset" by $n+\alpha$ to make this look like a negative binomial random variable. To do this define $\tilde{N}_2 = N_2 + (n+\alpha)$, then since the range of N_2 is from $0, 1, 2 \cdots$ the range of $\tilde{N}_2$ is from $n+\alpha, n+\alpha+1, \cdots$. Thus

$$\begin{aligned}
P\{\tilde{N}_2 &= \tilde{m}\} = \binom{\tilde{m}-1}{r-1} \frac{(s+1)^r}{(s+2)^{\tilde{m}}} \\
&= \binom{\tilde{m}-1}{r-1} \left(\frac{s+1}{s+2}\right)^r \frac{1}{(s+2)^{\tilde{m}-r}},
\end{aligned}$$

Where we have defined $r = n+\alpha$ and $\tilde{m}$ take values in the range $r, r+1, r+2, \cdots$. Then further defining $p = \frac{s+1}{s+2}$ so that

$$1 - p = \frac{s+2}{s+2} - \frac{s+1}{s+2} = \frac{1}{s+2},$$

we have that the above is given by

$$P\{\tilde{N}_2 = \tilde{m}\} = \binom{\tilde{m}-1}{r-1} p^r (1-p)^{\tilde{m}-r} \quad \text{for} \quad \tilde{m} = r, r+1, r+2, \cdots$$

From this expression and Example 8f from the book we know that

$$E[\tilde{N}_2] = \frac{r}{p} = \left(\frac{s+2}{s+1}\right)(n+\alpha).$$

The expectation of N_2, the expected number of accidents is then given by

$$\begin{aligned}
E[N_2] &= E[\tilde{N}_2] - (n+\alpha) \\
&= \left(\frac{s+2}{s+1}\right)(n+\alpha) - (n+\alpha) \\
&= \left(\frac{s+2}{s+1} - 1\right)(n+\alpha) = \frac{n+\alpha}{s+1},
\end{aligned}$$

which is requested expression.

Problem 53

From the theorem on the transformation of coordinates for joint probability density functions we have that

$$f_{X,Y}(x,y) = f_{Z,U}(z,u)|J(z,u)|^{-1}.$$

Now in this case the Jacobian expression becomes

$$
\begin{aligned}
J(z,u) &= \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial u} \end{vmatrix} = \begin{vmatrix} \sqrt{2}z^{-1/2}(1/2)\cos(u) & -\sqrt{2}z^{1/2}\sin(u) \\ \sqrt{2}z^{-1/2}(1/2)\sin(u) & +\sqrt{2}z^{1/2}\cos(u) \end{vmatrix} \\
&= \frac{2}{2}\cos^2(u) + \frac{2}{2}\sin^2(u) = 1\,.
\end{aligned}
$$

We also have that $f_{X,Y}(x,y) = f_Z(z) f_U(u)$ by independence of the random variable Z and U. Since Z is uniform $[0, 2\pi]$ and U is exponential with a rate one we get for $f_{X,Y}(x,y)$ the following

$$
f_{X,Y}(x,y) = \left(\frac{1}{2\pi}\right) 1 e^{-z}\,.
$$

Now since

$$
\frac{X}{\sqrt{2Z}} = \cos(U) \quad \text{and} \quad \frac{Y}{\sqrt{2Z}} = \sin(U)\,,
$$

squaring both sides and adding we obtain

$$
\frac{X^2}{2Z} + \frac{Y^2}{2Z} = 1\,,
$$

or $Z = \frac{1}{2}(X^2 + Y^2)$, thus we have

$$
f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\,,
$$

which is the product of two probability density functions for standard normal random variables showing that X and Y are independent and normal.

Chapter 6: Theoretical Exercises

Problem 19

We are asked to compute $p(W|N=n)$, using Bayes' rule this can be expressed as

$$
\frac{P\{N=n|W\}P(W)}{\int_W P\{N=n|W\}P(W)dw}\,.
$$

Now ignoring the normalizing term the above is proportional to

$$
\frac{e^{-w}w^n}{\Gamma(n+1)} \frac{\beta e^{-\beta w}(\beta w)^{t-1}}{\Gamma(t)}\,,
$$

which by factoring only the terms that depend on w is proportional to

$$
e^{-(\beta+1)w} w^{n+t-1}\,,
$$

which is the functional form (in w) for a gamma distribution with parameters $n+t$ and $\beta+1$. Thus

$$
p(W|N=n) = \frac{(\beta+1)e^{-(\beta+1)w}((\beta+1)w)^{n+t-1}}{\Gamma(n+t)}\,.
$$

Problem 22

We desire to compute $P\{[X] = n, X - [X] \leq x\}$ for $n = 0, 1, 2, \cdots$ and $0 \leq x \leq 1$. This is given by the following integral

$$\int_{\xi=n}^{n+x} \lambda e^{-\lambda\xi} d\xi = \lambda \left. \frac{e^{-\lambda\xi}}{-\lambda} \right|_n^{n+x} = e^{-\lambda n} - e^{-\lambda(n+x)} .$$

To see if $[x]$ and $X - [X]$ are independent, consider if the joint distribution function is the product of the two marginalized distributions. The joint probability density function of this random variable is given by the derivative of this with respect to x i.e.

$$f_{N,X}(n, x) = \frac{d}{dx} P\{[X] = n, X - [X] \leq x\} = \lambda e^{-\lambda(n+x)} ,$$

computing the marginal distribution we first compute

$$\begin{aligned} P\{[X] = n\} &= \int_{x=0}^{1} f_{N,X}(n, x) dx = \lambda \int_0^1 e^{-\lambda(n+x)} dx \\ &= \left. \frac{-\lambda}{\lambda} e^{-\lambda(n+x)} \right|_0^1 \\ &= -(e^{-\lambda(n+1)} - e^{-\lambda n}) = e^{-\lambda n} - e^{-\lambda(n+1)} . \end{aligned}$$

Also

$$\begin{aligned} p(X - [X] = x) &= \sum_{n=0}^{\infty} \lambda e^{-\lambda(n+x)} \\ &= \lambda e^{-\lambda x} \sum_{n=0}^{\infty} e^{-\lambda n} \\ &= \lambda e^{-\lambda x} \sum_{n=0}^{\infty} (e^{-\lambda})^n \\ &= \lambda e^{-\lambda x} \frac{1}{1 - e^{-\lambda}} = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda}} , \end{aligned}$$

so if $f_{N,X}(n, x) = P\{[X] = n\} p(X - [X] = x)$, then our elements are independent. Computing the right hand side of this expression we have

$$(e^{-\lambda n} - e^{-\lambda(n+1)}) \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda}} = \lambda e^{-\lambda(x+n)} ,$$

which does equal $f_{N,X}(n, x)$ so these two random variables are independent.

Problem 23

Part (a): Given X_i for $i = 1, 2, \cdots, n$ with common distribution function $F(x)$ define $Y = \max(X_i)$ for $i = 1, 2, \cdots, n$. Then Y is called the nth order statistic and is often written $X_{(n)}$. To compute the cumulative distribution function for $X_{(n)}$ i.e. $P\{X_{(n)} \leq x\}$ we can either use the result in the book

$$P\{X_{(n)} \leq x\} = \sum_{k=j}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) F(y)^k (1 - F(y))^{n-k} \,,$$

for $j = n$ which gives $F_{X_{(n)}} = F(y)^n$, or we can reason more simply as follows. For the random variable $X_{(n)}$ to be less than x, each draw X_i must be less than x. Each draw is less than x with probability $F(x)$ and as this must happen n times, the probability that $X_{(n)}$ is less than x is $F(x)^n$, verifying the above.

Part (b): In this case define $Z = \min(X_i)$, then Z is another order statistic and this time corresponds to $X_{(1)}$. Using the distribution function from before we have

$$\begin{aligned} F_{X_{(1)}} &= \sum_{k=1}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) F(y)^k (1 - F(y))^{n-k} \\ &= \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) F(y)^k (1 - F(y))^{n-k} - \left(\begin{array}{c} n \\ 0 \end{array} \right) F(y)^0 (1 - F(y))^n \\ &= 1 - (1 - F(y))^n \,. \end{aligned}$$

Or we may reason as follows. The distribution function

$$F_Z(x) = P\{Z \leq x\} = P\{\min_i(X_i) \leq x\} = 1 - P\{\min(X_i) \geq x\} \,.$$

This last probability is the intersection of n events i.e. each X_i that is drawn must be drawn greater than the value of x. Each event of this type happens with probability $1 - F(x)$. Giving that

$$P\{\min_i(X_i) \leq x\} = 1 - (1 - F(x))^n \,.$$

Problem 27 (the sum of a uniform and an exponential)

Part (a): Since X and Y are independent, the density of the random variable $Z = X + Y$ is given by the convolution of the density function for X and Y. For example

$$f_Z(a) = \int_{-\infty}^{\infty} f_X(a - y) f_Y(y) dy = \int_{-\infty}^{\infty} f_X(y) f_Y(a - y) dy \,.$$

Since X is uniform we have that

$$f_X(x) = \left\{ \begin{array}{cc} 1 & 0 \leq x \leq 1 \\ 0 & \text{else} \end{array} \right. \,,$$

and since Y is exponential we have that

$$f_Y(y) = \left\{ \begin{array}{cc} \lambda e^{-\lambda y} & y \geq 0 \\ 0 & \text{else} \end{array} \right.$$

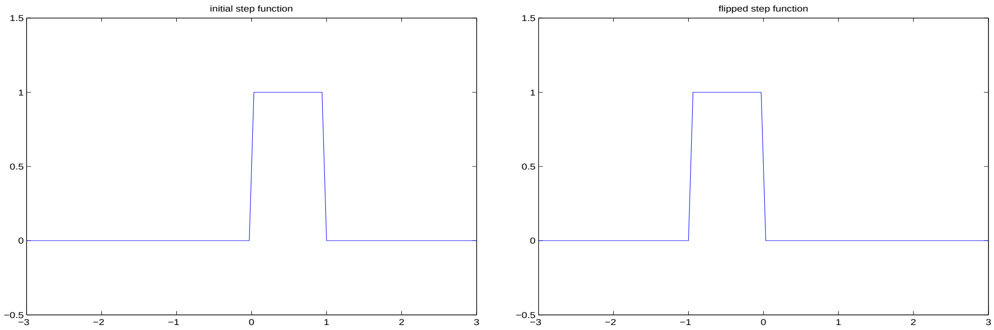

Figure 3: **Left:** The initial probability density function for X or $f_X(x)$ (a step function). **Right:** This function flipped or $f_X(-x)$.

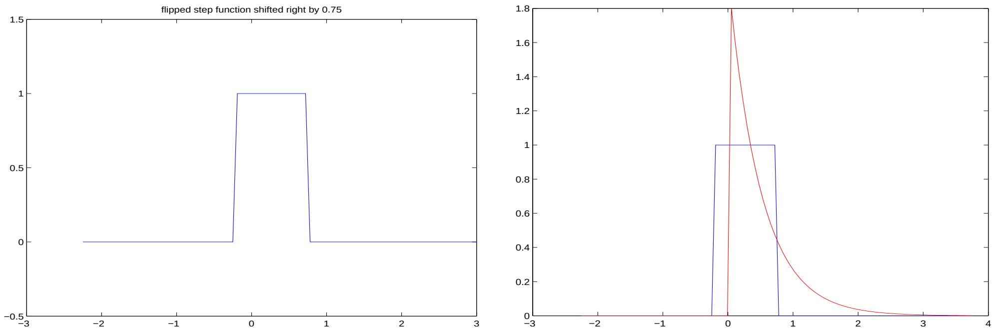

Figure 4: **Left:** The initial probability density function for X, flipped and shifted by $a = 3/4$ to the right or $f_X(-(x-a))$. **Right:** The flipped and shifted function plotted together with $f_Y(y)$ allowing visualizations of function overlap as a is varied.

To evaluate the above convolution we might as well select a formulation that is simple to evaluate. I'll pick the *first* formulation since it is easy to shift to the uniform distribution to produce $f_X(a-y)$. Now since $f_X(x)$ looks like the plot given in Figure 3 (left) we see that $f_X(-x)$ then looks like Figure 3 (right). Inserting a right shift by a we have $f_X(-(x-a)) = f_X(a-x)$, and this function looks like that shown in Figure 4 (left). Thus we can now evaluate the distribution function for $f_Z(a)$, we find that

$$f_Z(a) = \int_{-1+a}^{a} 1 f_Y(y) dy \,.$$

Now since $f_Y(y)$ is zero when y is negative, to further evaluate this we evaluate it for some specific a's. One easy case is if $a < 0$ then $f_Z(a) = 0$. If $a > 0$ but the lower limit of integration is negative, that is $-1 + a < 0$, i.e. $0 < a < 1$, then we have

$$\begin{aligned} f_Z(a) &= \int_{-1+a}^{a} f_Y(y) dy = \int_0^a \lambda e^{-\lambda y} dy = \lambda \int_0^a e^{-\lambda y} dy \\ &= \left. \frac{\lambda e^{-\lambda y}}{-\lambda} \right|_0^a = -(e^{-\lambda a} - 1) = 1 - e^{-\lambda a} \,. \end{aligned}$$

If $a > 1$ then the integral for $f_Z(a)$ then is

$$\begin{aligned} f_Z(a) &= \int_{-1+a}^{a} f_Y(y) dy = \int_{-1+a}^{a} \lambda e^{-\lambda y} dy \\ &= \left. \frac{\lambda e^{-\lambda y}}{-\lambda} \right|_{-1+a}^{a} \\ &= \frac{-\lambda}{\lambda} (e^{-\lambda a} - e^{-\lambda(-1+a)}) \\ &= e^{-\lambda(a-1)} - e^{-\lambda a} \,. \end{aligned}$$

In summary then

$$f_Z(a) = \begin{cases} 0 & a < 0 \\ 1 - e^{-\lambda a} & 0 < a < 1 \\ e^{-\lambda(a-1)} - e^{-\lambda a} & a > 1 \end{cases}$$

In the MATLAB function `chap_6_prob_27.m` we have code to duplicate the above figures.

Part (b): If $Z = \frac{X}{Y}$ then to compute the distribution function for Z is

$$\begin{aligned} F_Z(a) &= P\{Z \le a\} \\ &= P\{\frac{X}{Y} \le a\} \\ &= P\{X \le aY\} \\ &= \int\int_{X \le aY} f(x,y) dx dy \\ &= \int\int_{X \le aY} f(x) f(y) dx dy \end{aligned}$$

where we get the last equality from the independence of X and Y. The above integral can be evaluated by letting x range over 0 to 1, while y ranges over $\frac{1}{a}x$ to $+\infty$. With these limits

$\frac{x}{a} \leq y \leq \infty$ or $x \leq ay \leq \infty$. So the integral above becomes

$$\begin{aligned}
\int_{x=0}^{1}\int_{y=x/a}^{\infty} f(x)f(y)dydx &= \int_{0}^{1}\int_{x/a}^{\infty} \lambda e^{-\lambda y} dydx \\
&= \int_{x=0}^{1} \frac{-\lambda}{\lambda} e^{-\lambda y}\Big|_{x/a}^{\infty} = \int_{0}^{1} -(0 - e^{-\lambda x/a})dx \\
&= \int_{0}^{1} e^{-\lambda x/a} dx = \frac{e^{-\lambda x/a}}{\left(\frac{-\lambda}{a}\right)}\Bigg|_{0}^{1} \\
&= -\frac{a}{\lambda}(e^{-\lambda/a} - 1) = \frac{a}{\lambda}(1 - e^{-\lambda/a}) \,.
\end{aligned}$$

Problem 33 (the P.D.F. of the ratio of normals is a Cauchy distribution)

As stated in the problem, let X_1 and X_2 be distributed as standard normal random variables (i.e. they have mean 0 and variance 1). Then we want the distribution of the variable X_1/X_2. To this end define the random variables U and V as $U = X_1/X_2$ and $V = X_2$. The distribution function of U is then what we are after. From the definition of U and V in terms of X_1 and X_2 we see that $X_1 = UV$ and $X_2 = V$. To solve this problem we will derive the joint distribution function for U and V and then marginalize out V giving the distribution function for U, alone. Now from Theorem $2-4$ on page 45 of Schaums probability and statistics outline the distribution of the joint random variable (U, V), in term of the joint random variable (X_1, X_2) is given by

$$g(u, v) = f(x_1, x_2)\left|\frac{\partial(x_1, x_2)}{\partial(u, v)}\right| \,.$$

Now

$$\left|\frac{\partial(x_1, x_2)}{\partial(u, v)}\right| = \left|\begin{array}{cc} v & u \\ 0 & 1 \end{array}\right| = |v| \,,$$

so that

$$g(u, v) = f(x_1, x_2)|v| = p(x_1)\,p(x_2)|x_2| \,,$$

as $f(x_1, x_2) = p(x_1)p(x_2)$ since X_1 and X_2 are assumed independent. Now using the fact that the distribution of X_1 and X_2 are standard normals we get

$$g(u, v) = \frac{1}{2\pi}\,\exp(-\frac{1}{2}\,(u\,v)^2)\,\exp(-\frac{1}{2}\,v^2)\,|v| \,.$$

Marginalizing out the variable V we get

$$g(u) = \int_{-\infty}^{\infty} g(u, v)dv = \frac{1}{\pi}\int_{0}^{\infty} v\,e^{-\frac{1}{2}(1+u^2)v^2} dv$$

To evaluate this integral let $\eta = \sqrt{\frac{1+u^2}{2}}\,v$, and after performing the integration we then find that

$$g(u) = \frac{1}{\pi}\frac{1}{1+u^2} \,.$$

Which is the distribution function for a Cauchy random variable.

Chapter 7 (Properties of Expectations)

Chapter 7: Problems

Problem 1 (expected winnings with coins and dice)

If we roll a heads then we win twice the digits on the die roll. If we roll a tail then we win 1/2 the digit on the die. Now we have a 1/2 chance of getting a head (or a tail) and a 1/6 chance of getting any individual number on the die. Thus the expected winnings are given by

$$\frac{1}{2}\cdot\frac{1}{6}\left(\frac{1}{2}\cdot 1\right)+\frac{1}{2}\cdot\frac{1}{6}\left(\frac{1}{2}\cdot 2\right)+\cdots+\frac{1}{2}\cdot\frac{1}{6}\,(2\cdot 1)+\frac{1}{2}\cdot\frac{1}{6}\,(2\cdot 2)+\cdots$$

or factoring out the 1/2 and the 1/6 we obtain

$$\frac{1}{2}\cdot\frac{1}{6}\left(\frac{1}{2}+\frac{2}{2}+\frac{3}{2}+\frac{4}{2}+\frac{5}{2}+\frac{6}{2}+2+2\cdot 2+2\cdot 3+2\cdot 4+2\cdot 5+2\cdot 6\right)$$

which equals

$$\frac{105}{24}\,.$$

Problem 2 (the game of Clue)

Part (a): We have six choices for a suspect, six choices for a weapon and nine choices for a room giving in total $6\cdot 6\cdot 9=324$ possible combinations.

Part (b): Now let the random variables S, W, and R be the number of suspects, weapons, and rooms that the player receives and let X be the number of solutions possible after observing S, W, and R. Then X is given by

$$X=(6-S)(6-W)(9-R)\,.$$

Part (c): Now we must have

$$S+W+R=3 \quad \text{with} \quad 0\le S\le 3\,, \quad 0\le W\le 3\,, \quad 0\le R\le 3$$

Each specification of these three numbers (S, W, R) occurs with a uniform probability given by

$$\frac{1}{\binom{3+3-1}{3-1}}=\frac{1}{\binom{5}{2}}=\frac{1}{10}\,,$$

using the results given in Chapter 1. Thus the expectation of X is given by

$$\begin{aligned}
E[X] &= \frac{1}{10}\sum_S\sum_W\sum_R(6-S)(6-W)(9-R) \\
&= \frac{1}{10}\sum_{s=0}^{3}(6-s)\sum_{w=0}^{3}(6-w)\sum_{r=0}^{3}(9-r) \\
&= \frac{1}{10}\left[6\sum_{W+R=3}(6-W)(9-R)+5\sum_{W+R=2}(6-W)(9-R)\right. \\
&+ \left. 4\sum_{W+R=1}(6-W)(9-R)+3\sum_{W+R=0}(6-W)(9-R)\right] \\
&= 190.4\,.
\end{aligned}$$

Problem 3 (the expectation of $|X-Y|^\alpha$)

We have by definition

$$\begin{aligned}
E[|X-Y|^\alpha] &= \int\int|x-y|^\alpha f_{X,Y}(x,y)dxdy \\
&= \int\int|x-y|^\alpha dxdy\,.
\end{aligned}$$

Since the area of region of the $x-y$ plane where $y>x$ is equal to the area of the $x-y$ plane where $y<x$, we can compute the above integral by doubling the integration domain $y<x$ to give

$$\begin{aligned}
2\int_{x=0}^{1}\int_{y=0}^{x}(x-y)^\alpha dydx &= 2\int_{x=0}^{1}(-1)\frac{(x-y)^{\alpha+1}}{\alpha+1}|_0^x dx \\
&= \frac{2}{\alpha+1}\int_{x=0}^{1}x^{\alpha+1}dx \\
&= \frac{2}{\alpha+1}\frac{x^{\alpha+2}}{\alpha+2}|_0^1 \\
&= \frac{2}{(\alpha+1)(\alpha+2)}\,.
\end{aligned}$$

Problem 4

If X and Y are independent and uniform, then

$$P\{X=x,Y=y\}=P\{X=x\}P\{Y=y\}=\left(\frac{1}{m}\right)^2.$$

Then

$$E[|X-Y|]=\sum_X\sum_Y|x-y|\left(\frac{1}{m}\right)^2.$$

We can sum over the set $y < x$ and double the summation range to evaluate this expectation. Doing this we find that the above is equal to

$$\frac{2}{m^2} \sum_{x=1}^{m} \sum_{y=1}^{x-1} (x - y) \,.$$

Note that the second sum above goes to $x - 1$ since when $y = x$ the term in the above vanishes. Now evaluating the inner summation we have that

$$\sum_{y=1}^{x-1} (x - y) = (x - 1) + (x - 2) + (x - 3) + \cdots + 2 + 1 = \sum_{y=1}^{x-1} y = \frac{x(x-1)}{2} \,.$$

So the above double sum then becomes

$$\frac{2}{m^2} \sum_{x=1}^{m} \frac{x(x-1)}{2} = \frac{1}{m^2} \left(\sum_{x=1}^{m} x^2 - \sum_{x=1}^{m} x \right) \,.$$

Now remembering or looking up in tables we have that

$$\begin{aligned} \sum_{x=1}^{m} x &= \frac{m(m+1)}{2} \\ \sum_{x=1}^{m} x^2 &= \frac{m(m+1)(2m+1)}{6} \,, \end{aligned}$$

so that our expectation then becomes

$$\begin{aligned} E[|X - Y|] &= \frac{1}{m^2} \left(\frac{m(m+1)(2m+1)}{6} - \frac{m(m+1)}{2} \right) \\ &= \frac{(m+1)(m-1)}{3m} \,, \end{aligned}$$

as requested.

Problem 5

Then the distance traveled by the ambulance is given by $D = |X| + |Y|$ so our expected value is given by

$$\begin{aligned} E[D] &= \int_{-1.5}^{1.5}\int_{-1.5}^{1.5}(|x|+|y|)\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)dxdy \\ &= \frac{1}{9}\int_{-1.5}^{1.5}\int_{-1.5}^{1.5}(|x|+|y|)dxdy \\ &= \frac{1}{9}\left(3\int_{-1.5}^{1.5}|x|dx + 3\int_{-1.5}^{1.5}|y|dy\right) \\ &= \frac{1}{3}\left(2\int_{0}^{1.5}xdx + 2\int_{0}^{1.5}ydy\right) \\ &= \frac{2}{3}\left(\left.\frac{x^2}{2}\right|_0^{1.5} + \left.\frac{y^2}{2}\right|_0^{1.5}\right) \\ &= \frac{1}{3}\left(\frac{9}{4}+\frac{9}{4}\right) = \frac{3}{2}. \end{aligned}$$

Problem 6

Let X_i be the random variable that denotes the face on the die roll i. Then the total number of die is the random variable

$$Z = \sum_{i=1}^{10} X_i .$$

Then the expectation of Z is given by

$$E[Z] = \sum_{i=1}^{10} E[X_i] .$$

Now since X_i are uniformly distributed discrete random variables between 1, 2, 3, 4, 5, and 6, we have

$$E[X_i] = \frac{1}{6}(1+2+3+4+5+6) = 3.5 .$$

So we then have that $E[Z] = 10\left(\frac{7}{2}\right) = 35$.

Problem 7

Part (a): We want to know the expected number of objects chosen by both A and B. A will select three objects, then when B selects his three objects it is like the problem where we have three special items from ten and the number selected that are special (in this case selected by person A) is given by a hypergeomeric random variable with parameters $N = 10$, $m = 3$, and $n = 3$. So

$$P\{X = i\} = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \quad i = 0, 1, 2, \cdots, m$$

This distribution has an expected value given by

$$E[X] = \frac{nm}{N} = \frac{9}{10} = 0.9\,.$$

Part (b): After A has selected his three, then B will select three from ten where three of these ten are "special" i.e. the ones that are picked by A. Let X be the random variable that specifies the number of A objects that B selects. Then X is a hypergeometric random variable with parameters $N = 10$, $m = 3$, and $n = 3$. So

$$P\{X = i\} = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \quad i = 0, 1, 2, \cdots, n$$

Then if B selects X of A's selections then $(N-3)-(3-X)$ are *not* chosen by either A or B. Lets check this. If $X = 0$, meaning that B selects none of A picks we have that $N - 6$ items are not chosen by A or B. If $X = 3$, then all three of A's picks are selected by B and the number of unselected items is $N - 3$, so yes our formula seems true. Thus the expectation of the number of unselected items is

$$E[(N-3)-(3-X)] = N - 6 + E[X] = N - 6 + \frac{nm}{6} = 10 - 6 + \frac{9}{10} = 4.9\,.$$

Part (c): As in Part (b) if the number chosen by both A and B is a hypergeometric random variable with parameters $N = 10$, $m = 3$, and $n = 3$. Then the number of elements chosen by only one person is

$$(3 - X) + (3 - X)\,.$$

Where the first term is the the number of A's selections not selected by B and the second term is the number of B selections not selected by A. Thus our random variable is $6 - 2X$, so the expectation is given by

$$6 - 2E[X] = 6 - 2\left(\frac{mn}{N}\right) = 6 - 2\left(\frac{9}{10}\right) = \frac{21}{5} = 4.2\,.$$

Problem 8

Then a new table is started if and only if person i does not have any friends in the room. There will be no friends for person i with probability $P\{N = 0\} = q^{i-1} = (1-p)^{i-1}$.

Following the hint, we can let X_i be an indicator random variable denoting whether or not the ith person starts a new table. Then the total number of new tables is

$$T = \sum_{i=1}^{N} X_i \,.$$

So the expectation of the number of new tables is given by

$$E[T] = \sum_{i=1}^{N} E[X_i] = \sum_{i=1}^{N} P(X_i) \,,$$

Where the probability $P(X_i)$ is computed above so

$$\begin{aligned} E[T] &= \sum_{i=1}^{N} (1-p)^{i-1} \\ &= \sum_{i=1}^{N-1} (1-p)^{i} \\ &= \frac{(1-p)^N - (1-p)^0}{1-p-1} \\ &= \frac{1-(1-p)^N}{p} \,. \end{aligned}$$

Problem 9 (the number of empty bins)

Part (a): We want to calculate the expected number of empty bins. Let N be a random variable denoting the number of empty bins. Then

$$N = \sum_{i=1}^{n} I_i \,,$$

where I_i is an indicator random variable which is one if bin i is empty and is zero if bin i is not empty. The expectation of N is given by

$$E[N] = \sum_{i=1}^{n} E[I_i] = \sum_{i=1}^{n} P(I_i) \,.$$

Now $P(I_i)$ is the probability that bin i is empty, so

$$\begin{aligned} P(I_i) &= \left(\frac{i-1}{i}\right)\left(\frac{i}{i+1}\right)\left(\frac{i+1}{i+2}\right)\cdots\left(\frac{n-1}{n}\right) \\ &= \frac{i-1}{n} \quad \text{for} \quad 1 < i < n \end{aligned}$$

This is because bin i can only be filled when we insert the ith ball and it will remain empty with probability $\frac{i-1}{i}$. When we place the $i+1$st ball it will avoid the ith bin with probability

$\frac{i}{i+1}$, etc. Thus

$$\begin{aligned} E[N] &= \sum_{i=1}^{n} \frac{i-1}{n} = \frac{1}{n}\sum_{i=1}^{n}(i-1) \\ &= \frac{1}{n}\sum_{i=1}^{n-1} i = \frac{1}{n}\frac{n(n-1)}{2} = \frac{n-1}{2}\,. \end{aligned}$$

Part (b): We need to find the probability that none of the urns are empty. When we place the nth ball it must be placed in the nth bin (and no lower bin) for all bins to be filled. This happens with probability $\frac{1}{n}$. When placing the $n-1$th ball it must go in the $n-1$th bin and this will happen with probability $\frac{1}{n-1}$. Continuing in this manner we can work our way down to the first ball which must go in bin number one. Thus our probability is

$$p = 1\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\cdots\left(\frac{1}{n-1}\right)\left(\frac{1}{n}\right) = \prod_{k=1}^{n}\frac{1}{k} = \frac{1}{n!}\,.$$

Problem 10 (three trials)

Part (a): Let the random variable X denote the number of success in our three trials. Then X can be decomposed as $X = \sum_{i=1}^{3} X_i$ where X_i is a Bernoulli random variable with a value of one if trial i is a success or 0 otherwise. Then $E[X] = \sum_{i=1}^{2} E[X_i] = 1.8$. If each trial i, has the the same probability of success p then $E[X_i] = p$ and the equation earlier becomes

$$3p = 1.8 \quad \text{or} \quad p = 0.6\,,$$

I would then expect that $P\{X=3\} = \left(\begin{array}{c} 3 \\ 3 \end{array}\right) p^3 q^0 = p^3 = (0.6)^3$. Now by definition

$$E[X] = 0P\{X=0\} + 1P\{X=1\} + 2P\{X=2\} + 3P\{X=3\} = 1.8\,.$$

So if we assume that the three trials do not have the same probability of success we can maximize $P\{X=3\}$ by taking all of the other probabilities to be zero. This means that

$$3P\{X=3\} = 1.8 \Rightarrow P\{X=3\} = 0.6\,,$$

to impose the unit sum condition we can take $P\{X=4\} = 0.4$ and $P\{X=1\} = P\{X=2\} = 0$. With $P\{X=3\} = 0.6$ provides a desired realization of our probability space.

Part (b): To make $P\{X=3\}$ as small as possible take it to be zero. We now need to specify the remaining probabilities such that they sum to one i.e.

$$P\{X=0\} + P\{X=1\} + P\{X=2\} = 1\,,$$

the expectation of X is correct

$$0P\{X=0\} + 1P\{X=1\} + 2P\{X=2\} = 1.8\,,$$

and of course $0 \leq P\{X = i\} \leq 1$, the expectation calculation given $P\{X = 1\} = 1.8 - 2P\{X = 2\}$ which when put into the sum to unity constraint gives

$$P\{X = 0\} + (1.8 - 2P\{X = 2\}) + P\{X = 2\} = 1\,,$$

so

$$p\{X = 0\} - P\{X = 2\} = -0.8\,.$$

One easy solution to this equation (there are multiple) is to take $P\{X = 2\} = 0.8$ then $P\{X = 0\} = 0$ and $P\{X = 1\} = 0.2$ so a probability scenario that result in $P\{X = 3\}$ minimum is

$$P\{X = 0\} = 0\,, P\{X = 1\} = 0.2\,, P\{X = 2\} = 0.8\,, P\{X = 3\} = 0\,.$$

Problem 11 (changeovers)

We will have a change over, if a head changes to a tail (this happen with probability $1 - p$) or a tail changes to a head (this happens with probability p). Lets let the number of change over be represented by the random variable N. Then following the hint, we can decomposed N into a sum of Bernoulli random variables X_i. This is

$$N = \sum_{i=1}^{n-1} X_i\,,$$

Now the random variable X_i takes the value one if the coin changes from a head to a tail or from a tail to a head and is equal to zero if it does not change type (i.e. stays heads or tails). Then

$$E[N] = \sum_{i=1}^{n-1} E[X_i] = \sum_{i=1}^{n-1} P(X_i)\,.$$

To evaluate $E[N]$ we need to compute $P(X_i)$. This will be

$$p(1 - p) + (1 - p)p = 2p(1 - p)\,,$$

which can be seen by conditioning on the type of the coin flip. That is

$$\begin{aligned} P(X_i) &= P\{X_i = 1|(T,T)\}P(T,T) + P\{X_i = 1|(T,H)\}P(T,H) \\ &+ P\{X_i = 1|(H,T)\}P(H,T) + P\{X_i = 1|(H,H)\}P(H,H) \\ &= 0 + (1 - p)p + p(1 - p) + 0 = 2p(1 - p)\,. \end{aligned}$$

Thus we have that

$$E[N] = 2(n - 1)p(1 - p)\,.$$

Problem 13 (1000 cards for 1000 people)

Lets assume that the oldest person lives to be M years old (no one yet lives to be 1000 years old). Let A_i be the indicator random variable denoting if the person holding card i has an age that matches the number on the card he is holding. Then let N be the random variable representing the total number of people who's age matches the card that they are holding. Then

$$N = \sum_{i=1}^{1000} A_i \,.$$

So that $E[N]$ is given by

$$E[N] = \sum_{i=1}^{1000} E[A_i] = \sum_{i=1}^{1000} P\{A_i\} \,.$$

Now if we assume that people of all ages (L to M) are represented in our sample. We have $P\{X_i\} = 0$ if $i \leq L - 1$ or $i \geq M + 1$ so the above equals

$$\sum_{i=L}^{M} P\{A_i\} = \sum_{i=L}^{M} \left(\frac{1}{M - L + 1} \right) = 1 \,.$$

Since for the people holding the cards numbered $L, L+1, L+2, \cdots, M-1, M$ each has a $\frac{1}{M-L+1}$ chance of having the correct date on it.

Problem 14 (the expected number of draws to remove all black balls)

Let X_m be the number of draws or iterations required to take the urn from m black balls to $m-1$ black balls. Let X_{m-1} be the number of iterations needed to take the urn from $m-1$ black balls to $m-2$ black balls, etc. Then the total number of stages needed is given by

$$N = \sum_{i=m}^{1} X_i \,,$$

so that $E[N] = \sum_{i=m}^{1} E[X_i]$. Now to complete this problem we compute each $E[X_i]$ in tern.

Now $E[X_m]$ is the expected number of draws to reduce the number of black balls by one (to $m-1$). This will happen with probability $1-p$. Thus X_m is a geometric random variable with probability of success given by $1-p$. Thus

$$P\{X_m = i\} = p^{i-1}(1-p) \quad \text{for} \quad i = 1, 2, \cdots$$

This variable X_m has an expected value of $\frac{1}{1-p}$. This result holds for every random variable X_i. Thus we have

$$E[N] = \sum_{i=m}^{1} \left(\frac{1}{1-p} \right) = \frac{m}{1-p} \,.$$

Problem 15

Let E_{ij} be an indicator random variable denoting if the man i and j form a matched pair. Then let N be the random variable denoting the number of matched pairs. Then

$$N = \sum_{(i,j)} E_{i,j} \,,$$

so the expectation of N is given by

$$E[N] = \sum_{(i,j)} E[E_{i,j}] = \sum_{(i,j)} P(E_{i,j}) \,.$$

Now $P(E_{i,j}) = \frac{1}{N}\left(\frac{1}{N-1}\right)$ and there are $\left(\begin{array}{c} N \\ 2 \end{array}\right)$ total pairs in the sum. Thus

$$E[N] = \left(\begin{array}{c} N \\ 2 \end{array}\right)\left(\frac{1}{N(N-1)}\right) = \frac{N(N-1)}{2}\frac{1}{N}\frac{1}{N-1} = \frac{1}{2} \,.$$

Problem 16

Let

$$X = \left\{ \begin{array}{cc} Z & Z > X \\ 0 & \text{otherwise} \end{array} \right. \,,$$

which defines a function $f(\cdot)$ of our random variable (this function has x as a parameter). Now using the definition of the expectation of a random variable we have for $E[X]$ the following expression

$$\begin{aligned} E[X] &= \int xp(x)dx = \int f(z)p(z)dz \\ &= \int_x^\infty zp(z)dz = \int_x^\infty z\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \,. \end{aligned}$$

To evaluate this integral let $v = \frac{z}{\sqrt{2}}$ so that $dv = \frac{dz}{\sqrt{2}}$ and $dz = \sqrt{2}dv$. Then the above integral becomes

$$\begin{aligned} E[X] &= \int_{x/\sqrt{2}}^\infty \frac{\sqrt{2}v}{\sqrt{2\pi}}e^{-v^2}\sqrt{2}dv = \frac{\sqrt{2}}{\sqrt{\pi}}\int_{x/\sqrt{2}}^\infty ve^{-v^2}dv \\ &= \frac{\sqrt{2}}{\sqrt{\pi}}\left.\frac{e^{-v^2}}{2(-1)}\right|_{x/\sqrt{2}}^\infty = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \end{aligned}$$

Problem 17 (counting cards or not)

Part (a): If we are not given any information about our earlier guesses then we must pick one of the $n!$ orderings of the cards and just count the number of matches we have. Let

A_i be an indicator random variable determining if the cards at position i match. Then let N be the random variable denoting the number of matches. Thus $N = \sum_{i=1}^{n} A_i$ so $E[N] = \sum_{i=1}^{n} E[A_i] = \sum_{i=1}^{n} P(A_i)$. But $P(A_i) = \frac{1}{n}$, since at position i we have one chance in n of finding a match. Thus we have that

$$E[N] = \sum_{i=1} nP(A_i) = \sum_{i=1}^{n} \frac{1}{n} = \frac{n}{n} = 1\,,$$

as claimed.

Part (b): The best strategy is to obviously not guess any of the cards that one is shown but at each stage pick uniformly from among the possible remaining $n - i$ unrevealed cards. Thus with this prescription and the definition of A_i as above we have that

$$E[N] = \sum_{i=1}^{n} P(A_i)\,.$$

Now in this case we have that

$$\begin{aligned}
P(A_1) &= \frac{1}{n} \\
P(A_2) &= \frac{1}{n-1} \\
P(A_3) &= \frac{1}{n-2} \\
&\vdots \\
P(A_2) &= \frac{1}{2} \\
P(A_1) &= 1
\end{aligned}$$

Thus we have that

$$E[N] = \frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{2} + 1 \approx \int_1^n \frac{dx}{x} = \ln(n)\,.$$

Part (c): The optimal strategy in this case is to guess a single card repeatedly until we are told it is correct. Once we have been told that our card is correct we guess a second card repeatedly until we are told that we have guessed that card correctly. This procedure is repeated until all of the cards are turned over. This procedure relies on some ordering of the cards we guess. Let i be the C_ith card we guess for $1 \leq i \leq n$. Let I_i be an indicator random variable indicating whether we eventually guess this ith card correctly at some point. Then N the total number of correct guess is given by

$$N = \sum_{i=1}^{n} I_i \quad \text{so} \quad E[N] = \sum_{i=1}^{n} E[I_i] = \sum_{i=1}^{n} P(I_i)\,.$$

We now evaluate $P(I_i)$. Note that $P(I_1) = 1$ since for any ordering of cards we will eventually guess our first card correctly. We next have $P(I_2) = \frac{1}{2!}$ because we will be able to attempt to guess our second card correctly only if we first guess our first card correctly and then guess the second card correctly. This will only happen if the second card we choose to guess occurs *after* our first guess in the shuffled deck. There are 2! possible ordering of these two cards in the shuffled deck and in only one of them does this happen. Thus the probability is $\frac{1}{2!}$. For $P(E_3)$ the same logic holds. That is when we look sequentially in the shuffled deck to get three cards correct we must first observe our first guess, then our second guess, and finally our third guess. There are 3! ways to arrange these three guessed cards and only 1 of them has this ordering. Thus $P(I_3) = \frac{1}{3!}$. Using this logic we find that

$$E[N] = 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots + \frac{1}{(n-1)!} + \frac{1}{n!} \approx e^1 .$$

Problem 18 (counting matched cards)

Let A_i be the event that when we turn over card i it matches the required cards face. For example A_1 is the event that turning over card one reveals an ace, A_2 is the event that turning over the second card reveals a deuce etc. The the number of matched cards N is give by the sum of these indicator random variable as

$$N = \sum_{i=1}^{52} A_i .$$

Taking the expectation of this result and using linearity requires us to evaluate $E[A_i] = P(A_i)$. For card i the probability that when we turn it over it matches the expected face is given by

$$P(A_i) = \frac{4}{52} ,$$

since there are four suites that could match a given face. Thus we then have for the expected number of matching cards that

$$E[N] = \sum_{i=1}^{52} E[A_i] = \sum_{i=1}^{52} P(A_i) = 52 \cdot \frac{4}{52} = 4 .$$

Problem 19

Part (a): We will catch type $i = 2, 3, \cdots, r$ insects with probability $\sum_{i=2}^{r} P_i = 1 - P_1$. Thus the probability that we catch $i - 1$ insects of type $2, 3, \cdots, r$ before we catch one of of type one is given by a geometric random variable. This means that if X is the random variable representing the trial that resulted in catching a type one insect then

$$P\{X = i\} = (1 - P_1)^{i-1} P_1 .$$

We are interested in the expected value of the number of type $2, 3, \cdots, r$ insects caught before catching one of type one. Thus as a random variable N in terms of X is given by

$N = X - 1$. With this we see that

$$E[N] = E[X] - 1 = \frac{1}{P_1} - 1 = \frac{P_1 - 1}{P_1}\,.$$

Part (b): We can compute the mean number of insect types before catching one of type one by using conditional expectation. Let N be the random variable that denotes the different number of insects caught before catching one of type one. Also let the random variable K be the total number of insects caught when we catch our first type one insect. This means that on catch K we catch our first type one insect. Then conditioning on this random variable K we have

$$E[N] = E[E[N|K]]\,.$$

Now to evaluate $E[N|K]$ we recognize that this is the expectation number of insects from $2, 3, \cdots, r$ caught when we catch our type one insect on catch k. Then since for the catches $1, 2, 3, \cdots, k-1$ we are selecting from the insect types $2, 3, \cdots, r$, each specific insect type is caught with probability

$$\frac{P_2}{\sum_{i=2}^r P_i}\,, \quad \frac{P_3}{\sum_{i=2}^r P_i}\,, \quad \cdots\,, \quad \frac{P_r}{\sum_{i=2}^r P_i}\,.$$

Now since

$$\sum_{r=2}^{r} P_i = 1 - P_1\,,$$

the above terms are equivalent to

$$\frac{P_2}{1-P_1}\,, \quad \frac{P_3}{1-P_1}\,, \quad \cdots \quad \frac{P_r}{1-P_1}\,.$$

From Example 3d in this chapter, the expected number of different insects caught is given by

$$\begin{aligned} E[N|K] &= (r-1) - \sum_{i=2}^{r} \left(1 - \tilde{P}_i\right)^{k-1} \\ &= (r-1) - \sum_{i=2}^{r} \left(1 - \frac{P_i}{1-P_1}\right)^{k-1}\,. \end{aligned}$$

Taking the outer expectation over the random variable K we have

$$E[N] = \sum_{k=1}^{\infty} \left((r-1) - \sum_{r=2}^{r} \left(1 - \frac{P_i}{1-P_1}\right)^{k-1} \right) P\{K = k\}\,.$$

Since K is a geometric random variable, we have that

$$P\{K = k\} = (1-P_1)^{k-1} P_1\,,$$

which gives $E[N]$ as

$$\begin{aligned} E[N] &= \sum_{k=1}^{\infty}(r-1)P\{K=k\} - \sum_{i=2}^{r}\left(1-\frac{P_i}{1-P_1}\right)^{k-1} P_1 \\ &= (r-1) - P_1\sum_{k=1}^{\infty}\sum_{i=2}^{r}(1-P_1-P_i)^{k-1} \\ &= (r-1) - P_1\sum_{i=2}^{r}\sum_{k=1}^{\infty}(1-P_1-P_i)^{k-1} \\ &= (r-1) - P_1\sum_{i=2}^{r}\frac{1}{P_1+P_i}. \end{aligned}$$

Problem 21 (more birthday problems)

Part (a): Let $A_{i,j,k}$ be an indicator random variable if persons i, j, and k have the same birthday *and no one else does*. Then if we let N denote the random variable representing the number of groups of three people all of whom have the same birthday we see that N is given by a sum of these random variables as

$$N = \sum_{i<j<k} A_{i,j,k}.$$

Then taking the expectation of the above expression we have

$$E[N] = \sum_{i<j<k} E[A_{i,j,k}].$$

Now there are $\binom{100}{3}$ terms in the above sum (since there are one hundred total people and our sum involves all subsets of three people), and the probability of each event $A_{i,j,k}$ happening is given by

$$\begin{aligned} P(A_{i,j,k}) &= \frac{1}{365^2}\left(1-\frac{1}{365}\right)^{100-3} \\ &= \frac{1}{365^2}\left(\frac{364}{365}\right)^{97} \end{aligned}$$

since person j and person k's birthdays must match that of person i, and the remaining 97 people must have different birthdays (the problem explicitly states we are looking for the expected number days that are the birthday of *exactly* three people and not more). Thus the total expectation of the number of groups of three people that have the same birthday is then given by

$$E[N] = \binom{100}{3}\frac{1}{365^2}\left(\frac{364}{365}\right)^{97} = 0.93014,$$

in agreement with the back of the book.

Part (b): To find the expected number of distinct birthdays we define an indicator random variable X_i to be 1 if the i-th day has *at least* one birthday falling on it (for $i = 1, 2, \ldots, 365$). Then defining N to be the total number of distinct birthdays we have

$$N = \sum_{i=1}^{365} X_i \,.$$

Since the variables X_i are independent the expectation of this expression is given by

$$E[N] = \sum_{i=1}^{365} E[X_i] = \sum_{i=1}^{365} P\{X_i = 1\} = 365 P\{X_1 = 1\} \,.$$

To compute $P\{X_1 = 1\}$ we recognized this as the probability that one specific date has at least one person with a birthday one it. This is the converse of the probability that no one has a birthday on this date. No one has a birthday on this date with probability $\left(\frac{364}{365}\right)^{100}$. Thus

$$P\{X_1 = 1\} = 1 - \left(\frac{364}{365}\right)^{100} \,.$$

With this we find that

$$E[N] = 365\left(1 - \left(\frac{364}{365}\right)^{100}\right) = 87.576 \,.$$

Problem 22 (number of times to roll a fair die to get all six sides)

This is exactly like the coupon collecting problem where we have six coupons with a probability of obtaining any one of them given by 1/6. Then this problem is equivalent to determining the expected number of coupons we need to collect before we get a complete set. From Example 2i from the book we have the expected number of rolls X to be given by

$$E[X] = N\left[1 + \frac{1}{2} + \cdots + \frac{1}{N-1} + \frac{1}{N}\right]$$

when $N = 6$ this becomes

$$E[X] = 6\left[1 + \frac{1}{2} + \cdots + \frac{1}{5} + \frac{1}{6}\right] = 14.7 \,.$$

Problem 26

Part (a): The density for the random variable $X_{(n)}$ defined as $X_{(n)} = \max(X_1, X_2, \cdots, X_n)$ where X_i is drawn from the distribution function $F(x)$ (density function $f(x)$) is given by

$$f_{X_{(n)}}(x) = \frac{n!}{(n-1)!}(F(x))^{n-1} f(x) = nF(x)^{n-1} f(x) \,.$$

When X_i is a uniform random variable between $[0, 1]$ we have $f(x) = 1$, and $F(x) = x$, so that the above becomes

$$f_{X_{(n)}}(x) = nx^{n-1}\,.$$

Then the expectation of this random variable is given by

$$E[X_{(n)}] = \int_0^1 x(nx^{n-1})dx = \left.\frac{nx^{n+1}}{n+1}\right|_0^1 = \frac{n}{n+1}\,.$$

Part (b): The minimum random variable $X_{(1)}$ defined by $X_{(1)} = \min(X_1, X_2, \cdots, X_n)$ has a distribution function given by

$$f_{X_{(1)}}(x) = n(1 - F(x))^{n-1} f(x)\,.$$

Again when X_i is a uniform random variable our expectation is given by

$$\begin{aligned} E[X_{(1)}] &= \int_0^1 xn(1-x)^{n-1}dx \\ &= n\left[\left.\frac{x(1-x)^n(-1)}{n}\right|_0^1 - \int_0^1 \frac{(1-x)^n}{n}(-1)dx\right] \\ &= \int_0^1 (1-x)^n dx = \left.\frac{(1-x)^{n+1}(-1)}{n+1}\right|_0^1 \\ &= \frac{1}{n+1}\,. \end{aligned}$$

Problem 30 (a squared expectation)

We find, by expanding the quadratic and using independence, that

$$E[(X-Y)^2] = E[X^2 - 2XY + Y^2] = E[X^2] - 2E[X]E[Y] + E[Y^2]\,.$$

In terms of the variance $E[X^2]$ is given by $E[X^2] = \text{Var}(X) + E[X]^2$ so the above becomes

$$\begin{aligned} E[(X-Y)^2] &= \text{Var}(X) + E[X]^2 - 2E[X]E[Y] + \text{Var}(Y) + E[Y]^2 \\ &= \sigma^2 + \mu^2 - 2\mu^2 + \sigma^2 + \mu^2 = 2\sigma^2\,. \end{aligned}$$

Problem 33 (evaluating expectations and variances)

Part (a): We find, expanding the quadratic and using the linearity property of expectations that

$$E[(2+X)^2] = E[4 + 4X + X^2] = 4 + 4E[X] + E[X^2]\,.$$

In terms of the variance, $E[X^2]$ is given by $E[X^2] = \text{Var}(X) + E[X]^2$, both terms of which we know from the problem statement. Using this the above becomes

$$E[(2+X)^2] = 4 + 4(1) + (5 + 1^2) = 14\,.$$

Part (b): We find, using properties of the variance that

$$\text{Var}(4 + 3X) = \text{Var}(3X) = 9\text{Var}(X) = 9 \cdot 5 = 45\,.$$

Problem 40 (computing an expectation)

If

$$f(x,y) = \frac{1}{y}e^{-(y+x/y)} \quad \text{for} \quad x > 0\,, y > 0$$

then by the definition of the expectation we have that

$$\begin{aligned} E[X] &= \int_{x=0}^{\infty} x f(x) dx = \int_{x=0}^{\infty} x \int_{y=0}^{\infty} f(x,y) dy dx \\ &= \int_{x=0}^{\infty} \int_{y=0}^{\infty} \left(\frac{x}{y}\right) e^{-\left(y+\frac{x}{y}\right)} dy dx \\ &= \int_{y=0}^{\infty} \int_{x=0}^{\infty} \left(\frac{x}{y}\right) e^{-\left(y+\frac{x}{y}\right)} dx dy\,. \end{aligned}$$

The last two lines are obtained by exchanging the order of the integration. To integrate this expression with respect to x, let v be defined as $v = \frac{x}{y}$, so that $dv = \frac{dx}{y}$, and the above expression becomes

$$\begin{aligned} E[X] &= \int_{y=0}^{\infty} \int_{v=0}^{\infty} v e^{-(y+v)} y dv dy \\ &= \int_{y=0}^{\infty} y e^{-y} \int_{v=0}^{\infty} v e^{-v} dv dy\,. \end{aligned}$$

Now evaluating the v integral using integration by parts we have

$$\begin{aligned} \int_0^{\infty} v e^{-v} dv &= -v e^{v}\Big|_0^{\infty} + \int_0^{\infty} e^{-v} dv \\ &= -e^{-v}\Big|_0^{\infty} = -(0-1) = 1\,. \end{aligned}$$

With this the expression for $E[X]$ becomes

$$E[X] = \int_{y=0}^{\infty} y e^{-y} 1 dy = 1\,.$$

Now in the same way

$$E[Y] = \int_{y=0}^{\infty} y \int_{x=0}^{\infty} f(x,y) dx dy = \int_{y=0}^{\infty} y \int_{x=0}^{\infty} \frac{1}{y} e^{-\left(y+\frac{x}{y}\right)} dx dy\,.$$

To evaluate the x integral, let $v = \frac{x}{y}$ then $dv = \frac{dx}{y}$ and we have the above equal to

$$\begin{aligned} E[Y] &= \int_{y=0}^{\infty} y \int_{v=0}^{\infty} \frac{1}{y} e^{-y} e^{-v} y dv dy \\ &= \int_{y=0}^{\infty} y e^{-y} \int_{v=0}^{\infty} e^{-v} dv dy \\ &= \int_{y=0}^{\infty} y e^{-y} \left(-e^{-v}\Big|_{v=0}^{\infty}\right) dy \\ &= \int_{y=0}^{\infty} y e^{-y} dy = -\, y e^{-y}\Big|_0^{\infty} + \int_{y=0}^{\infty} e^{-y} dy = 1\,. \end{aligned}$$

Finally to compute $\text{Cov}(X,Y)$ using the definition we require the calculation of $E[XY]$. This is given by

$$\begin{aligned} E[XY] &= \int_0^\infty \int_0^\infty xy\frac{1}{y}e^{-(y+\frac{x}{y})}dxdy \\ &= \int_{y=0}^\infty \int_{x=0}^\infty xe^{-(y+\frac{x}{y})}dxdy \,. \end{aligned}$$

To perform the x integration let $v = \frac{x}{y}$ so that $dv = \frac{dx}{y}$ and the above becomes

$$\begin{aligned} E[XY] &= \int_{y=0}^\infty \int_{v=0}^\infty yve^{-(y+v)}ydvdy \\ &= \int_{y=0}^\infty y^2e^{-y}\int_{v=0}^\infty ve^{-v}dvdy \,. \end{aligned}$$

Since $\int_0^\infty ve^{-v}dv = 1$, the above equals

$$\int_0^\infty y^2e^{-y}dy = -y^2e^{-y}\Big|_0^\infty + 2\int_0^\infty ye^{-y}dy = 2 \,.$$

Then

$$\text{Cov}(X,Y) = E[XY] - E[X]E[Y] = 2 - 1(1) = 1 \,.$$

Problem 45 (pairwise uncorrelated)

To be pairwise uncorrelated means that $\text{Cor}(X_i, X_j) = 0$ if $i \neq j$.

Part (a): We have

$$\begin{aligned} \text{Cov}(X_1+X_2, X_1+X_2) &= \sum_{i=1}^2\sum_{j=2}^3 \text{Cov}(X_i,X_j) \\ &= \text{Cov}(X_1,X_2) + \text{Cov}(X_1,X_3) \\ &+ \text{Cov}(X_2,X_2) + \text{Cov}(X_2,X_3) \,. \end{aligned}$$

Using the fact that that these variables are *pairwise* uncorrelated the right hand side of the above equals

$$0 + 0 + 1^2 + 0 = 1 \,.$$

The *correlation* between to random variables X and Y is (defined as)

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

follows once we have the two variances. We now compute these variances

$$\begin{aligned} \text{Var}(X_1+X_2) &= \text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1,X_2) \\ &= \text{Var}(X_1) + \text{Var}(X_2) = 1 + 1 = 2 \,. \end{aligned}$$

In the same way $\text{Var}(X_2 + X_3) = 2$, so that

$$\rho(X_1 + X_2, X_2 + X_3) = \frac{1}{\sqrt{2}}$$

Part (b): We have that

$$\begin{aligned} \text{Cov}(X_1 + X_2, X_3 + X_4) &= \text{Cov}(X_1, X_3) + \text{Cov}(X_1, X_4) \\ &+ \text{Cov}(X_2, X_3) + \text{Cov}(X_2, X_4) = 0\,. \end{aligned}$$

so obviously then $\rho(X_1 + X_2, X_3 + X_4) = 0$ regardless of the value of the variances.

Problem 48 (conditional expectation of die rolling)

Part (a): The probability that the first six is rolled on the nth roll is given by a geometric random variable with parameter $p = 1/6$. Thus the expected number of rolls to get a six is given by

$$E[X] = \frac{1}{p} = 6\,.$$

Part (b): We want to evaluate $E[X|Y = 1]$. Since in this expectation we are told that the first roll of our dice results in a five we have that

$$E[X|Y = 1] = 1 + E[X] = 1 + \frac{1}{p} = 1 + 6 = 7\,,$$

since after the first roll we again have that the number of rolls to get the first six is a geometric random variable with $p = 1/6$.

Part (c): We want to evaluate $E[X|Y = 5]$, which means that the first five happens on the fifth roll. Thus the rolls $1, 2, 3, 4$ all have a probability of $1/5$ to show a six. After the fifth roll, there are again six possible outcomes of the die so the probability of obtaining a six is given by $1/6$. Defining the event A to be the event that we do not roll a six in any of the first four rolls (and implicitly given that the first five happens on the fifth roll) we see that

$$P(A) = \left(\frac{4}{5}\right)^4 = 0.4096\,,$$

since with probability of $1/5$ we will roll a six and with probability $4/5$ we will not roll a six. With this definition and using the definition of expectation we find that

$$\begin{aligned} E[X|Y = 5] &= 1\left(\frac{1}{5}\right) + 2\left(\frac{4}{5}\right)\frac{1}{5} + 3\left(\frac{4}{5}\right)^2\frac{1}{5} + 4\left(\frac{4}{5}\right)^3\frac{1}{5} \\ &+ \sum_{k=6}^{\infty} k\left(P(A)\left(\frac{5}{6}\right)^{k-6}\frac{1}{6}\right)\,. \end{aligned}$$

We will evaluate this last sum numerically. This is done in the Matlab file `chap_7_prob_48.m`, where we find that

$$[X|Y = 5] = 5.8192\,,$$

in agreement with the book.

Problem 49 (misshaped coins)

We desire to compute the conditional expected number of heads in our ten flips. Let N be the random variable specifying the number of heads obtained when we flip our coin ten ten times. Let E be the event that on the first three flips we obtain two heads and one tail. Then $E[N|E]$ can be computed by conditioning on the misshapen coin chosen. Let A be the event that we select the coin with $p = 0.4$. Then

$$E[N|E] = E[N|A, E]P(A|E) + E[N|A^c, E]P(A^c|E)\,.$$

We first compute $E[N|A, E]$. The easiest way to do this is to notice that this is two plus the expectation of a binomial random variable with parameters $(n, p) = (7, 0.4)$. Since two of the first three flips resulted in a head. Thus

$$E[N|A, E] = 2 + 0.4(7) = \frac{24}{5}\,.$$

In the same way

$$E[N|A^c, E] = 2 + 0.7(7) = \frac{69}{10}\,.$$

We now need to compute $P(A|E)$ using Bayes' rule we find

$$P(A|E) = \frac{P(E|A)P(A)}{P(E)} = \frac{P(E|A)P(A)}{P(E)} = \frac{P(E|A)P(A)}{P(E|A)P(A) + P(E|A^c)P(A^c)}\,.$$

Now we have $P(E|A)$ and $P(E|A^c)$ given by

$$P(E|A) = \left(\begin{array}{c} 3 \\ 2 \end{array}\right)(0.4)^2(0.6) = \frac{36}{125}$$
$$P(E|A^c) = \left(\begin{array}{c} 3 \\ 2 \end{array}\right)(0.7)^2(0.3) = \frac{441}{1000}\,.$$

Thus we find assuming that $P(A) = P(B) = \frac{1}{2}$ that $P(A|E) = \frac{32}{81}$ and $P(A^c|E) = \frac{49}{81}$ to get

$$E[N|E] = \frac{32}{81}\left(\frac{48}{10}\right) + \frac{49}{81}\left(\frac{69}{10}\right) = \frac{1639}{270} = 6.0704\,.$$

Problem 50 (compute $E[X^2|Y = y]$)

By definition, the requested expectation is given by

$$E[X^2|Y = y] = \int_0^\infty x^2 f(x|Y = y)dx\,.$$

Lets begin by computing $f(x|Y = y)$, using the definition of this density in terms of the joint density

$$f(x|y) = \frac{f(x, y)}{f(y)}\,.$$

Since we are given $f(x, y)$ we begin by first computing $f(y)$. We find that

$$\begin{aligned} f(y) &= \int_0^\infty f(x,y)dx = \int_0^\infty \frac{e^{-x/y}e^{-y}}{y}dx \\ &= \frac{e^{-y}}{y}\int_0^\infty e^{-x/y}dx = \frac{e^{-y}}{y}(-y)\,e^{-x/y}\Big|_0^\infty \\ &= e^{-y}\,. \end{aligned}$$

So that $f(x|y)$ is given by

$$f(x|y) = \frac{e^{-x/y}e^{-y}}{y}e^y = \frac{e^{-x/y}}{y}\,.$$

With this expression we can evaluate our expectation above. We have (using integration by parts several times)

$$\begin{aligned} E[X^2|Y=y] &= \int_0^\infty x^2\frac{e^{-x/y}}{y}dx \\ &= \frac{1}{y}\int_0^\infty x^2e^{-x/y}dx \\ &= \frac{1}{y}\left(x^2(-y)e^{-x/y}\Big|_0^\infty - \int_0^\infty 2x(-y)e^{-x/y}dx\right) \\ &= 2\int_0^\infty xe^{-x/y}dx \\ &= 2\left(x(-y)e^{-x/y}\Big|_0^\infty - \int_0^\infty (-y)e^{-x/y}dx\right) \\ &= 2y\int_0^\infty e^{-x/y}dx \\ &= 2y(-y)\,e^{-x/y}\Big|_0^\infty = 2y^2\,. \end{aligned}$$

Problem 51 (compute $E[X^3|Y=y]$)

By definition, the requested expectation is given by

$$E[X^3|Y=y] = \int x^3 f(x|Y=y)dx\,.$$

Lets begin by computing $f(x|Y=y)$, using the definition of this density in terms of the joint density

$$f(x|y) = \frac{f(x,y)}{f(y)}\,.$$

Since we are given $f(x, y)$ we begin by first computing $f(y)$. We find that

$$f(y) = \int_0^y f(x,y)dx = \int_0^y \frac{e^{-y}}{y}dx = e^{-y}\,.$$

So that $f(x|y)$ is given by

$$f(x|y) = \frac{e^{-y}}{y}e^y = \frac{1}{y}\,.$$

With this expression we can evaluate our expectation above. We have

$$E[X^3|Y=y] = \int_0^y x^3 \frac{1}{y} dx = \frac{1}{y} \left. \frac{x^4}{4} \right|_0^y = \frac{y^3}{4}.$$

Problem 52 (the average weight)

Let W denote the random variable representing the weight of a person selected from the total population. Then we can compute $E[W]$ by conditioning on the subgroups. Letting G_i denote the event we are drawing from subgroup i, we have

$$E[W] = \sum_{i=1}^{r} E[W|G_i]P[G_i] = \sum_{i=1}^{r} w_i p_i .$$

Problem 53 (the time to escape)

Let T be the random variable denoting the number of days until the prisoner reaches freedom. We can evaluate $E[T]$ by conditioning on the door selected. If we denote D_i be the event the prisoner selects door i then we have

$$E[T] = E[T|D_1]P(D_1) + E[T|D_2]P(D_2) + E[T|D_3]P(D_3) .$$

Each of the above expressions can be evaluated. For example if the prisoner selects the first door then after two days he will be right back where he started and thus has in expectation $E[T]$ more days left. Thus

$$E[T|D_1] = 2 + E[T] .$$

Using logic like this we see that $E[T]$ can be expressed as

$$\begin{aligned} E[T] &= E[T|D_1]P(D_1) + E[T|D_2]P(D_2) + E[T|D_3]P(D_3) \\ &= (2 + E[T])(0.5) + (4 + E[T])(0.3) + (1)(0.2) . \end{aligned}$$

Solving the above expression for $E[T]$ we find that $E[T] = 12$.

Problem 55 (shooting ducks or not)

Let D be a random variable denoting the total number of ducks that are shot by our ten hunters. The expectation of this number will depend on the number of ducks that were observed and is denoted X. We are told that X is given by a Poisson random variable with mean 6. Thus

$$P\{X = x\} = \frac{e^{-6} 6^x}{x!} .$$

Then to compute the desired expectation we can condition on the variable X. We have

$$E[D] = \sum_{x=1}^{\infty} E[D|X = x] P\{X = x\} .$$

We now need to compute $E[D|X = x]$. If we assume that the variable $D|X = x$ or the total number of hit ducks when x ducks are seen, is a Binomial random variable with a probability that at least one of the 10 hunters shots hitting a single duck given by p_{D^*}, the expectation $E[D|X = x]$ is then xp_{D^*}. Thus we have

$$E[D] = \sum_{x=1}^{\infty} xp_{D^*} P\{X = x\} ,$$

and we now need to evaluate p_{D^*} or the probability that at least one of the 10 hunters hits a given duck. To determine that, consider the probability that a given duck would *not* get hit by the ith hunter denoted by p_i for $1 \leq i \leq 10$. Then in that case none of the hunters will hit this duck with probability p_i^{10} and at least one will hit it with probability $1 - p_i^{10}$. This duck will not get hit by our ith hunter if

- the hunter does not choose to shoot at this particular duck.
- the hunter does choose to shoot at this particular duck but misses when shooting at him.

Thus conditioning on each of these two events we have

$$\begin{aligned} p_i &= P(\text{Hunter misses}|\text{Hunter chooses this duck})P(\text{Hunter chooses this duck}) \\ &+ 1 \cdot P(\text{Hunter does not choose this duck}) \\ &= (1 - 0.6)\frac{1}{x} + \left(1 - \frac{1}{x}\right) = 1 - \frac{0.6}{x} . \end{aligned}$$

As already discussed all 10 hunters miss with probability $\left(1 - \frac{0.6}{x}\right)^{10}$ and at least one will hit this duck with probability $1 - \left(1 - \frac{0.6}{x}\right)^{10}$. Putting everything together we have

$$E[D] = \sum_{x=1}^{\infty} x \left[1 - \left(1 - \frac{0.6}{x}\right)^{10} \right] \left(\frac{e^{-6}6^x}{x!} \right) = 3.6989 ,$$

when we evaluate this sum numerically.

Problem 56 (the number of elevator stops)

Let M be the random variable representing the number of people who enter an elevator on the ground floor. Then once we are loaded up with the M people then we can envision each of the M people uniformly selecting one of the N floors to get off on. This is exactly the same as counting the number of different coupons collected with probability of selecting each type to be $\frac{1}{N}$. Thus we can compute the expected number of stops made by the elevator by conditioning on the number of passengers loaded in the elevator initially and the result from Example 3d (the expected number of distinct coupons when drawing M). For example, let X be the random variable denoting the number of stops made when M passengers are on board. Then we want to compute

$$E[X] = E[E[X|M]] .$$

Now $E[X|M = m]$ is given by the result of Example 3d so that

$$E[X|M = m] = N - \sum_{i=1}^{N} \left(1 - \frac{1}{N}\right)^m = N - N\left(1 - \frac{1}{N}\right)^m .$$

Thus the total expectation of X is then given by

$$\begin{aligned}
E[X] &= \sum_{m=0}^{\infty} E[X|M = m]P\{M = m\} \\
&= \sum_{m=0}^{\infty} \left(N - N\left(1 - \frac{1}{N}\right)^m\right) P\{M = m\} \\
&= N\sum_{m=0}^{\infty} P\{M = m\} - N\sum_{m=0}^{\infty}\left(1 - \frac{1}{N}\right)^m P\{M = m\} \\
&= N - N\sum_{m=0}^{\infty}\left(1 - \frac{1}{N}\right)^m \frac{e^{-10}10^m}{m!} \\
&= N - Ne^{-10}\sum_{m=0}^{\infty} \frac{\left(10\left(1 - \frac{1}{N}\right)\right)^m}{m!} \\
&= N - Ne^{-10}\exp\{10\left(1 - \frac{1}{N}\right)\} \\
&= N\left(1 - \exp\{-10 + 10 - \frac{1}{N}\}\right) = N\left(1 - e^{-\frac{10}{N}}\right)
\end{aligned}$$

Problem 57

Let N_A be the random variable denoting the number of accidents in a week. Then $E[N_A] = 5$. Let N_I be the random variable denoting the number of injured when an accident occurs. Let N be the total number of workers injured each week. Then $E[N]$ can be calculated by conditioning on the number of accidents in a given week N_A as

$$E[N] = E[E[N|N_A]] .$$

Now we are told that the number of workers injured in each accident is independent of the number of accidents that occur. We then have that

$$E[E[N|N_A]] = E[N|N_A] \cdot E[N_A] = 2.5 \cdot 5 = 12.5 .$$

Problem 58 (flipping a biased coin until a head and a tail appears)

Part (a): We reason as follows if the first flip lands heads then we will continue to flip until a tail appears at which point we stop. If the first flip lands tails we will continue to flip until a head appears. In both cases the number of flips required until we obtain our desired outcome (a head and a tail) is a geometric random variable. Thus computing the desired expectation is easy once we condition on the result of the first flip. Let H denote the event that the first flip lands heads then with N denoting the random variable denoting the number of flips until both a head and a tail occurs we have

$$E[N] = E[N|H]P\{H\} + E[N|H^c]P\{H^c\}\,.$$

Since $P\{H\} = p$ and $P\{H^c\} = 1 - p$ the above becomes

$$E[N] = pE[N|H] + (1-p)E[N|H^c]\,.$$

Now we can compute $E[N|H]$ and $E[N|H^c]$. Now $E[N|H]$ is one plus the expected number of flips required to obtain a tail. The expected number of flips required to obtain a tail is the expectation of a geometric random variable with probability of success $1 - p$ and thus we have that

$$E[N|H] = 1 + \frac{1}{1-p}\,.$$

The addition of the one in the above expression is due to the fact that we were required to performed one flip to determining what the first flip was. In the same way we have

$$E[N|H^c] = 1 + \frac{1}{p}\,.$$

With these two sub-results we have that $E[N]$ is given by

$$E[N] = p + \frac{p}{1-p} + (1-p) + \frac{1-p}{p} = 1 + \frac{p}{1-p} + \frac{1-p}{p}\,.$$

Part (b): We can reason this probability as follows. Since once the outcome of the first coin flip is observed we repeatedly flip our coin as many times as needed to obtain the opposite face we see that we will end our experiment on a head only if the first coin flip is a *tail*. Since this happens with probability $1 - p$ this must also be the probability that the last flip lands heads.

Problem 61

Part (a): Conditioning on N we have that

$$P\{M \le x\} = \sum_{n=1}^{\infty} P\{M \le x|N = n\}P\{N = n\}\,.$$

Now for a geometric random variable N with parameter p we have that $P\{N = n\} = p(1-p)^{n-1}$ for $n \geq 1$ so we have that

$$P\{M \leq x\} = \sum_{n=1}^{\infty} P\{M \leq x | N = n\} p(1-p)^{n-1} .$$

From the discussion in Chapter 6 we see that $P\{M \leq x | N = n\} = F(x)^n$ so the above becomes

$$\begin{aligned} P\{M \leq x\} &= \sum_{n=1}^{\infty} F(x)^n p(1-p)^{n-1} = pF(x) \sum_{n=1}^{\infty} F(x)^{n-1}(1-p)^{n-1} \\ &= pF(x) \sum_{n=0}^{\infty} (F(x)(1-p))^n = pF(x) \frac{1}{1 - F(x)(1-p)} \\ &= \frac{pF(x)}{1 - (1-p)F(x)} . \end{aligned}$$

Part (b): By definition we have $P\{M \leq x | N = 1\} = F(x)$

Part (c): To evaluate $P\{M \leq x | N > 1\}$ we can again condition on N to obtain

$$\begin{aligned} P\{M \leq x | N > 1\} &= \sum_{n=1}^{\infty} P\{M \leq x, N = n | N > 1\} \\ &= \sum_{n=2}^{\infty} P\{M \leq x, N = n | N > 1\} \\ &= \sum_{n=2}^{\infty} P\{M \leq x | N = n, N > 1\} P\{N = n | N > 1\} \\ &= \sum_{n=2}^{\infty} P\{M \leq x | N = n\} P\{N = n | N > 1\} . \end{aligned}$$

Now as before we have that $P\{M \leq x | N = n\} = F(x)^n$ and that

$$\begin{aligned} P\{N = n | N > 1\} &= \frac{P\{N = n, N > 1\}}{P\{N > 1\}} \\ &= \frac{P\{N = n\}}{1 - p} \\ &= \frac{p(1-p)^{n-1}}{1 - p} \\ &= p(1-p)^{n-2} . \end{aligned}$$

Thus we have that

$$
\begin{aligned}
P\{M \leq x | N > 1\} &= \frac{p}{1-p} \sum_{n=2}^{\infty} F(x)^n (1-p)^{n-2} \\
&= \frac{pF(x)^2}{1-p} \sum_{n=0}^{\infty} (F(x)(1-p))^n \\
&= \frac{pF(x)^2}{1-p} \left(\frac{1}{1 - F(x)(1-p)} \right) .
\end{aligned}
$$

Part (d): Conditioning on $N = 1$ and $N > 1$ we have that

$$
\begin{aligned}
P\{M \leq x\} &= P\{M \leq x | N = 1\} P\{N = 1\} + P\{M \leq x | N > 1\} P\{N > 1\} \\
&= F(x)p + \frac{pF(x)^2(1-p)}{1 - F(x)(1-p)} = pF(x) \left[1 + \frac{F(x)(1-p)}{1 - F(x)(1-p)} \right] \\
&= \frac{pF(x)}{1 - F(x)(1-p)} .
\end{aligned}
$$

This is the same as in Part (a)!.

Problem 62

Defining $N(x) = \min\{n : \sum_{i=1}^{n} U_i > x\}$

Part (a): Let $n = 0$ then $P\{N(x) \geq 1\} = 1$ since we must have at least one term in our sum. Lets also derive the expression for the case $n = 1$. We see that $P\{N(x) \geq 2\} = P\{U_1 < x\}$, because the event that we need at least two random draws is equivalent to the event that the first random draw is less than x. This is equivalent to the cumulative distribution function for a uniform random variable so equals $F_U(a) = a$ and therefore

$$
P\{N(x) \geq 2\} = x .
$$

Now lets assume (to be shown by induction) that

$$
P\{N(x) \geq k + 1\} = \frac{x^k}{k!} \quad \text{for} \quad k \leq n ,
$$

We want to compute $P\{N(x) \geq k + 2\}$ which we will do by conditioning on the value U_1. Thus we have

$$
\begin{aligned}
P\{N(x) \geq k + 2\} &= \int_{u_1=0}^{x} P\{N(x) \geq k + 2 | U_1 = u_1\} P\{U_1 = u_1\} du_1 \\
&= \int_{u_1=0}^{x} P\{N(x) \geq k + 2 | U_1 = u_1\} du_1 .
\end{aligned}
$$

To evaluate $P\{N(x) \geq k + 2 | U_1 = u_1\}$ we note that it is the probability that we require $k + 2$ or more terms in our sum to create a sum that is larger than x. Given that the first random variable U_1 is equal to u_1. We know that because we require a at least $k+1$ terms that this value of u_1 must be less than x. This puts the upper limit on the integral of x and we see that the expression $P\{N(x) \geq k+2 | U_1 = u_1\}$ is equivalent to

$$P\{N(x-u_1) \geq k+1\} = \frac{(x-u_1)^k}{k!},$$

by the induction hypothesis. Our integral above becomes

$$\int_0^x \frac{(x-u_1)^k}{k!} du_1 = -\left.\frac{(x-u_1)^{k+1}}{(k+1)!}\right|_0^x = -\left(0 - \frac{x^{k+1}}{(k+1)!}\right) = \frac{x^{k+1}}{(k+1)!}$$

which is what we were trying to prove.

With this expression we can evaluate the expectation of $N(x)$ by using the identity that

$$E[N] = \sum_{n=0}^{\infty} P\{N \geq n+1\},$$

which is proven in Problem 2 of the theoretical exercises. With the expression for $P\{N \geq n+1\}$ above we find that

$$E[N] = \sum_{n \geq 0} \frac{x^n}{n!} = e^x,$$

as expected.

Problem 63 (Cov(X, Y))

Warning: For some reason this solution does not match the answer given in the back of the book. If anyone knows why please let me know. I have not had as much time as I would have liked to go over this problem, careat emptor.

Part (a): With X and Y as suggested we have

$$\text{Cov}(X, Y) = \text{Cov}\left(\sum_{i=1}^{10} X_i, \sum_{j=1}^{3} Y_j\right) = \sum_{i=1}^{10}\sum_{j=1}^{3} \text{Cov}(X_i, Y_j).$$

Here X_i is a Bernoulli random variable specifying if red ball i is drawn or not. Since defined in this way X_i and Y_j are independent and the above factorization is valid. For two Bernoulli random variables we have

$$\text{Cov}(X_i, Y_j) = E[X_i Y_j] - E[X_i]E[Y_j].$$

We have

$$\begin{aligned} E[X_i Y_j] &= P(X_i Y_j) = \frac{10}{30}\frac{8}{29} \\ E[X_i] &= P(X_i) = \frac{10}{30} \\ E[Y_j] &= P(Y_j) = \frac{8}{30}\frac{1}{18}. \end{aligned}$$

Thus $\text{Cov}(X_i, Y_j) = \frac{80}{30}\left(\frac{1}{29} - \frac{1}{30}\right)$.

Part (b): $\text{Cov}(XY) = E[XY] - E[X]E[Y]$. Now to compute $E[X]$, we recognized that X is a hypergeomeric random variable with parameters $N = 18$, $m = 10$, and $n = 12$, so that

$$E[X] = \frac{10(12)}{18} = \frac{20}{3}.$$

To compute $E[Y]$ we recognize that Y is a hypergeometric random variable with parameter $N = 18$, $m = 8$, $n = 12$ so

$$E[Y] = \frac{8(12)}{18} = \frac{16}{3}.$$

Finally to compute $E[XY]$ we condition on X (or Y) as suggested in the book as $E[XY] = E[E[XY|Y]]$. Now

$$E[XY|Y = y] = E[Xy|Y = y] = yE[X|Y = y].$$

The probability $X|Y = y$ is a hypergeometric random variable with parameters $N = 18 - y$, $m = 10$, $n = 12 - y$, for $0 \le y \le 12$ and so has an expectation given by

$$\frac{10(12 - y)}{18 - y}.$$

Thus we have $E[XY|Y] = Y\left(\frac{10(12-Y)}{18-Y}\right)$, so that

$$\begin{aligned} E[XY] &= \sum_Y Y\left(\frac{10(12 - Y)}{18 - Y}\right) P\{Y\} \\ &= \sum_{y=0}^{8} y\left(\frac{10(12 - y)}{18 - y}\right) \frac{\binom{8}{y}\binom{18-8}{12-y}}{\binom{18}{12}} \end{aligned}$$

Problem 64

Part (a): We can compute this expectation by conditioning on the type of light bulb selected. Let the event T_1 be the event that we select the type one light bulb and T_2 be the event that we select the type two light bulb. Then

$$E[X] = E[X|T_1]P(T_1) + E[X|T_2]P(T_2) = \mu_1 p + \mu_2(1 - p).$$

Part (b): Again conditioning on the type of light bulb selected we have

$$E[X^2] = E[X^2|T_1]P(T_1) + E[X^2|T_2]P(T_2).$$

Now for the these Gaussians we have in terms of the variables of the problem that $E[X^2|T_1] = \text{Var}(X|T_1) + E[X]^2 = \sigma_1^2 + \mu_1^2$. So the value of $E[X^2]$ the becomes

$$E[X^2] = p(\sigma_1^2 + \mu_1^2) + (1 - p)(\sigma_1^2 + \mu_1^2).$$

Thus $\text{Var}(X)$ is then given by

$$\begin{aligned}\text{Var}(X) &= E[X^2] - E[X]^2 \\ &= p(\sigma_1^2 + \mu_1^2) + (1-p)(\sigma_1^2 + \mu_1^2) - (\mu_1 p + \mu_2(1-p))^2 \\ &= p(1-p)(\mu_1^2 + \mu_2^2) + p\sigma_1^2 + (1-p)\sigma_2^2 - 2p(1-p)\mu_1\mu_2\,,\end{aligned}$$

after some simplification. Not that this problem can also be solved using the conditional variance formula. The conditional variance formula is given by

$$\text{Var}(X) = E(\text{Var}(X|Y)) + \text{Var}(E[X|Y])\,.$$

Since $E[X|T_1] = \mu_1$, and $E[X|T_1] = \mu_2$, the variance of $E[X|T]$ and the second term in the conditional variance formula is given by

$$\begin{aligned}\text{Var}(E[X|T]) &= E[E[X|T]^2] - E[E[X|T]]^2 \\ &= \mu_1^2 p + \mu_2^2(1-p) - (\mu_1 p + \mu_2(1-p))^2\,.\end{aligned}$$

Also the random variable $\text{Var}(X|Y)$ can be computed by recognizing that

$$\begin{aligned}\text{Var}(X|T_1) &= \sigma_1^2 \quad \text{and} \\ \text{Var}(X|T_2) &= \sigma_2^2\,,\end{aligned}$$

so that

$$E[\text{Var}(X|T)] = \sigma_1^2 p + \sigma_2^2(1-p)\,.$$

Putting all of these pieces together we find that

$$\begin{aligned}\text{Var}(X) &= \sigma_1^2 p + \sigma_2^2(1-p) + \mu_1^2 p + \mu_2^2(1-p) - (\mu_1 p + \mu_2(1-p))^2 \\ &= p\mu_1^2(1-p) + (1-p)p\mu_2^2 + p\sigma_1^2 + (1-p)\sigma_2^2 - 2p(1-p)\mu_1\mu_2,\,,\end{aligned}$$

the same result as before.

Problem 65 (bad winter storms)

We can compute the expectation of the number of storms by conditioning on the type of winter we will have. If we let G be the event that the winter is good and B be the event that the winter is bad then we have (with N the random variable denoting the number of winter storms) the following

$$\begin{aligned}E[N] &= E[N|G]P(G) + E[N|B]P(B) \\ &= 3(0.4) + 5(0.6) = 4.2\,.\end{aligned}$$

To compute the variance we will use the conditional variance formula given by

$$\text{Var}(N) = E[\text{Var}(N|Y)] + \text{Var}(E[N|Y])\,,$$

where Y is the random variable denoting the type of winter. We will compute the first term on the right hand side of this expression first. Since the variances given the type of storm are known i.e.

$$\begin{aligned}\text{Var}(N|Y=G) &= 3 \quad \text{and} \\ \text{Var}(N|Y=B) &= 5\,,\end{aligned}$$

by the fact that a Poisson random variable has equal means and variances. Thus the expectation of these variances can be calculated as

$$E[\text{Var}(N|Y)] = 3(0.4) + 5(0.6) = 4.2\,.$$

Now to compute the second term in the conditional variance formula we recall that

$$\begin{aligned} E[N|Y = G] &= 3 \quad \text{and} \\ E[N|Y = B] &= 5\,, \end{aligned}$$

so that using the definition of the variance, the variance of the random variable $E[N|Y]$ is given by

$$\text{Var}(E[N|Y = G]) = (3 - 4.2)^2(0.4) + (5 - 4.2)^2(0.6) = 0.96\,.$$

Combining these two components we see that

$$\text{Var}(N) = 4.2 + 0.96 = 5.16\,.$$

Problem 66 (our miners variance)

Following the example in the book we can compute $E[X^2]$ in much the same way as in example 5c. By conditioning on the door taken we have that

$$\begin{aligned} E[X^2] &= E[X^2|Y = 1]P\{Y = 1\} + E[X^2|Y = 2]P\{Y = 2\} + E[X^2|Y = 3]P\{Y = 3\} \\ &= \frac{1}{3}(E[X^2|Y = 1] + E[X^2|Y = 2] + E[X^2|Y = 3])\,. \end{aligned}$$

But now we have to compute $E[X^2|Y]$ for the various possible Y values. The easiest to compute is $E[X^2|Y = 1]$ which would equal $3^2 = 9$ since when our miner selects the first door he is able to leave the mine in three hours. The other two expectations are computed using something like a "no memory" property of this problem. As an example if the miner takes the second door $Y = 2$ then after five hours he returns back to the mine exactly where he started. Thus the expectation of X^2, given that he takes the second door, is equal to the expectation of $(5+X)^2$ with no information as to the next door he may take. Mathematically, expressing this we then have

$$\begin{aligned} E[X^2|Y = 2] &= E[(5 + X)^2] \quad \text{and} \\ E[X^2|Y = 3] &= E[(7 + X)^2]\,. \end{aligned}$$

Expanding the quadratic in the above expectations we find that

$$\begin{aligned} E[X^2|Y = 2] &= E[25 + 10X + X^2] = 25 + 10E[X] + E[X^2] = 175 + E[X^2] \\ E[X^2|Y = 3] &= E[49 + 14X + X^2] = 49 + 14E[X] + E[X^2] = 259 + E[X^2]\,. \end{aligned}$$

Using the previously computed result that $E[X] = 15$. Thus when we put these expressions in our expansion of $E[X^2]$ above we find that

$$E[X^2] = \frac{1}{3}(9 + 175 + E[X^2] + 259 + E[X^2])\,,$$

or upon solving for $E[X^2]$ gives $E[X^2] = 443$. We can then easily compute the variance of X. We find that

$$\text{Var}(X) = E[X^2] - E[X]^2 = 443 - 15^2 = 218\,.$$

Problem 67 (gambling with the Kelly strategy)

Let E_n be the expected fortune after n gambles of a gambler who uses the Kelly strategy. Then we are told that $E_0 = x$ (in fact in this case this is his exact fortune i.e. there no expectation). Now we can compute in terms of E_{n-1} by conditioning on whether we win or loose. At time $n-1$ we have a fortune of E_{n-1} and we bet $2p-1$ of this fortune. Thus if we win (which happens with probability p) we will then have $E_{n-1} + (2p-1)E_{n-1}$. While if we loose (which happens with probability $1-p$) we will then have $E_{n-1} - (2p-1)E_{n-1}$. Thus E_n our expected fortune at time n is then given by

$$\begin{aligned} E_n &= (E_{n-1} + (2p-1)E_{n-1})p + (E_{n-1} - (2p-1)E_{n-1})(1-p) \\ &= E_{n-1}p + E_{n-1}(1-p) + E_{n-1}\{(2p-1)p - (2p-1)(1-p)\} \\ &= E_{n-1} + (2p-1)^2 E_{n-1} \\ &= (1 + (2p-1)^2)E_{n-1} \quad \text{for} \quad n = 1, 2, \cdots . \end{aligned}$$

Writing this expression out for $n = 1, 2, \cdots$ and by using induction we see that E_n is given by

$$E_n = (1 + (2p-1)^2)^n E_0 = (1 + (2p-1)^2)^n x \,.$$

Problem 68 (Poisson accidents)

Let E_2 be the event that the person has a number of accidents (per year) given by a Poisson random variable with $\lambda = 2$ and E_3 the event that the person has a number of accidents (again per year) given by a Poisson random variable with $\lambda = 3$. Then the probability a person has k accidents can be computed by conditioning on the type of person someone is i.e. whether they are of E_2 or of E_3 type. We then find (if N is the random variable denoting the number of accidents a person has this year)

$$\begin{aligned} P\{N = k\} &= P\{N = k|E_2\}P(E_2) + P\{N = k|E_3\}P(E_3) \\ &= 0.6\left(\frac{e^{-2}2^k}{k!}\right) + 0.4\left(\frac{e^{-3}3^k}{k!}\right) . \end{aligned}$$

Part (a): Evaluating the above for $k = 0$ we find that

$$P\{N = 0\} = 0.6e^{-2} + 0.4e^{-3} = 0.101 \,.$$

Part (b): Evaluating the above for $k = 3$ we find that

$$P\{N = 3\} = 0.6\left(\frac{e^{-2}2^3}{3!}\right) + 0.4\left(\frac{e^{-3}3^3}{3!}\right) = 0.1978 \,.$$

If we have no accidents in the previous year this information will change the probability that a person is a type E_2 or a type E_3 person. Specifically, if Y_0 is the information/event that our person had no accidents in the previous year, the calculation we now want to evaluate is

$$\begin{aligned} P\{N = k|Y_0\} &= P\{N = k|E_2, Y_0\}P(E_2|Y_0) + P\{N = k|E_3, Y_0\}P(E_3|Y_0) \\ &= P\{N = k|E_2\}P(E_2|Y_0) + P\{N = k|E_3\}P(E_3|Y_0) \,. \end{aligned}$$

Where $P(E_i|Y_0)$ is the probability the person is of "type", E_i, given the information about no accidents. We are also assuming that N is conditionally independent of Y_0 given E_i i.e. $P\{N = k|E_i, Y_0\} = P\{N = k|E_i\}$. We can compute the conditional probabilities $P(E_i|Y_0)$ with Bayes' rule. We find

$$P(E_2|Y_0) = \frac{P(Y_0|E_2)P(E_2)}{P(Y_0|E_2)P(E_2) + P(Y_0|E_3)P(E_3)},$$

and the same type formula for $P(E_3|Y_0)$. Now we have computed the denominator of the above expression in Part (a) above. Thus we find that

$$\begin{aligned} P(E_2|Y_0) &= \frac{(e^{-2})(0.6)}{P\{N = 0\}} = 0.803 \\ P(E_3|Y_0) &= \frac{(e^{-3})(0.4)}{P\{N = 0\}} = 0.196\,. \end{aligned}$$

With these two expressions we can calculate the probability we obtain any number of accidents in the next year. Incorporating the information that the event Y_0 happened that $P\{N = k\}$ is given by

$$P\{N = k|Y_0\} = 0.803\left(\frac{e^{-2}2^k}{k!}\right) + 0.196\left(\frac{e^{-3}3^k}{k!}\right).$$

Evaluating this expression for $k = 3$ we find that $P\{N = 3|Y_0\} = 0.18881$. The information that the our person had no accidents in the previous year reduced the probability that they will have three accidents this year (computed above) as one would expect. These calculations can be found in the file `chap_7_prob_68.m`.

Problem 70

Part (a): We want to calculate $P\{F_1 = H\}$ or the probability that the first flip is heads. We will do this by conditioning on the coin that is chosen. Let $P\{F_1 = H|C = p\}$ be the probability the first flip is heads given that the chosen coin has p as its probability of heads. Then

$$P\{F_1 = H\} = \int_0^1 P\{F_1 = H|C = p\}P\{C = p\}dp\,.$$

Since we are assuming a uniform distribution of probabilities for our coins we have that the above is given by

$$\int_0^1 P\{F_1 = H|C = p\}dp\,.$$

Now $P\{F_1 = H|C = p\} = p$ so the above becomes

$$\int_0^1 pdp = \frac{1}{2}\,.$$

Part (b): In this case let E be the event that the first two flips are both heads. Then in exactly the same way as in Part (a) we have

$$P\{E\} = \int_0^1 P\{E|C = p\}dp\,.$$

Now let $P\{E|C = p\} = p^2$ so that the above becomes $\frac{1}{3}$.

Problem 71

In exactly the same way as for Problem 70 let E be the event that i heads occur given that the coin selected has a probability of landing heads of p. Then conditioning on this probability we have that

$$P\{E\} = \int_0^1 P\{E|C = P\}dp\,.$$

But $P\{E|C = p\} = \left(\begin{array}{c} n \\ i \end{array}\right) p^i(1-p)^{n-i}$ and we have that

$$P\{E\} = \int_0^1 \left(\begin{array}{c} n \\ i \end{array}\right) p^i(1-p)^{n-i}dp\,.$$

Remembering the definition of the Beta function and the hint provided in the book we see that

$$P\{E\} = \frac{n!}{i!(n-i)!}\left(\frac{i!(n-i)!}{(n+1)!}\right) = \frac{1}{n+1}\,,$$

as claimed.

Problem 72

Again following the framework provide in Problems 70 and 71 we can calculate these probabilities by conditioning on the coin selected (and given that its corresponding probability of heads is p). We have

$$P\{N \geq i\} = \int_0^1 P\{N \geq i|C = p\}dp\,.$$

Now given the coin we are considering has probability p of obtaining heads

$$P\{N \geq i|C = p\} = 1 - P\{N < i|C = p\} = 1 - \sum_{n=1}^{i-1} P\{N = n|C = p\}\,.$$

Where $P\{N = n|C = p\}$ is the probability that our first head appears on flip n. The random variable N is geometric so we know that $P\{N = n|C = p\} = p(1-p)^{n-1}$ for $n = 1, 2, \cdots$ and the above becomes

$$1 - \sum_{n=1}^{i-1} p(1-p)^{n-1}\,.$$

Integrating this with respect to p we have that

$$\begin{aligned} P\{N \geq i\} &= \int_0^1 \left(1 - \sum_{n=1}^{i-1} p(1-p)^{n-1}\right) dp \\ &= 1 - \sum_{n=1}^{i-1} \int_0^1 p(1-p)^{n-1} dp \\ &= 1 - \sum_{n=1}^{i-1} \frac{1!(n-1)!}{(n+1)!} \\ &= 1 - \sum_{n=1}^{i-1} \frac{1}{n(n+1)} \,. \end{aligned}$$

Using partial fractions to evaluate the sum above we have that

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1} \,,$$

from which we recognize that the above sum is of a telescoping type so that we find that

$$\sum_{n=1}^{i-1} \frac{1}{n(n+1)} = \left(\frac{1}{1} - \frac{1}{i}\right) \,.$$

Thus in total we find that

$$P\{N \geq i\} = 1 - \left(1 - \frac{1}{i}\right) = \frac{1}{i} \,.$$

Part (b): We could follow the same procedure as in Part (a) by conditioning on the coin selected and noting that $P\{N = i | C = p\} = p(1-p)^{i-1}$ or we could simply notice that

$$\begin{aligned} P\{N = i\} &= P\{N \geq i\} - P\{N \geq i+1\} \\ &= \frac{1}{i} - \frac{1}{i+1} = \frac{1}{i(i+1)} \,. \end{aligned}$$

Part (c): Given the probabilities computed in Part (b) the expression $E[N]$ is easily computed

$$E[N] = \sum_{n=1}^{\infty} n \left(\frac{1}{n(n+1)}\right) = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty$$

Problem 75

From the chart in the book we see that X is a Poisson random variable with parameters $\lambda = 2$ and that Y is a Binomial random variable with parameters $(n, p) = (10, \frac{3}{4})$.

Part (a): The moment generating function for $X+Y$ is the product of the moment generating function for X and Y. Thus

$$M_{X+Y}(t) = \exp\{2e^t - 2\}\left(\frac{3}{4}e^t + \frac{1}{4}\right)^{10},$$

Then $P\{X+Y=2\}$ is the third term in the Taylor expansion centered on e^t i.e.

$$P\{X+Y=2\} = \left.\frac{d^2}{d(e^t)^2}M_{X+Y}(t)\right|_{e^t=0}$$

Computing the first derivative of the above (with respect to the variable is e^t) we find

$$\begin{aligned}\frac{d}{d(e^t)}M_{X+Y} &= 2\exp\{2e^t-2\}\left(\frac{3}{4}e^t+\frac{1}{4}\right)^{10} \\ &+ 10\exp\{2e^t-2\}\left(\frac{3}{4}e^t+\frac{1}{4}\right)^{9}\left(\frac{3}{4}\right).\end{aligned}$$

So that the second derivative is given by

$$\begin{aligned}\frac{d^2}{d(e^t)^2}M_{X+Y} &= 4\exp\{2e^t-2\}\left(\frac{3}{4}e^t+\frac{1}{4}\right)^{10} \\ &+ 20\left(\frac{3}{4}\right)\exp\{2e^t-2\}\left(\frac{3}{4}e^t+\frac{1}{4}\right)^{9} \\ &+ 20\left(\frac{3}{4}\right)\exp\{2e^t-2\}\left(\frac{3}{4}e^t+\frac{1}{4}\right)^{9} \\ &+ 90\left(\frac{3}{4}\right)^2\exp\{2e^t-2\}\left(\frac{3}{4}e^t+\frac{1}{4}\right)^{8}.\end{aligned}$$

Evaluating this expression at $e^t = 0$ gives

$$\frac{d^2M_{X+Y}}{d(e^t)^2} = 4e^{-2}\left(\frac{1}{4}\right)^{10} + 40\left(\frac{3}{4}\right)e^{-2}\left(\frac{1}{4}\right)^{9} + 90\left(\frac{3}{4}\right)^2 e^{-2}\left(\frac{1}{4}\right)^{8},$$

which can easily be further evaluated.

Part (b): Now $P\{XY=0\}$ can be computed by summing the probabilities of the mutually exclusive individual terms that could result in the product XY being zero. We find

$$P\{XY=0\} = P\{(X=0,Y=0)\} + P\{(X=0,Y\neq 0)\} + P\{(X\neq 0,Y=0)\}.$$

Now the first of these is given by

$$P\{(X=0,Y=0)\} = e^{-2}\left(\frac{1}{4}\right)^{10} + e^{-2}\left(1-\left(\frac{1}{4}\right)^{10}\right) + (1-e^{-2})\left(\frac{1}{4}\right)^{10}.$$

Part (c): Now $E[XY]=E[X]E[Y]$ since X and Y are independent. Since X is a Poisson random variable we know that $E[X]=2$ and since Y is a binomial random variable we know that $E[Y]=10\left(\frac{3}{4}\right)=\frac{15}{2}$. So that

$$E[XY] = 15.$$

Chapter 7: Theoretical Exercises

Problem 2

Following the hint we have that

$$E[|X-a|] = \int |x-a|f(x)dx = -\int_{-\infty}^{a}(x-a)f(x)dx + \int_{a}^{\infty}(x-a)f(x)dx\,.$$

Then taking the derivative of this expression with respect to a we have that

$$\begin{aligned}\frac{dE[|X-a|]}{da} &= 0 - \int_{-\infty}^{a}(-1)f(x)dx - 0 - \int_{a}^{\infty} f(x)dx \\ &= \int_{-\infty}^{a} f(x)dx - \int_{a}^{\infty} f(x)dx\,,\end{aligned}$$

Setting this expression equal to zero gives that a must satisfy

$$\int_{-\infty}^{a} f(x)dx = \int_{a}^{\infty} f(x)dx\,.$$

Which is the exact definition of the median of the distribution $f(\cdot)$. That is a is the point where one half of the probability is to left of a and where one half of the probability is to the right of a.

Problem 5 (sums of the probabilities that at least one event takes place)

Let X_i be defined as

$$X_i = \begin{cases} 1 & A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}\,,$$

Then X, the total number of events A_i, that occur can be written as $X = \sum_{i=1}^{n} X_i$. Now since X_i are Bernoulli random variables $E[X_i] = P(A_i)$ so we have one expression for $E[X]$ given by

$$E[X] = \sum_{i=1}^{n} P(A_i)\,.$$

Recall that C_k is the event that at least k of the A_i events occur. Thus $P(C_k) = P\{X \geq k\}$ and we note that

$$\begin{aligned}
\sum_{k=1}^{n} P(C_k) &= \sum_{k=1}^{n} P\{X \geq k\} \\
&= P\{X \geq n\} + P\{X \geq n-1\} + \cdots P\{X \geq 2\} + P\{X \geq 1\} \\
&= P\{X = n\} \\
&+ P\{(X = n) \cup (X = n-1)\} \\
&+ P\{(X = n) \cup (X = n-1) \cup (X = n-2)\} \\
&+ \cdots \\
&+ P\{(X = n) \cup (X = n-1) \cup \cdots \cup (X = 2) \cup (X = 1)\} \\
&= nP\{X = n\} + (n-1)P\{X = n-1\} + \cdots + 2P\{X = 2\} + 1P\{X = 1\} = E[X]\,,
\end{aligned}$$

since that is the definition of $E[X]$.

Problem 6 (the integral of the complement of the distribution function)

We desire to prove that

$$E[X] = \int_0^\infty P\{X > t\} dt\,.$$

Following the hint in the book define the random variable $X(t)$ as

$$X(t) = \begin{cases} 1 & \text{if} \quad t < X \\ 0 & \text{if} \quad t \geq X \end{cases}$$

Then integrating the variable $X(t)$ we see that

$$\int_0^\infty X(t)dt = \int_0^X 1dt = X\,.$$

Thus taking the expectation of both sides we have

$$E[X] = E\left[\int_0^\infty X(t)dt\right]\,.$$

This allows us to use the assumed identity that we can pass the expectation inside the integration as

$$E\left[\int_0^\infty X(t)dt\right] = \int_0^\infty E[X(t)]dt\,,$$

so applying this identity to the expression we have for $E[X]$ above we see that $E[X] = \int_0^\infty E[X(t)]dt$. From the definition of $X(t)$ we have that $E[X(t)] = P\{X > t\}$ and we then finally obtain the fact that

$$E[X] = \int_0^\infty P\{X > t\} dt\,,$$

as we were asked to prove.

Problem 7 (stochastically larger)

Part (a): When X is a nonnegative random variable $E[X]$ can be written as $E[X] = \int_0^\infty P\{X > t\}dt$. Then if X is stochastically larger than Y using this and the definition of stochastically larger we have that

$$E[X] = \int_0^\infty P\{X > t\}dt \geq \int_0^\infty P\{Y > t\}dt = E[Y]\,,$$

where Y is another nonnegative random variable.

Problem 10 (the expectation of a sum of random variables)

We begin by defining $R(k)$ to be

$$R(k) \equiv E\left[\frac{\sum_{i=1}^k X_i}{\sum_{i=1}^n X_i}\right]\,.$$

Then we see that $R(k)$ satisfies a recursive expression given by

$$R(k) - R(k-1) = E\left[\frac{X_k}{\sum_{i=1}^n X_i}\right] \quad \text{for} \quad 2 \leq k \leq n\,.$$

To further simplify this we would like to evaluate the expectation on the right hand side of the above. Now by the assumed independence of all X_i's the expectation on the right hand-side of the above is *independent* of k, and is a constant C. Thus it can be evaluated by considering

$$\begin{aligned} 1 &= E\left[\frac{\sum_{k=1}^n X_k}{\sum_{i=1}^n X_i}\right] \\ &= \sum_{k=1}^n E\left[\frac{X_k}{\sum_{i=1}^n X_i}\right] \\ &= nC\,. \end{aligned}$$

Which when we solve for C gives $C = 1/n$ or in terms of the original expectations

$$E\left[\frac{X_k}{\sum_{i=1}^n X_i}\right] = \frac{1}{n} \quad \text{for} \quad 1 \leq k \leq n\,.$$

Thus using our recursive expression $R(k) = R(k-1) + 1/n$, we see that since

$$R(1) = E\left[\frac{X_1}{\sum_{i=1}^n X_i}\right] = \frac{1}{n}\,,$$

that

$$R(2) = \frac{1}{n} + \frac{1}{n} = \frac{2}{n}\,.$$

Continuing our iterations in this way we find that

$$R(k) = E\left[\frac{\sum_{i=1}^k X_i}{\sum_{i=1}^n X_i}\right] = \frac{k}{n} \quad \text{for} \quad 1 \leq k \leq n\,.$$

Problem 11

Let X_i denote the Bernoulli indicator random variable that is one if outcome i never occurs in all n trials and is zero if it does occur. Then

$$X = \sum_{i=1}^{r} X_i \,.$$

The expected number of outcomes that never occur is given by $E[X] = \sum_{i=1}^{r} E[X_i]$. But $E[X_i] = P(X_i) = (1 - P_i)^n$, since with probability $1 - P_i$ the ith event won't happen with one draw. Thus

$$E[X] = \sum_{i=1}^{r} (1 - P_i)^n \,.$$

To find that maximum or minimum of this expression with respect to the P_i we can't simply take the derivative of $E[X]$ and set it equal to zero because that won't enforce the constraint that $\sum_{i=1}^{r} P_i = 1$. To enforce this constraint we introduce a Lagrangian multiplier λ and a Lagrangian L defined by

$$L \equiv \sum_{i=1}^{r} (1 - P_i)^n + \lambda \left(\sum_{i=1}^{r} P_i - 1 \right) \,.$$

Then taking the derivatives of L with respect to P_i *and* λ and setting all of these expressions equal to zero we get

$$\begin{aligned} \frac{\partial L}{\partial P_i} &= n(1 - P_i)^{n-1}(-1) + \lambda = 0 \quad \text{for} \quad i = 1, 2, \cdots, r \\ \frac{\partial L}{\partial \lambda} &= \sum_{i=1}^{r} P_i - 1 = 0 \,. \end{aligned}$$

It is this system that we solve for $P_1, P_2, \cdots, P_r$ and λ. We can solve the first equation for P_i in terms of λ and obtain the following

$$P_i = 1 - \left(\frac{\lambda}{n} \right)^{\frac{1}{n-1}} \,.$$

When this is put in the constraint equation $\frac{\partial L}{\partial \lambda} = 0$ gives

$$\sum_{i=1}^{r} \left(1 - \left(\frac{\lambda}{n} \right)^{\frac{1}{n-1}} \right) - 1 = 0 \,.$$

Solving this for λ gives

$$\lambda = n \left(1 - \frac{1}{r} \right)^{n-1} \,.$$

Putting this value of into the expression we derived earlier for $P_i = \frac{1}{r}$. We can determine if this solution is a minimum for $E[X]$ by computing the second derivative of this expression.

Specifically

$$
\begin{aligned}
\frac{\partial E[X]}{\partial P_i} &= -n(1-P_i)^{n-1} \\
\frac{\partial^2 E[X]}{\partial P_i \partial P_j} &= n(n-1)(1-P_i)^{n-2}\delta_{ij} \,.
\end{aligned}
$$

So that the matrix $\frac{\partial^2 E[X]}{\partial P_i \partial P_j}$ is diagonal with positive entries and is therefore positive definite. Thus the values $P_i = \frac{1}{r}$ corresponds to a minimum of $E[X]$.

Problem 12

Part (a): Let I_i be an indicator random variable that is one if trial i results in a success and is zero if trial i results in a failure. Then defining $X = \sum_{i=1}^{n} I_i$ we see that X represents the random variable that denotes the total number of successes. Then

$$
E[X] = \sum_{i=1}^{n} E[I_i] = \sum_{i=1}^{n} P(I_i = 1) = \sum_{i=1}^{n} P_i \,.
$$

Part (b): Since $\left(\begin{array}{c} X \\ 2 \end{array} \right)$ is the number of paired events that occur, we have that

$$
\left(\begin{array}{c} X \\ 2 \end{array} \right) = \sum_{i<j} I_i I_j \,.
$$

Taking the expectation of both sides gives

$$
E[\left(\begin{array}{c} X \\ 2 \end{array} \right)] = E[\frac{X(X-1)}{2}] = \sum_{i<j} P\{I_i = 1, I_j = 1\} = \sum_{i<j} P_i P_j \,,
$$

by using independence of the events $I_i = 1$ and $I_j = 1$. Expanding the quadratic in the expectation on the left hand side we find that

$$
E[X^2] - E[X] = 2 \sum_{i<j} P_i P_j \,,
$$

so that $E[X^2]$ is given by

$$
E[X^2] = \sum_{i=1}^{n} P_i + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} P_i P_j \,.
$$

with the definition that $\sum_{j=1}^{0} (\cdot) = 0$.

Using these expressions the variance of X is given by

$$\begin{aligned}\text{Var}(X) &= E[X^2] - E[X]^2 \\ &= \sum_{i=1}^{n} P_i + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} P_i P_j - \left(\sum_{i=1}^{n} P_i\right)^2 \\ &= \sum_{i=1}^{n} P_i + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} P_i P_j - \sum_{i=1}^{n} P_i^2 - 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} P_i P_j \\ &= \sum_{i=1}^{n} P_i - \sum_{i=1}^{n} P_i^2 .\end{aligned}$$

The independence assumption makes no difference in Part (a) but in Part (b) to evaluate the probability $P\{I_i = 1, I_j = 1\}$ we explicitly invoked independence.

Problem 13 (record values)

Part (a): Let R_j be an indicator random variable denoting whether or not the j-th random variable (from n) is a record value. This is that $R_j = 1$ if and only if X_j is a record value i.e. $X_j \geq X_i$ for all $1 \leq i \leq j$, and X_j is zero otherwise. Then the number N of record values is given by summing up these indicator

$$N = \sum_{j=1}^{n} R_j .$$

Taking the expectation of this expression we find that

$$E[N] = \sum_{j=1}^{n} E[R_j] = \sum_{j=1}^{n} P\{R_j\} .$$

Now $P\{R_j\}$ is the probability that X_j is the maximum from among all X_i samples where $1 \leq i \leq j$. Since each X_i is equally likely to be the maximum we have that

$$P\{R_j\} = P\{X_j = \max_{1\leq i \leq j}(X_i)\} = \frac{1}{j} ,$$

and the expected number of record values is given by

$$E[N] = \sum_{j=1}^{n} \frac{1}{j} ,$$

as claimed.

Part (b): From the discussion in the text if N is a random variable denoting the number of record values that occur then we have

$$\binom{N}{2} = \sum_{i<j} R_i R_j .$$

Thus taking the expectation and expanding the expression $\binom{N}{2}$ in the above we have

$$E[N^2 - N] = E\left[2\sum_{i<j} R_i R_j\right] = 2\sum_{i<j} P(R_i, R_j)\,.$$

Now $P(R_i, R_j)$ is the probability that X_i and X_j are record values. Since there is no constraint on R_j if R_i is a record value this probability is given by

$$P(R_i, R_j) = \frac{1}{j}\frac{1}{i}\,.$$

Thus we have that

$$\begin{aligned} E[N^2] &= E[N] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{1}{j}\frac{1}{i} \\ &= \sum_{j=1}^{n}\frac{1}{j} + 2\sum_{i=1}^{n-1}\frac{1}{i}\sum_{j=i+1}^{n}\frac{1}{j}\,, \end{aligned}$$

so that the variance is given by

$$\begin{aligned} \mathrm{Var}(N) &= E[N^2] - E[N]^2 \\ &= \sum_{j=1}^{n}\frac{1}{j} + 2\sum_{i=1}^{n-1}\frac{1}{i}\sum_{j=i+1}^{n}\frac{1}{j} - \left(\sum_{j=1}^{n}\frac{1}{j}\right)^2 \\ &= \sum_{j=1}^{n}\frac{1}{j} + \left(2\sum_{i=1}^{n-1}\frac{1}{i}\sum_{j=i+1}^{n}\frac{1}{j}\right) - \sum_{j=1}^{n}\frac{1}{j^2} - \left(2\sum_{i=1}^{n-1}\frac{1}{i}\sum_{j=i+1}^{n}\frac{1}{j}\right) \\ &= \sum_{j=1}^{n}\frac{1}{j} - \sum_{j=1}^{n}\frac{1}{j^2}\,. \end{aligned}$$

where we have used the fact that $(\sum_i a_i)^2 = \sum_i a_i^2 + 2\sum_{i<j} a_i a_j$, thus

$$\mathrm{Var}(N) = \sum_{j=1}^{n}\frac{1}{j} - \frac{1}{j^2} = \sum_{j=1}^{n}\frac{j-1}{j^2}\,,$$

as claimed.

Problem 14

We begin by first computing the variance of the number of coupons needed to amass a full set. Following Example 2i from the book the total number of coupons that are collected, X, can be decomposed as sum of random variables X_i which are recognized as the number of *additional* coupons needed after i distinct types have been obtained to obtain a new type. Now

$$\mathrm{Var}\left(\sum_{i=0}^{N-1} X_i\right) = \sum_{i=0}^{N-1}\mathrm{Var}(X_i) - 2\sum_{i<j}\mathrm{Cov}(X_i, X_j)\,.$$

Here $\text{Var}(X_i)$ is the variance of a geometric random variable with parameter $\frac{N-i}{N}$. This is given by

$$\frac{1-\left(\frac{N-i}{N}\right)}{\left(\frac{N-i}{N}\right)^2} = \frac{N^2}{(N-i)^2}\left(\frac{N-N+i}{N}\right) = \frac{Ni}{(N-i)^2}\,.$$

Since X_i and X_j are pairwise independent, introducing the value of X_i does not affect the value of X_j. Thus

$$\text{Var}(X) = \sum_{i=0}^{N-1}\frac{Ni}{(N-i)^2}\,,$$

as we were to show.

Problem 15

Part (a): Define X_i to be an indicator random variable such that if trial i is a success then $X_i = 1$ otherwise $X_i = 0$. Then if X is a random variable representing the number of successes from all n trials we have that

$$X = \sum_i X_i\,,$$

taking the expectation of both sides we find that $E[X] = \sum_i E[X_i] = \sum_i P_i$. Thus an expression for the mean μ is given by

$$\mu = \sum_i P_i\,.$$

Part (b): Using the result from the book we have that

$$\left(\begin{array}{c} X \\ 2 \end{array}\right) = \sum_{i<j} X_i X_j\,,$$

so that taking the expectation of the above gives

$$E[\left(\begin{array}{c} X \\ 2 \end{array}\right)] = \frac{1}{2}E[X^2 - X] = \sum_{i<j} E[X_i X_j]\,.$$

But the expectation of $X_i X_j$ is given by (using independence of the trials X_i and X_j) $E[X_i X_j] = P\{X_i X_j\} = P\{X_i\}P\{X_j\}$. Thus the above expectation becomes

$$E[X^2] = E[X] + 2\sum_{i<j} P_i P_j = \mu + 2\sum_{i=1}^{n-1} P_i \sum_{j=i+1}^{n} P_j\,.$$

From which we can compute the variance of X as

$$\begin{aligned}
\text{Var}(X) &= E[X^2] - E[X]^2 \\
&= \mu + 2\sum_{i=1}^{n-1} P_i \sum_{j=i+1}^{n} P_j - \left(\sum_{i=1}^{n} P_i\right)^2 \\
&= \mu + 2\sum_{i=1}^{n-1} P_i \sum_{j=i+1}^{n} P_j - \sum_{i=1}^{n} P_i^2 - 2\sum_{i=1}^{n-1} P_i \sum_{j=i+1}^{n} P_j \\
&= \sum_{i=1}^{n} P_i(1-P_i)\,.
\end{aligned}$$

To find the values of P_i that maximize this variance we use the method of Lagrange multipliers. Consider the following Lagrangian

$$L = \sum_{i=1}^{n} P_i(1-P_i) + \lambda\left(\sum_{i=1}^{n} P_i - 1\right)\,.$$

Taking the derivatives of this expression with respect to P_i and λ gives

$$\begin{aligned}
\frac{\partial L}{\partial P_i} &= 1 - P_i - P_i + \lambda \quad \text{for} \quad 1 \le i \le n \\
\frac{\partial L}{\partial \lambda} &= \sum_{i=1}^{n} P_i - 1\,.
\end{aligned}$$

The first equation gives for P_i (in terms of λ) the expression that $P_i = \frac{1+\lambda}{2}$ which when put into the second constraint gives

$$\lambda = \frac{2}{n} - 1 = \frac{2-n}{n}\,.$$

Which means that

$$P_i = \frac{1}{n}\,.$$

To determine if this maximizes or minimizes the functional $\text{Var}(X)$ we need to consider the second derivative of the $\text{Var}(X)$ expression, i.e.

$$\frac{\partial^2 \text{Var}(X)}{\partial P_i \partial P_j} = -2\delta_{ij}\,,$$

with δ_{ij} the Kronecker delta. Thus the matrix of second derivatives is negative definite implying that our solutions $P_i = \frac{1}{n}$ will *maximize* the variance.

Part (c): To select a choice of P_i's that minimizes this variance we note that $\text{Var}(X) = 0$ if $P_i = 0$ or $P_i = 1$ for every i. In this case the random variable X is a constant.

Problem 17

Define the random variable Y as $Y \equiv \lambda X_1 + (1-\lambda)X_2$. Then the variance of Y is given by

$$\begin{aligned}
\text{Var}(Y) &= \text{Var}(\lambda X_1 + (1-\lambda)X_2) \\
&= \lambda^2 \text{Var}(X_1) + (1-\lambda)^2 \text{Var}(X_2) \\
&+ 2\sum_{i<j} \text{Cov}(\lambda X_i, (1-\lambda)X_j)\,.
\end{aligned}$$

Since X_1 and X_2 are independent their covariance is zero so the above becomes

$$\text{Var}(Y) = \lambda^2 \sigma_1^2 + (1-\lambda)^2 \sigma_2^2\,.$$

To make this variance as small as possible we desire to minimize this function with respect to λ. Taking the derivative of this expression with respect to λ and setting it equal to zero gives

$$2\lambda\sigma_1^2 + 2(1-\lambda)(-1)\sigma_2^2 = 0\,,$$

which when we solve for λ gives the following

$$\lambda = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\,.$$

A second derivative gives the expression $2\sigma_1^2 + 2\sigma_2^2$ a positive quantity and shows that at this value of λ $E[Y]$ is indeed a minimum. This value of λ weights the samples X_1 and X_2 explicitly as

$$Y = \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right) X_1 + \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right) X_2\,,$$

which we see is in *inverse* proportion to the variance of the individual X_i, so if X_1 has a small variance we weight the value of X_2 less than that of X_1.

Problem 18

Part (a): The distribution of $N_i + N_j$ is binomial with probability of success given by $p_i + p_j$.

Part (b): Since a binomial distribution has a variance given by npq we have that

$$\begin{aligned}
\text{Var}(N_i) &= mp_i(1-p_i) \\
\text{Var}(N_j) &= mp_j(1-p_j) \\
\text{Var}(N_i + N_j) &= m(p_i + p_j)(1 - p_i - p_j)\,.
\end{aligned}$$

So that the expression

$$\text{Var}(N_i + N_j) = \text{Var}(N_i) + \text{Var}(N_j) + 2\text{Cov}(N_i, N_j)\,,$$

becomes

$$m(p_i + p_j)(1 - p_i - p_j) - mp_i(1-p_i) - mp_j(1-p_j) = 2\text{Cov}(N_i, N_j)\,.$$

This simplifies to

$$\text{Cov}(N_i, N_j) = -mp_ip_j\,,$$

as claimed.

Problem 19

Expanding the given expression we have that

$$\begin{aligned}\text{Cov}(X+Y,X-Y) &= \text{Cov}(X,X) - \text{Cov}(X,Y) + \text{Cov}(Y,X) - \text{Cov}(Y,Y) \\ &= \text{Cov}(X,X) - \text{Cov}(Y,Y)\,.\end{aligned}$$

If X and Y are identically distributed then $\text{Cov}(X,X) = \text{Cov}(Y,Y)$ and the above expression is zero.

Problem 20

To solve this problem we will use the definition of conditional variance which is defined by

$$\text{Cov}(X,Y|Z) = E[(X - E[X|Z])(Y - E[Y|Z])]\,.$$

Part (a): By expanding the expression inside the expectation above we have

$$(X - E[X|Z])(Y - E[Y|Z]) = XY - XE[Y|Z] - YE[X|Z] + E[X|Z]E[Y|Z]\,.$$

Then taking the expectation (given Z) i.e. $E[\cdot|Z]$ of the above we find that

$$\begin{aligned}\text{Cov}(X,Y|Z) &= E[XY|Z] - E[XE[Y|Z]|Z] - E[YE[X|Z]|Z] + E[X|Z]E[Y|Z] \\ &= E[XY|Z] - E[Y|Z]E[X|Z] - E[Y|Z]E[X|Z] + E[X|Z]E[Y|Z] \\ &= E[XY|Z] - E[X|Z]E[Y|Z]\end{aligned}$$

Part (b): Considering the expectation with respect to Z of the expression derived in Part (a) we have that

$$E[\text{Cov}(X,Y|Z)] = E[E[XY|Z]] - E[E[X|Z]E[Y|Z]]\,.$$

Since $E[E[XY|Z]] = E[XY]$ we can add and subtract $E[X]E[Y]$ to the right hand side of the above to get

$$E[\text{Cov}(X,Y|Z)] = E[XY] - E[X]E[Y] + E[X]E[Y] - E[E[X|Z]]E[Y|Z]\,.$$

Since $\text{Cov}(X,Y) = E[XY] - E[X]E[Y]$ and $E[X] = E[E[X|Z]]$ (similarly for $E[Y]$) the above becomes

$$E[\text{Cov}(X,Y|Z)] = \text{Cov}(X,Y) + E[E[X|Z]]E[E[Y|Z]] - E[E[X|Z]E[Y|Z]]\,.$$

Finally defining $\text{Cov}(E[X|Z],E[Y|Z])$ as

$$E[E[X|Z]E[Y|Z]] - E[E[X|Z]]E[E[Y|Z]]\,,$$

we see that the above gives for $\text{Cov}(X,Y)$ the following

$$\text{Cov}(X,Y) = E[\text{Cov}(X,Y|Z)] + \text{Cov}(E[X|Z],E[X|Z])\,.$$

Part (c): If $X = Y$, the expression in Part (b) becomes

$$\begin{aligned}\text{Var}(X) &= E[\text{Var}(X|Z)] + \text{Cov}(E[X|Z],E[X|Z]) \\ &= E[\text{Var}(X|Z)] + \text{Var}(E[X|Z])\,.\end{aligned}$$

Problem 21

Part (a): By expanding the definition of the variance we have that $\text{Var}(X_{(i)}) = E[X_{(i)}^2] - E[X_{(i)}]^2$. Using the definition of expectation we can compute each of these expectations. By the definition of $E[X_{(i)}]$ we have that

$$E[X_{(i)}] = \frac{n!}{(i-1)!(n-i)!} \int_0^1 x^i (1-x)^{n-i} dx \,.$$

Remembering the definition of the Beta function

$$B(a, b) \equiv \int_0^1 x^{a-1}(1-x)^{b-1} dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \,,$$

and the fact that $\Gamma(k) = (k-1)!$ when k is an integer, we find that the expectation of $X_{(i)}$ is given by

$$\begin{aligned} E[X_{(i)}] &= \frac{n!}{(i-1)!(n-i)!} \left(\frac{\Gamma(i+1)\Gamma(n-i+1)}{\Gamma(i+n-i+2)} \right) \\ &= \frac{n!}{(i-1)!(n-i)!} \left(\frac{i!(n-i)!}{(n+1)!} \right) = \frac{i}{n+1} \,. \end{aligned}$$

In the same way we have

$$\begin{aligned} E[X_{(i)}^2] &= \frac{n!}{(i-1)!(n-i)!} \int_0^1 x^{i+1}(1-x)^{n-i} dx \\ &= \frac{n!}{(i-1)!(n-i)!} \left(\frac{(i+1)!(n-i)!}{(n+2)!} \right) \\ &= \frac{i(i+1)}{(n+1)(n+2)} \,. \end{aligned}$$

Combining these two we have finally that

$$\begin{aligned} \text{Var}(X_{(i)}) &= \frac{i(i+1)}{(n+1)(n+2)} - \frac{i^2}{(n+1)^2} \\ &= \frac{i(n+1-i)}{(n+1)^2(n+2)} \quad \text{for} \quad i = 1, 2, \cdots, n \end{aligned}$$

Part (b): Since the denominator of $\text{Var}(X_{(i)})$ for all i is a constant, to minimize (or maximize) this expression we can study the numerator $i(n+1-i)$. Then the minimum/maximum for this expression occurs at $i = 1$ or n or the index where $\frac{d}{di}(i(n+1-i)) = 0$. Taking this derivative we find that the first order necessary condition is

$$n + 1 - i - i = 0 \,,$$

or that $i = \frac{n+1}{2}$. Note this is effectively the sample median, i.e. if n is odd this is an integer otherwise this is non-integer. Since the second derivative of this expression is given by

$$\frac{d^2(i(n+1-i))}{di^2} = -2 < 0 \,,$$

The value of $\text{Var}(X_{(i)})$ at $i = \frac{n+1}{2}$ corresponds to a local maximum and has a value given by

$$\left(\frac{n+1}{2}\right)\left(n+1-\left(\frac{n+1}{2}\right)\right) = \left(\frac{n+1}{2}\right)^2 .$$

This is to be compared to the value of the numerator $i(n+1-i)$ when $i = 1$ or n both of which equal n. Thus $\text{Var}(X_{(1)}) = \text{Var}(X_{(n)})$ and the maximum and minimum statistic ($i = 1$ and $i = n$) have the smallest variance while the "median" element $i = \frac{i+1}{2}$ (or the nearest integer) has the largest variance.

Problem 22

We begin by remembering the definition of the correlation coefficient between two random variables X and Y

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} .$$

Since $Y + a + bX$ we have that $\text{Var}(Y) = \text{Var}(a + bX) = b^2\text{Var}(X)$, and $\text{Cov}(X, Y) = \text{Cov}(X, X) = b\text{Var}(X)$. With these ρ becomes

$$\rho(X, Y) = \frac{b\text{Var}(X)}{\sqrt{\text{Var}(X)}|b|\sqrt{\text{Var}(X)}} = \frac{b}{|b|} = \begin{cases} -1 & b < 0 \\ +1 & b > 0 \end{cases} .$$

Problem 23

To compute $\rho(Y, Z)$ we need to compute $\text{Cov}(Y, Z)$. Since $Y = a + bZ + cZ^2$, we see that

$$\begin{aligned} \text{Cov}(Y, Z) &= a\text{Cov}(1, Z) + b\text{Cov}(Z, Z) + c\text{Cov}(Z^2, Z) \\ &= 0 + b + c\text{Cov}(Z^2, Z) . \end{aligned}$$

Now from Problem 54 in this Chapter we know that $\text{Cov}(Z^2, Z) = 0$, $\text{Var}(Z) = 1$, and we can compute $\text{Var}(Y)$ as

$$\begin{aligned} \text{Var}(Y) &= \text{Var}(a + bZ + cZ^2) \\ &= \text{Var}(bZ + cZ^2) \\ &= E[(bZ + cZ^2)^2] - E[(bZ + cZ^2)]^2 . \end{aligned}$$

Now $(bZ + cZ^2)^2 = b^2Z^2 + 2bcZ^3 + c^3Z^4$, so the expectation of this expression becomes $b^2 \cdot 1 + c^2E[Z^4]$. Now to compute $E[Z^4]$ when Z is a standard normal we can use the definition of expectation and evaluate

$$E[Z^4] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^4 e^{-\frac{1}{2}z^2} dz .$$

Introduce the variable $v = \frac{1}{2}z^2$, so that $dv = zdz$, and $z = \sqrt{2}\sqrt{v}$ so that our integral above becomes (using the evenness of the integrand and doubling the integral)

$$\begin{aligned} E[Z^4] &= \frac{2}{\sqrt{2\pi}}\int_0^\infty 4v^2 e^{-v}\frac{dv}{\sqrt{2}v^{1/2}} \\ &= \frac{4}{\sqrt{\pi}}\int_0^\infty v^{3/2}e^{-v}dv\,. \end{aligned}$$

Remembering the definition of the Gamma function $\Gamma(x) \equiv \int_0^\infty v^{x-1}e^{-x}dx$, we see that the above is equal to $\frac{4}{\sqrt{\pi}}\Gamma(\frac{5}{2})$ and from the identities $\Gamma(x+1) = x\Gamma(x)$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ we have that

$$\Gamma(\frac{5}{2}) = \frac{3}{2}\Gamma(\frac{3}{2}) = \frac{3}{2}\frac{1}{2}\Gamma(\frac{1}{2}) = \frac{3\sqrt{\pi}}{4}\,.$$

Thus our expectation becomes $E[Z^4] = 3$.

Problem 24

Following the hint we see that

$$0 < E[(tX+Y)^2] = E[t^2X^2 + 2tXY + Y^2] = t^2E[X^2] + 2tE[XY] + E[Y^2]\,,$$

so that the roots (in the variable t) of this equation must be imaginary and we must have that "$b^2 - 4ac < 0$" which using the expressions for this problem becomes

$$(2E[XY])^2 - 4E[X^2]E[Y^2] < 0\,,$$

or

$$E[XY]^2 \le E[X^2]E[Y^2]\,,$$

as claimed.

Problem 33 (expected number of flips to get r heads in a row)

We want to compute the expected number of flips until we get r heads in a row. Following the hint, we let X be a random variable representing the number of flips made until a string of r heads in a row obtained. Let Y be the number of flips until the first occurrence of a tail occurs. Then lets compute $E[X]$ by conditioning on the value of Y. We have

$$E[X] = \sum_{i=1}^{\infty} E[X|Y=i]P\{Y=i\}\,.$$

Note that we have

- If $Y = i \le r$ we have that $E[X|Y=i] = i + E[X]$ since if we observe a tail in the first r flips we must start the flipping process over.

- If $Y = i > r$ then $E[X|Y = i] = r$ since we obtained r heads in r flips.

In addition, we have $P\{Y = i\} = p^{i-1}(1-p)$ and from the above we have

$$E[X] = \sum_{i=1}^{r}(i + E[X])p^{i-1}(1-p) + \sum_{i=r+1}^{\infty} rp^{i-1}(1-p)\,.$$

Simplifying we get

$$\begin{aligned} E[X] &= (1-p)\sum_{i=1}^{r} ip^{i-1} + (1-p)E[X]\sum_{i=1}^{r} p^{i-1} + r(1-p)\sum_{i=r+1}^{\infty} p^{i-1} \\ &= \frac{1-p^r - r(1-p)p^r}{1-p} + (1-p)E[X]\left(\frac{1-p^r}{1-p}\right) + r(1-p)\left(\frac{p^r}{1-p}\right) \\ &= \frac{1-p^r}{1-p} + (1-p^r)E[X]\,. \end{aligned}$$

When we solve for $E[X]$ we find

$$E[X] = \frac{1-p^r}{p^r(1-p)}\,.$$

Problem 38 (minimizing $E[(Y - (a + bX_1 + cX_2))^2]$)

We are told that we want to pick a, b, and c to minimize the expression

$$f(a, b, c) \equiv E[(Y - (a + bX_1 + cX_2))^2]\,.$$

To find this minimum we take derivatives of the above with respect to the variables a, b, and c, set each result equal to zero, and then solve the resulting equations for a, b, and c. To do this we need to compute three derivatives

$$\begin{aligned} \frac{\partial}{\partial a}E[(Y - (a + bX_1 + cX_2))^2] &= E\left[2\left(Y - (a + bX_1 + cX_2)\right)(-1)\right] \\ &= -2E[Y] + 2a + 2bE[X_1] + 2cE[X_2] \\ \frac{\partial}{\partial b}E[(Y - (a + bX_1 + cX_2))^2] &= E\left[2\left(Y - (a + bX_1 + cX_2)\right)(-X_1)\right] \\ &= -2E[X_1Y] + 2aE[X_1] + 2bE[X_1^2] + 2cE[X_1X_2] \\ \frac{\partial}{\partial c}E[(Y - (a + bX_1 + cX_2))^2] &= E\left[2\left(Y - (a + bX_1 + cX_2)\right)(-X_2)\right] \\ &= -2E[X_2Y] + 2aE[X_2] + 2bE[X_1X_2] + 2cE[X_2^2]\,. \end{aligned}$$

Setting these three equations equal to zero we must find a, b, and c that satisfy

$$\begin{aligned} a + E[X_1]b + E[X_2]c &= E[Y] \\ E[X_1]a + E[X_1^2]b + E[X_1X_2]c &= E[X_1Y] \\ E[X_2]a + E[X_1X_2]b + E[X_2^2]c &= E[X_2Y]\,. \end{aligned}$$

Problem 39 (minimizing $E[(Y - (a + bX + cX^2))^2]$)

We are told that we want to pick a, b, and c to minimize the expression

$$f(a, b, c) \equiv E[(Y - (a + bX + cX^2))^2] \,.$$

To find this minimum we take derivatives of the above with respect to the variables a, b, and c, set each result equal to zero, and then solve the resulting equations for a, b, and c. To do this we need to compute three derivatives

$$\begin{aligned}
\frac{\partial}{\partial a} E[(Y - (a + bX + cX^2))^2] &= E\left[2\left(Y - (a + bX + cX^2)\right)(-1)\right] \\
&= -2E[Y] + 2a + 2bE[X] + 2cE[X^2] \\
\frac{\partial}{\partial b} E[(Y - (a + bX + cX^2))^2] &= E\left[2\left(Y - (a + bX + cX^2)\right)(-X)\right] \\
&= -2E[XY] + 2aE[X] + 2bE[X^2] + 2cE[X^3] \\
\frac{\partial}{\partial c} E[(Y - (a + bX + cX^2))^2] &= E\left[2\left(Y - (a + bX + cX^2)\right)(-X^2)\right] \\
&= -2E[X^2Y] + 2aE[X^2] + 2bE[X^3] + 2cE[X^4] \,.
\end{aligned}$$

Setting these three equations equal to zero we must find a, b, and c that satisfy

$$\begin{aligned}
a + E[X]b + E[X^2]c &= E[Y] \\
E[X]a + E[X^2]b + E[X^3]c &= E[XY] \\
E[X^2]a + E[X^3]b + E[X^4]c &= E[X^2Y] \,.
\end{aligned}$$

Problem 43 (a conditional expectation)

We want to evaluate

$$E[X^2 - 2XY + Y^2] = E[X^2] - 2E[XY] + E[Y^2] \,.$$

Using the law of iterated expectation in the form $E[X] = E_Z[E[X|Z]]$, since $Y \equiv E[X|Z]$ we find that the middle term above is given by

$$E[XY] = E_Z[E[XY|Z]] = E_Z[E[X(E[X|Z])|Z]] = E_Z[E[X|Z]E[X|Z]] = E_Z[Y^2] \,,$$

since $E[X|Z]$ is a function of Z only. Thus we have shown that

$$E[(X - Y)^2] = E[X^2] - 2E[Y^2] + E[Y^2] = E[X^2] - E[Y^2] \,,$$

as we were to show.

Problem 45 (the moment generating function for the uniform distribution)

The uniform distribution has a moment generating function that can be computed

$$M(t) = E[e^{tX}] = \int_a^b e^{tx}\frac{1}{b-a}dx = \frac{1}{b-a}\left(\frac{e^{tb}-e^{ta}}{t}\right).$$

We now compute $E(X)$ using the moment generating function $M(t)$ for a uniform random variable. Beginning this calculation we have

$$\begin{aligned}
E[X] &= \left.\frac{\partial M(t)}{\partial t}\right|_{t=0} \\
&= \frac{1}{b-a}\left.\left[\frac{1}{t}(be^{tb}-ae^{ta}) - \frac{1}{t^2}(e^{tb}-e^{ta})\right]\right|_{t=0} \\
&= \frac{1}{b-a}\left.\left[\frac{t(be^{tb}-ae^{ta})-(e^{tb}-e^{ta})}{t^2}\right]\right|_{t=0}.
\end{aligned}$$

To evaluate this expression requires the use of L'Hopital's rule where we find

$$\begin{aligned}
E[X] &= \frac{1}{b-a}\lim_{t\to 0}\left[\frac{(be^{tb}-ae^{at})+t(b^2e^{tb}-a^2e^{ta})-(be^{tb}-ae^{ta})}{2t}\right] \\
&= \frac{1}{b-a}\lim_{t\to 0}\left[\frac{b^2e^{tb}-a^2e^{ta}}{2}\right] = \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right) = \frac{a+b}{2}.
\end{aligned}$$

We now compute $E(X^2)$ using the moment generating function $M(t)$ for a uniform random variable.

$$\begin{aligned}
E[X^2] &= \left.\frac{\partial^2 M(t)}{\partial t^2}\right|_{t=0} \\
&= \frac{1}{b-a}\left.\left[-\frac{1}{t^2}(be^{tb}-ae^{ta}) + \frac{1}{t}(b^2e^{tb}-a^2e^{ta}) + \frac{2}{t^3}(e^{tb}-e^{ta}) - \frac{1}{t^2}(be^{tb}-ae^{ta})\right]\right|_{t=0} \\
&= \frac{1}{b-a}\left.\left[\frac{-tbe^{tb}+tae^{ta}+b^2t^2e^{tb}-a^2t^2e^{ta}+2e^{tb}-2e^{ta}-tbe^{tb}+tae^{ta}}{t^3}\right]\right|_{t=0} \\
&= \frac{1}{b-a}\left.\left[\frac{-2tbe^{tb}+2tae^{ta}+b^2t^2e^{tb}-a^2t^2e^{ta}+2e^{tb}-2e^{ta}}{t^3}\right]\right|_{t=0} \\
&= \frac{1}{b-a}\left.\left[\frac{(-2tb+b^2t^2+2)e^{tb}+(2ta-a^2t^2-2)e^{ta}}{t^3}\right]\right|_{t=0}.
\end{aligned}$$

Evaluating this again requires L'Hopital's rule where we find

$$
\begin{aligned}
E[X^2] &= \frac{1}{b-a}\lim_{t\to 0}\left[\frac{(-2b+2b^2t-2tb^2+b^3t^2+2b)e^{tb}+(2a-2a^2t+2ta^2-a^3t^2-2a)e^{ta}}{3t^2}\right] \\
&= \frac{1}{b-a}\lim_{t\to 0}\left[\frac{b^3t^2e^{tb}-a^3t^2e^{ta}}{3t^2}\right] \\
&= \frac{1}{b-a}\lim_{t\to 0}\left[\frac{(2b^3t+b^4t^2)e^{tb}-(2a^3t+a^4t^2)e^{ta}}{6t}\right] \\
&= \frac{1}{b-a}\lim_{t\to 0}\left[\frac{(2b^3+2b^4t+2b^4t+b^5t^2)e^{tb}-(2a^3+2a^4t+2a^4t+a^5t^2)e^{ta}}{6}\right] \\
&= \frac{1}{b-a}\lim_{t\to 0}\left[\frac{(2b^3+4b^4t+b^5t^2)e^{tb}-(2a^3+4a^4t+a^5t^2)e^{ta}}{6}\right] \\
&= \frac{1}{b-a}\left[\frac{2b^3-2a^3}{6}\right] = \frac{b^2+ab+a^2}{3}\,.
\end{aligned}
$$

Thus given this expression for $E[X^2]$ we can compute $\text{Var}(X)$ using $E[X]$ as

$$
\begin{aligned}
\text{Var}(X) = E[X^2]-E[X]^2 &= \frac{b^2+ab+a^2}{3}-\left(\frac{a+b}{2}\right)^2 \\
&= \frac{4(b^2+ab+a^2)-3(a^2+2ab+b^2)}{12} \\
&= \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}\,.
\end{aligned}
$$

Problem 46 (moments of the standard normal)

From the book we have $M_Z(t) = E[e^{tZ}] = e^{\frac{t^2}{2}}$. Expanding the right-hand-side of this expression in a Taylor series about 0 we have

$$
M_Z(t) = \sum_{j=0}^{\infty}\frac{t^{2j}}{2^j j!}\,. \qquad (33)
$$

The general theory of Taylor series states that this must equal the expression

$$
M_Z(t) = \sum_{j=0}^{\infty}\frac{M_Z^{(j)}(0)}{j!}t^j = \sum_{j=0}^{\infty}\frac{\mu_j}{j!}t^j\,.
$$

By equating coefficients of t^j between these two expression for $M_Z(t)$ we first see that $\mu_j = 0$ when j is odd. Using this fact, the general expression for $M_Z(t)$ then becomes

$$
M_Z(t) = \sum_{j=0}^{\infty}\frac{\mu_{2j}}{(2j)!}t^{2j}\,.
$$

Equating powers of t in this expression to those in Equation 33 we have that

$$
\frac{1}{(2j)!}\mu_{2j} = \frac{1}{2^j j!}\,,
$$

or

$$\mu_{2j} = \frac{(2j)!}{2^j j!} \quad \text{for} \quad j \geq 0\,,$$

as we were to show.

Problem 47 (moments of a normal RV with mean μ and variance σ^2)

From the book we have

$$\begin{aligned}
M_X(t) &= e^{t\mu} M_Z(t\sigma) = e^{t\mu} e^{\frac{(t\sigma)^2}{2}} \\
&= \left(\sum_{l=0}^{\infty} \frac{(t\mu)^l}{l!} \right) \left(\sum_{j=0}^{\infty} \frac{(t\sigma)^{2j}}{2^j j!} \right) = \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \frac{\mu^l \sigma^{2j}}{l! j! 2^j} t^{l+2j} \\
&= \sum_{n=0}^{\infty} \left(\sum_{j=0}^{[n/2]} \frac{\mu^{n-2j}\sigma^{2j}}{(n-2j)! j! 2^j} \right) t^n\,.
\end{aligned}$$

Thus from the coefficient of t^n we have

$$\frac{E[X^n]}{n!} = \sum_{j=0}^{[n/2]} \frac{\mu^{n-2j}\sigma^{2j}}{(n-2j)! j! 2^j}\,,$$

so

$$E[X^n] = \sum_{j=0}^{[n/2]} \frac{n!}{(n-2j)!(2j)!} \left(\frac{(2j)!}{2^j j!} \right) \mu^{n-2j}\sigma^{2j} = \sum_{j=0}^{[n/2]} \left(\begin{array}{c} n \\ 2j \end{array} \right) \frac{\mu^{n-2j}\sigma^{2j}(2j)!}{2^j j!}\,, \qquad (34)$$

as we were to show. Consider the case $n = 1$. Then Equation 34 has only one term $k = 0$ and we get

$$E[X] = \left(\begin{array}{c} 1 \\ 0 \end{array} \right) \frac{\mu^1 \sigma^0 0!}{2^0 0!} = \mu\,.$$

When $n = 2$ Equation 34 has two terms $k = 0$ and $k = 1$ and we get

$$E[X^2] = \left(\begin{array}{c} 2 \\ 0 \end{array} \right) \mu^2 + \left(\begin{array}{c} 2 \\ 2 \end{array} \right) \frac{\mu^{2-2}\sigma^2 2!}{21!} = \mu^2 + \sigma^2\,.$$

Another way of doing this problem is to note that X can be written as $X = \mu + \sigma Z$ where Z is a standard normal random variable. Then using the binomial theorem we can write

$$X^n = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \sigma^k Z^k \mu^{n-k}\,.$$

Taking the expectation of this and using the formula for the moments of the standard normal we get

$$E[X^n] = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \sigma^k \mu^{n-k} E[Z^k] = \sum_{j=0}^{[n/2]} \left(\begin{array}{c} n \\ 2j \end{array} \right) \sigma^{2j} \mu^{n-2j} \frac{(2j)!}{2^j j!}\,,$$

the same result.

Problem 48 (the moment generating function for $Y = aX + b$)

We have

$$M_Y(t) = E[e^{tY}] = E[e^{t(aX+b)}] \\ = e^{tb}E[e^{taX}] = e^{tb}M_X(ta)\,.$$

Problem 49 (the mean and variance of a lognormal random variable)

Let X be the lognormal random variable so that if we let Y be its logarithm as $Y = \log(X)$ then since Y is a normal random variable we have that $M_Y(t) = E[e^{tY}] = \exp\{\mu t + \frac{\sigma^2}{2}t^2\}$. If we evaluate this expression at $t = 1$ we get

$$E[e^Y] = \exp\{\mu + \frac{\sigma^2}{2}\} = E[X]\,,$$

since $e^Y = X$. In addition we have that

$$E[X^2] = E[e^{2Y}] = M_Y(2) = \exp\{2\mu + 2\sigma^2\}\,.$$

From these two moments we can compute the variance

$$\text{Var}(X) = E[X^2] - E[X]^2 = \exp\{2\mu + 2\sigma^2\} - \exp\{2\mu + \sigma^2\} = \exp\{2\mu + \sigma^2\}(\exp\{\sigma^2\} - 1)\,.$$

Problem 50 (logs of moment generating functions)

When we have $\psi(t) = \log(M(t))$ we have that

$$\psi'(t) = \frac{M'(t)}{M(t)}$$
$$\psi''(t) = \frac{M''(t)}{M(t)} - \frac{M'(t)^2}{M(t)^2}\,.$$

To evaluate these at $t = 0$ we recall that $M(0) = 1$, $M'(0) = E[X]$, and $M''(0) = E[X^2]$ and find

$$\psi'(0) = E[X^2] - E[X]^2 = \text{Var}(X)\,,$$

as we were to show.

Problem 51 (the distribution of $\sum X_i$ when X_i is an exponential)

Since the sum of independent random variables has a moment generating function that is the product of the moment generating function of the summands. In the case given here we then have

$$M_{\sum_i X_i}(t) = \prod_{i=1}^{N} M_{X_i}(t) = \prod_{i=1}^{N}\left(\frac{\lambda}{\lambda - t}\right) = \left(\frac{\lambda}{\lambda - t}\right)^N\,,$$

which is the moment generating function for a gamma random variable.

Problem 52 (computing $\mathrm{Cov}(X, Y)$ from the moment generating function)

When we have two random variables X and Y the moment generating function looks like $M(t_1, t_2) = E[e^{t_1 X + t_2 Y}]$ from which we see

$$\frac{\partial^2 M}{\partial t_1 \partial t_2}(t_1, t_2) = E[XY e^{t_1 X + t_2 Y}] \,.$$

Thus

$$\frac{\partial^2 M}{\partial t_1 \partial t_2}(0, 0) = E[XY] \,.$$

In the same way we have $\frac{\partial M}{\partial t_1}(0, 0) = E[X]$ and $\frac{\partial M}{\partial t_2}(0, 0) = E[Y]$. Thus we can write $\mathrm{Cov}(X, Y)$ in terms of the moment generating function as

$$\mathrm{Cov}(X, Y) = \frac{\partial^2 M}{\partial t_1 \partial t_2}(0, 0) - \frac{\partial M}{\partial t_1}(0, 0) \frac{\partial M}{\partial t_2}(0, 0) \,.$$

Problem 53 (multivariate random variables)

We will use the fact that $X_1, X_2, \cdots, X_n$ are independent if and only if the multidimensional moment generating function factors into the product of individual moment generating functions as

$$M(t_1, t_2, \cdots, t_n) = M_{X_1}(t_1) \cdots M_{X_n}(t_n) \,.$$

Since the moment generating function for a set of m multivariate normal random variables looks like

$$M(t_1, t_2, \cdots, t_n) = \exp \left\{ \sum_{i=1}^{m} t_i \mu_i + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} t_i t_j \mathrm{Cov}(X_i, X_j) \right\} \,,$$

in order for this expression to factor into the product of moment generating functions we must have $\mathrm{Cov}(X_i, X_j) = 0$ when $i \neq j$. In this case, using $\mathrm{Cov}(X_i, X_i) = \mathrm{Var}(X_i)$, we have that

$$M(t_1, t_2, \cdots, t_n) = \exp \left\{ \sum_{i=1}^{m} t_i \mu_i + \frac{1}{2} \sum_{i=1}^{m} t_i^2 \mathrm{Var}(X_i) \right\} = \prod_{i=1}^{m} \exp \left\{ t_i \mu_i + \frac{1}{2} t_i^2 \mathrm{Var}(X_i) \right\} \,,$$

which is the product of moment generating functions for normal random variables.

Problem 54 ($\mathrm{Cov}(Z, Z^2)$ when Z is a standard normal)

We have that $\mathrm{Cov}(Z, Z^2) = E[Z^3] - E[Z]E[Z^2]$. Since Z is a standard normal random variable we know that $E[Z] = 0$ and $E[Z^3] = 0$. Both of these can be seen from the identity of integrating an odd function over an symmetric integral. Thus $\mathrm{Cov}(Z, Z^2) = 0$.

Problem 55 (joint normal RVs)

Part (a): From the description the conditional density of $X|Y$ is normal with a mean y and a variance 1, while the conditional distribution of $Y+Z|Y$ is exactly the same. Since the conditional distributions are the same and the distribution of Y is the same in both cases we can conclude that the two distributions are equal. In symbols

$$p(X,Y) = p(X|Y)p(Y) = p(Y+Z|Y)p(Y) = P(Y+Z,Y)\,.$$

Part (b): Since Y and Z are independent normal random variables $Y+Z$ being the sum of independent normal random variables is also normal so the density of $(Y+Z,Y)$ has a bivariate normal density.

Part (c): We find

$$\begin{aligned}
E[X] &= E[Y+Z] = \mu + 0 = \mu \\
\text{Var}(X) &= \text{Var}(Y+Z) = \sigma^2 + 1 \\
\rho = \text{Corr}(X,Y) &= \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\text{Cov}(Y+Z,Y)}{\sqrt{(\sigma^2+1)\sigma^2}} \\
&= \frac{\text{Cov}(Y,Y)+\text{Cov}(Z,Y)}{\sqrt{(\sigma^2+1)\sigma^2}} = \frac{\sigma^2}{\sqrt{(\sigma^2+1)\sigma^2}} = \frac{\sigma}{\sqrt{\sigma^2+1}}\,.
\end{aligned}$$

Part (d): We have

$$\begin{aligned}
E[Y|X=x] &= E[Y] + \rho\sqrt{\frac{\text{Var}(Y)}{\text{Var}(X)}}(x - E[X]) \\
&= \mu + \frac{\sigma}{\sqrt{\sigma^2+1}}\sqrt{\frac{\sigma^2}{\sigma^2+1}}(x-\mu) \\
&= \mu + \frac{\sigma^2}{\sigma^2+1}(x-\mu)\,.
\end{aligned}$$

Part (e): As X, Y has a bivariate normal distribution $p(Y|X=x)$ has a normal distribution with mean given by the expression in Part (d) and a variance given by

$$\text{Var}(Y|X=x) = \sigma^2(1-\rho^2) = \sigma^2\left(1 - \frac{\sigma^2}{\sigma^2+1}\right) = \frac{\sigma^2}{\sigma^2+1}\,.$$

Chapter 8 (Limit Theorems)

Chapter 8: Problems

Problem 1 (bounding the probability we are between two numbers)

We are told that $\mu = 20$ and $\sigma^2 = 20$ so that

$$P\{0 < X < 40\} = P\{-20 < X - 20 < 20\} = 1 - P\{|X - 20| > 20\}.$$

Now by Chebyshev's inequality

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2},$$

we know that

$$P\{|X - 20| > 20\} \leq \frac{20}{20^2} = 0.05.$$

This implys that (negating both sides that)

$$-P\{|X - 20| > 20\} > -0.05,$$

so that $1 - P\{|X - 20| > 20\} > 0.95$. In summary then we have that $P\{0 < X < 40\} > 0.95$.

Problem 2 (distribution of test scores)

We are told, that if X is the students score in taking this test then $E[X] = 75$.

Part (a): Then by Markov's inequality we have

$$P\{X \geq 85\} \leq \frac{E[X]}{85} = \frac{75}{85} = \frac{15}{17}.$$

If we also know the variance of X is given by $\text{Var}X = 25$, then we can use the one-sided Markov inequality given by

$$P\{X - \mu \geq a\} \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

With $\mu = 75$, $a = 10$, $\sigma^2 = 25$ this becomes

$$P\{X \geq 85\} \leq \frac{25}{25 + 10^2} = \frac{1}{5}.$$

Part (b): Using Chernoff's inequality given by

$$P\{|X - \mu| \geq k\sigma\} \leq \frac{1}{k^2},$$

we have (since we want $5k = 10$ or $k = 2$) that

$$P\{|X - 75| \geq 2 \times 5\} \leq \frac{1}{2^2} = 0.25\,,$$

Thus

$$P\{|X - 75| \leq 10\} = 1 - P\{|X - 75| \geq 10\} = 1 - \frac{1}{4} = \frac{3}{4}\,.$$

Part (c): We desire to compute

$$P\{75 - 5 \leq \frac{1}{n}\sum_{i=1}^{n} x_i \leq 75 + 5\} = P\{|\frac{1}{n}\sum_{i=1}^{n} x_i - 75| \leq 5\}$$

Defining $X = \sum_{i=1}^{n} X_i$, we have that $\mu = E[X] = 75$ and $\text{Var}(X) = \frac{1}{n^2} \times n\text{Var}(X) = \frac{25}{n}$. So to use Chernoff' inequality on this problem we desire a k such that $k\left(\frac{5}{\sqrt{n}}\right) = 5$ so $k = \sqrt{n}$ and then Chernoff's bound gives

$$P\{|\frac{1}{n}\sum_{i=1}^{n} x_i - 75| > 5\} \leq \frac{1}{n}\,.$$

So to make $P\{|\frac{1}{n}\sum_{i=1}^{n} x_i - 75| > 5\} \leq 0.1$ we must take

$$\frac{1}{n} \leq 0.1 \Rightarrow n \geq 10\,.$$

Problem 3 (an example with the central limit theorem)

We want to compute n such that

$$P\{\left|\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - 75}{5/\sqrt{n}}\right| \leq \frac{5}{5/\sqrt{n}}\} \geq 0.9\,.$$

Now by the central limit theorem the expression

$$\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - 75}{5/\sqrt{n}}\,,$$

we have that the above can be written (first removing the absolute values)

$$\begin{aligned} P\{\left|\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - 75}{5/\sqrt{n}}\right| \leq \sqrt{n}\} &= 1 - 2P\{\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - 75}{5/\sqrt{n}} \leq \sqrt{n}\} \\ &= 1 - 2\Phi(-\sqrt{n})\,. \end{aligned}$$

Setting this equal to 0.9 gives $\Phi(-\sqrt{n}) = 0.05$, or when we solve for n we obtain

$$n > (-\Phi^{-1}(0.05))^2 = 2.7055\,.$$

In the file `chap_8_prob_3.m` we use the Matlab command `norminv` to compute this value. We see that we should take $n \geq 3$.

Problem 4 (sums of Poisson random variables)

Part (a): The Markov inequality is $P\{X \geq a\} \leq \frac{E[X]}{a}$, so if $X = \sum_{i=1}^{20} X_i$ then $E[X] = \sum_{i=1}^{20} E[X_i] = 20$, and the Markov inequality becomes in this case

$$P\{X \geq 15\} \leq \frac{20}{15} = \frac{4}{3}.$$

Note that since all probabilities must be less than one, this bound is not informative.

Part (b): We desire to compute (using the central limit theorem) $P\{\sum_{i=1}^{20} X_i > 15\}$. Thus the desired probability is given by (since $\sigma = \sqrt{\text{Var}(X_i)} = 1$)

$$\begin{aligned} P\{\frac{\sum_{i=1}^{20} X_i - 20}{\sqrt{20}} > \frac{15-20}{\sqrt{20}}\} &= 1 - P\{Z < -\frac{5}{\sqrt{20}}\} \\ &= 0.8682. \end{aligned}$$

This calculation can be found in `chap_8_prob_4.m`.

Problem 5 (rounding to integers)

Let $R = \sum_{i=1}^{50} R_i$ be the approximate sum where each R_i is the rounded variable and let $X = \sum_{i=1}^{50} X_i$ be the exact sum. We desire to compute $P\{|X - R| > 3\}$, which can be simplified to give

$$\begin{aligned} P\{|X-R| > 3\} &= P\left\{\left|\sum_{i=1}^{50} X_i - \sum_{i=1}^{50} R_i\right| > 3\right\} \\ &= P\left\{\left|\sum_{i=1}^{50}(X_i - R_i)\right| > 3\right\}. \end{aligned}$$

Now $X_i - R_i$ are independent uniform random variables between $[-0.5, 0.5]$ so the above can be evaluated using the central limit theorem. For this sum of random variables the mean of the individual random variables $X_i - R_i$ is zero while the standard deviation σ is given by

$$\sigma^2 = \frac{(0.5 - (-0.5))^2}{12} = \frac{1}{12}.$$

Thus by the central limit theorem we have that

$$\begin{aligned} P\left\{\left|\sum_{i=1}^{50}(X_i - R_i)\right| > 3\right\} &= P\left\{\left|\frac{\sum_{i=1}^{50}(X_i - R_i)}{50/\sqrt{12}}\right| > \frac{3}{50/\sqrt{12}}\right\} \\ &= 2P\left\{\frac{\sum_{i=1}^{50}(X_i - R_i)}{50/\sqrt{12}} < \frac{-3}{50/\sqrt{12}}\right\} \\ &= 2\Phi(\frac{-3}{50/\sqrt{12}}) = 0.8353. \end{aligned}$$

This calculation can be found in `chap_8_prob_5.m`.

Problem 6 (rolling a die until our sum exceeds 800)

The sum of n die rolls is given by $X = \sum_{i=1}^{n} X_i$ with X_i a random variable taking values $1, 2, 3, 4, 5, 6$ all with probability of $1/6$. Then

$$\mu = E[\sum_{i=1}^{n} E[X_i] = nE[X_i] = \frac{n}{6}(1+2+3+4+5+6) = \frac{7}{2}n$$

In addition, because of the independence of our X_i we have that $\text{Var}(X) = n\text{Var}(X_i)$. For the individual random variables X_i we have that $\text{Var}(X_i) = E[X_i^2] - E[X_i]^2$. For die we have

$$E[X_i^2] = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}.$$

so that our variance is given by

$$\text{Var}(X_i) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = 2.916\,.$$

Now the probability we want to calculate is given by $P\{X > 300\}$, which we can maniuplate into a form where we can apply the central limit theorm. We have

$$P\left\{\frac{X - \frac{7n}{2}}{\sqrt{2.916}\sqrt{n}} > \frac{300 - \frac{7n}{2}}{\sqrt{2.916}\sqrt{n}}\right\}$$

Now if $n = 80$ we have the above given by

$$P\left\{\frac{X - \frac{7}{2}\cdot 80}{\sqrt{2.916}\sqrt{80}} > \frac{300 - \frac{7}{2}\cdot 80}{\sqrt{2.916}\sqrt{80}}\right\} = 1 - P\{Z < 1.309\} = 1 - \Phi(1.309) = 0.0953\,.$$

Problem 7 (working bulbs)

The total lifetime of all the bulbs is given by

$$T = \sum_{i=1}^{100} X_i\,,$$

where X_i is an exponential random variable with mean five hours. Then since the random variable T is the sum of independent identically distributed random variables we can use the central limit theorm to derive estimates about T. For example we know that

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}\,,$$

is approximatly a standard normal. Thus to evaluate (since $\sigma^2 = 25$) we have that

$$\begin{aligned}
P\{T > 525\} &= P\left\{\frac{T - 100(5)}{10(5)} > \frac{525 - 500}{50}\right\} \\
&= 1 - P\{Z < 1/2\} \\
&= 1 - \Phi(0.5) = 1 - 0.6915 = 0.3085\,.
\end{aligned}$$

Problem 8 (working bulbs with replacement times)

Our expression for the total time that there is a working bulb in problem 7 without any replacment time is given by

$$T = \sum_{i=1}^{100} X_i \,.$$

If there is a random time required to replace each bulb then we our random variable T must now include this randomness and becomes

$$T = \sum_{i=1}^{100} X_i + \sum_{i=1}^{99} U_i \,.$$

Again we desire to evaluate $P\{T \leq 550\}$. To evaluate this let

$$T = \sum_{i=1}^{99} (X_i + U_i) + X_{100} \,,$$

which motivates us to define the random variables V_i as

$$V_i = \begin{cases} X_i + U_i & i = 1, \cdots, 99 \\ X_{100} & i = 100 \end{cases}$$

Then $T = \sum_{i=1}^{100} V_i$ and the V_i's are all independent. Below we will introduce the variables μ_i and σ_i to be the mean and the standard deviation respectivly of the random variable V_i. Taking the expectation of T we find

$$\begin{aligned} E[T] &= \sum_{i=1}^{100} E[V_i] = \sum_{i=1}^{99} (E[X_i] + E[U_i]) + E[X_{100}] \\ &= 100 \cdot 5 + 99 \left(\frac{1}{4} \right) = 524.75 \,. \end{aligned}$$

In the same way the variance of this summation is also given by

$$\begin{aligned} \text{Var}(T) &= \sum_{i=1}^{99} (\text{Var}(X_i) + \text{Var}(U_i)) + \text{Var}(X_{100}) \\ &= 100 \cdot 5 + 99 \cdot \frac{1}{4} \left(\frac{1}{12} \right) = 502.0625 \,. \end{aligned}$$

By the central limit theorm we have that

$$P\left\{ \sum_{i=1}^{100} V_i \leq 550 \right\} = P\left\{ \frac{\sum_{i=1}^{100} (V_i - \mu_i)}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \leq \frac{550 - \sum_{i=1}^{100} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \right\} \,.$$

Where the variables μ_i and σ_i the means and standard deviations of the variables V_i. Calculating the expression on the right handside of the inequality above i.e.

$$\frac{550 - \sum_{i=1}^{100} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \,,$$

we find it equal to $\frac{550-524.75}{\sqrt{502.0625}} = 1.1269$. Therefore we see that

$$P\left\{\sum_{i=1}^{100} V_i \leq 550\right\} \approx \Phi(1.1269) = 0.8701\,,$$

using the Matlab function `normcdf`.

Problem 9 (how large n needs to be)

Warning: This result does not match the back of the book. If anyone can find anything incorrect with this problem please let me know.

A gamma random variable with parameters $(n, 1)$ is equivalent to a sum of n exponential random variables each with parameter $\lambda = 1$. i.e. $X = \sum_{i=1}^n X_i$, with each X_i an exponential random variable with $\lambda = 1$. This result is discussed in Example 3b Page 282 Chapter 6 in the book. Then the requested problem seems equivalent to computing n such that

$$P\left\{\left|\frac{\sum_{i=1}^n X_i}{n} - 1\right| > 0.01\right\} < 0.01\,.$$

which we will do by converting this into an expression that looks like the central limit theorem and then evaluate. Recognizing that X is a sum of exponential with parameters $\lambda = 1$, we have that

$$\mu = E[X] = E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n \frac{1}{\lambda} = n\,.$$

In the same way since $\text{Var}(X_i) = \frac{1}{\lambda^2} = 1$, we have that

$$\sigma^2 = \text{Var}(X) = \sum_{i=1}^n \text{Var}(X_i) = n\,.$$

Then the central limit theorem applied to the random variable X claims that as $n \to \infty$, we have

$$P\left\{\left|\frac{\sum_{i=1}^n X_i - n}{\sqrt{n}}\right| < a\right\} = \Phi(a) - \Phi(-a)\,.$$

or taking the requested probabilistic statement and converting it we find that

$$\begin{aligned}
P\left\{\left|\frac{\sum_{i=1}^n X_i}{n} - 1\right| > 0.01\right\} &= 1 - P\left\{\left|\frac{\sum_{i=1}^n X_i}{n} - 1\right| \leq 0.01\right\} \\
&= 1 - P\left\{\left|\frac{\sum_{i=1}^n X_i - n}{n}\right| \leq 0.01\right\} \\
&= 1 - P\left\{\left|\frac{\sum_{i=1}^n X_i - n}{\sqrt{n}}\right| \leq 0.01\sqrt{n}\right\} \\
&\approx 1 - (\Phi(0.01\sqrt{n}) - \Phi(-0.01\sqrt{n}))\,.
\end{aligned}$$

From the following identity on the cumulative distribution of a normal random variable we have that $\Phi(x) - \Phi(-x) = 1 - 2\Phi(-x)$, so that the above equals

$$1 - (1 - 2\Phi(-0.01\sqrt{n})) = 2\Phi(-0.01\sqrt{n}) \,.$$

To have this be less that 0.01 requires a value of n such that

$$2\Phi(-0.01\sqrt{n}) \le 0.01 \,.$$

Solving for n then gives $n \ge (-100\Phi^{-1}(0.005))^2 = (257.58)^2$.

Problem 11 (a simple stock model)

Given the recurrence relationship $Y_n = Y_{n-1} + X_n$ for $n \ge 1$, with $Y_0 = 100$, we see that a solution to this is given by

$$Y_n = \sum_{k=1}^{n} X_k + Y_0 \,.$$

If we assume that the X_k's are independent identically distributed random variables with mean 0 and variance σ^2, we are asked to evaluate

$$P\{Y_{10} > 105\} \,.$$

Which we will do by transforming this problem into something that looks like an application of the central limit theorem. We find that

$$\begin{aligned} P\{Y_{10} > 105\} &= P\{\sum_{k=1}^{10} X_k > 5\} \\ &= P\left\{ \frac{\sum_{k=1}^{10} X_k - 10\cdot(0)}{\sqrt{10}} > \frac{5 - 10\cdot(0)}{\sqrt{10}} \right\} \\ &= 1 - P\left\{ \frac{\sum_{k=1}^{10} X_k - 10\cdot(0)}{\sqrt{10}} < \frac{5}{\sqrt{10}} \right\} \\ &\approx 1 - \Phi(\frac{5}{\sqrt{10}}) = 0.0569 \,. \end{aligned}$$

Problem 12

The total life of our 100 components is given by $L = \sum_{i=1}^{100} X_i$ with X_i exponentially distributed with rate $\lambda_i = \frac{1}{10+\frac{i}{10}} = \frac{10}{100+i}$. We want to estimate the following probability

$$P\{L > 1200\} = P\left\{ \sum_{i=1}^{100} X_i > 1200 \right\} \,.$$

From the properties of exponential random variables the mean of each X_i is given by $\mu_i = \frac{1}{\lambda_i} = 10 + \frac{i}{10}$ and the variance is $\text{Var}(X_i) = \frac{1}{\lambda_i^2} = \left(10 + \frac{i}{10}\right)^2$. Then to compute the above the probability with respect to L we transforme the right handside in the usual manner. The central limit theorem for independent random variables gives

$$P\left\{\frac{\sum_{i=1}^{100}(X_i - \mu_i)}{\sqrt{\sum_{i=1}^{100}\sigma_i^2}} \le a\right\} \to \Phi(a) \quad \text{as} \quad n \to \infty\,.$$

So the above probability can be calculated as

$$\begin{aligned} P\{L > 1200\} &= 1 - P\left\{\frac{\sum_{i=1}^{100}(X_i - \mu_i)}{\sqrt{\sum_{i=1}^{100}\sigma_i^2}} \le \frac{1200 - \sum_{i=1}^{100}\mu_i}{\sqrt{\sum_{i=1}^{100}\sigma_i^2}}\right\} \\ &\approx 1 - \Phi\left(\frac{1200 - \sum_{i=1}^{100}\mu_i}{\sqrt{\sum_{i=1}^{100}\sigma_i^2}}\right) \end{aligned}$$

If we change the distribution of X_i such that X_i is uniform over $(0, 20 + \frac{i}{5})$ we then from properties of uniform random variable we know that

$$\begin{aligned} \mu_i &= 10 + \frac{i}{10} \\ \sigma_i^2 &= \frac{\left(20 + \frac{i}{5} - 0\right)^2}{12} = \frac{1}{3}\left(10 + \frac{i}{10}\right)^2\,. \end{aligned}$$

which is different from the previous variance calculation by the reduction of each individual variance by three. Thus the $\sum_{i=1}^{100}\sigma_i^2$ is also reduced by $\frac{1}{3}$ from the earlier result. This propagates through the calculation and we see that

$$P\{L > 1200\} = 1 - \Phi\left(\frac{1200 - \sum_{i=1}^{100}\mu_i}{\sqrt{3}\sqrt{\sum_{i=1}^{100}\sigma_i^2}}\right)\,,$$

with μ_i and σ_i in the above evaluated for the exponential variable.

Problem 13

Warning: Here are some notes I had on this problem. I've not had the time to check these in as much detail as I would have liked. Caveat emptor.

Part (a): Let X_i be the score of the ith student. Then since X_i is drawn from a distribution with mean 74 and standard deviation of 14. Then the average test scores for this class of 25 is given by

$$A = \frac{1}{25}\sum_{i=1}^{25} X_i\,.$$

Then the probability that A exceeds 80 is $P\{A \geq 80\} = 1 - P\{A \leq 80\}$. From the central limit theorem we see that the probability $P\{A \leq 80\}$ can be expressed in terms of the standard normal. Specifically

$$\begin{aligned} P\{A \leq 80\} &= P\{\frac{1}{25}\sum_{i=1}^{25} X_i \leq 80\} = P\{\sum_{i=1}^{25} X_i \leq 25(80)\} \\ &= P\left\{\frac{\sum_{i=1}^{25} X_i - 25(74)}{14\sqrt{25}} \leq \frac{25(80) - 25(74)}{14\sqrt{25}}\right\} = \Phi\left(\frac{25(6)}{14(5)}\right) = \Phi\left(\frac{15}{7}\right) \end{aligned}$$

Part (c): We have $\mu = 74$ and $\sigma = 14$. $S_{25} = \frac{1}{25}\sum_{i=1}^{25} X_i$ and $S_{64} = \frac{1}{64}\sum_{i=1}^{64} Y_i$. From the central limit theorem we know that

$$\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} \sim \mathcal{N}(0, 1)\,,$$

so that

$$\frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})\,.$$

Thus $S_{25} \sim \mathcal{N}(74, \frac{14^2}{25})$ and $S_{64} \sim \mathcal{N}(74, \frac{14^2}{64})$. Now

$$V = S_{64} - S_{25} \sim \mathcal{N}(0, \frac{14^2}{25} + \frac{14^2}{64})\,,$$

so that

$$\begin{aligned} P\{V \geq 2.2\} &= 1 - P\{V \leq 2.2\} \\ &= 1 - P\left\{\frac{V}{\sqrt{\frac{14^2}{25} + \frac{14^2}{64}}} \leq \frac{2.2}{\sqrt{\frac{14^2}{25} + \frac{14^2}{64}}}\right\} \\ &= 1 - \Phi\left(\frac{2.2}{\sqrt{\frac{14^2}{25} + \frac{14^2}{64}}}\right)\,. \end{aligned}$$

Problem 14

Let X_i be the random variable denoting the mean lifetime of the ith component. We are told that $E[X_i] = 100$ and $\text{Var}(X_i) = 30^2$. We assume that we have n components of this type in stock and can assume that each is replaced immediately when the previous one breaks. We then desire to compute the value of n such that

$$P\left\{\sum_{i=1}^{n} X_i > 2000\right\} > 0.95\,,$$

or equivalently

$$P\left\{\sum_{i=1}^{n} X_i > 2000\right\} < 1 - 0.95 = 0.05\,.$$

Now this can be done by using the central limit theorem for independent random variables. We have that

$$P\left\{\frac{\sum_{i=1}^{n}(X_i - 100)}{30n^{1/2}} < \frac{2000 - 100n}{30n^{1/2}}\right\} \to \Phi\left(\frac{2000 - 100n}{30\sqrt{n}}\right)\,.$$

Thus we should select n such that

$$\Phi\left(\frac{2000 - 100n}{30\sqrt{n}}\right) \approx 0.05\,,$$

which is a nonlinear function that needs to be solved to find the smallest value of n.

Problem 15

Let C be the random variable that denotes the total yearly claim for our 10,000 policy holders. Then $C = \sum_{i=1}^{10000} X_i$ with X_i the random claim made by the ith policy holder. We desire to evaluate $P\{C > 2.7\,10^6\}$ or

$$P\left\{\sum_{i=1}^{10000} X_i > 2.7\,10^6\right\} \quad \text{or} \quad 1 - P\left\{\sum_{i=1}^{10000} X_i < 2.7\,10^6\right\}\,.$$

Using the central limit theorem for independent random variables we have that

$$P\left\{\frac{\sum_{i=1}^{10000} X_i - \sum_{i=1}^{10000} 240}{\sqrt{\sum_{i=1}^{10000} 800^2}} < \frac{2.7\,10^6 - 240\,10^4}{800\sqrt{10^4}}\right\} = \Phi\left(\frac{2.7\,10^6 - 240\,10^4}{800\sqrt{10^4}}\right)\,,$$

which can easily be evaluated.

Problem 16

If we assume that the number N of men-women pairs is approximately normally distributed then we desire to calculate $P\{N > 30\}$. The mean of men-women pairs I would expect to be $\frac{1}{2}(100) = 50$, with a variance of "npq" or $\frac{1}{2}\left(\frac{1}{2}\right)(100) = 25$. Thus we can evaluate the above using

$$\begin{aligned} P\{N > 30\} &= 1 - P\{N < 30\} \\ &= 1 - P\left\{\frac{N - 50}{5} < \frac{-20}{5}\right\} \\ &= 1 - \Phi(-4) \approx 1\,. \end{aligned}$$

I would expect this to not be a good approximation.

Problem 18

Let Y denoted the random variable representing the number of fish that must be caught to obtain at least one of these types.

Part (a): The probability to catch any one given fish is $\frac{1}{4}$.

Problem 19 (expectations of functions of random variables)

For each of the various parts we will apply Jensen's inequality $E[f(X)] \geq f(E[X])$ which requires $f(x)$ to be convex i.e. $f''(x) \geq 0$. Now since we are told that $E[X] = 25$ we can compute the following.

Part (a): For the function $f(x) = x^3$, we have that $f''(x) = 6x \geq 0$ since we are told that X is a nonnegative random variable. Thus Jensen's inequality gives

$$E[X^3] \geq 25^3 = 15625\,.$$

Part (b): For the function $f(x) = \sqrt{x}$, we have that $f'(x) = \frac{1}{2\sqrt{x}}$, and $f''(x) = -\frac{1}{4\sqrt{x}} < 0$. Thus $f(x)$ is not a convex function but $-f(x)$ is. Applying Jensen's inequality to $-f(x)$ gives $E[-\sqrt{X}] \geq -\sqrt{25} = -5$ or

$$E[\sqrt{X}] \leq 5\,.$$

Part (c): For the function $f(x) = \log(x)$, we have that $f'(x) = \frac{1}{x}$, and $f''(x) = -\frac{1}{x^2} < 0$. Thus $f(x)$ is not a convex function but $-f(x)$ is. Applying Jensen's inequality to $-f(x)$ gives $E[-\log(X)] \geq -\log(25)$ or

$$E[\log(X)] \leq \log(25)\,.$$

Part (d): For the function $f(x) = e^{-x}$, we have that $f''(x) = e^{-x} > 0$. Thus $f(x)$ is a convex function. Applying Jensen's inequality to $f(x)$ gives

$$E[e^{-X}] \geq e^{E[X]} = e^{25}\,.$$

Problem 20 ($E[X] \leq (E[X^2]^{1/2} \leq (E[X^3])^{1/3} \leq \cdots$)

Now Jensens' inequality is that if $f(x)$ is a convex function then $E[f(x)] \geq f(E[X])$. If f is invertable then this is equivalent to

$$E[X] \leq f^{-1}(E[f(X)])\,,$$

which will be functional equation we will use to derive the requested results. Now to show the first stage of the inequality sequence let $f(x) = x^2$, then $f''(x) = 2 > 0$ so $f(\cdot)$ is convex and $f^{-1}(x) = x^{1/2}$. An application of the above functional expression gives

$$E[X] \leq (E[X^2])^{1/2} .$$

To show that $E[X] \leq E[X^3]^{1/3}$ one could perform the same logic with the function $f(x) = x^3$.

To show the expression $E[X^2]^{1/2} \leq E[X^3]^{1/3}$, we will apply Jensen's inequlity a second time. For this second application let $Y = f(X)$ then $E[f(X)] = E[Y] \leq g^{-1}(E[g(Y)])$ for any convex $g(\cdot)$. Thus

$$E[f(X)] \leq g^{-1}(E[g(f(X))]) .$$

On defining $Y = f(X)$ (and $X = f^{-1}(Y)$) we have that

$$E[f^{-1}(Y)] \leq f^{-1}(g^{-1}(E[g(Y)])) ,$$

or

$$f(E[f^{-1}(Y)]) \leq g^{-1}(E[g(Y)]) .$$

Thus beginning with the function pair $f(x) = x$ and $g(x) = x^2$ we have

$$E[Y] \leq (E[Y^2])^{1/2} ,$$

or the first inequality. For the second inequality we can take a function pair to consist of $f^{-1}(x) = x^2$ (so that $f(x) = x^{1/2}$) and $g(x) = x^3$ (so that $g^{-1}(x) = x^{1/3}$) then we have that

$$(E[Y^2])^{1/2} \leq (E[Y^3])^{1/3} .$$

For the third inequality we can take f and g, such that $f^{-1}(x) = x^3$ (so that $f(x) = x^{1/3}$) and $g(x) = x^4$ (so that $g^{-1}(x) = x^{1/4}$). Then

$$(E[Y^3])^{1/3} \leq (E[Y^4])^{1/4} ,$$

These relationships can be continued in general.

Problem 5

We desire to prove the following, if we define $B_n(x)$ as

$$B_n(x) = \sum_{k=0}^{n} f(\frac{k}{n}) \left(\begin{array}{c} n \\ k \end{array} \right) x^k (1-x)^{n-k} ,$$

then we want to show that $\lim_{n\to\infty} B_n(x) = f(x)$. To do this begin by defining $X_1, X_2, \cdots, X_k$ to be independent random variables each with mean x, then $E\left[f\left(\frac{X_1+X_2+X_3+\cdots+X_n}{n}\right)\right]$ can be evaluated by first noting that if X_i are Bernoulli random variables with mean x then $X_1 + X_2 + \cdots + X_n$ is a Binomial random variable with parameters (n, x) and thus

$$Pr\{\sum_{i=1}^{n} X_i = k\} = \left(\begin{array}{c} n \\ k \end{array} \right) x^k (1-x)^{n-k} ,$$

so that

$$E\left[f\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)\right] = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\Pr\left\{\sum_{i=1}^{n}X_i = k\right\},$$

by the definition of expectation. Continuing we have that

$$E\left[f\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)\right] = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\left(\begin{array}{c} n \\ k \end{array}\right)x^k(1-x)^{n-k}.$$

Now if we can show that when we define $Z_n = \frac{1}{n}\sum_{i=1}^{n}X_i$ that

$$\Pr\{|Z_n - x| > \epsilon\} \to 0 \quad \text{as} \quad n \to \infty,$$

then from Theoretical exercise number 4 we have that

$$E[f(Z_n)] \to f(z) \quad \text{as} \quad n \to \infty,$$

and we have proven the the famous Weierstrass theorem from analysis. Now from the central limit theorem we have that the random variable

$$\frac{\frac{1}{n}\sum_{i=1}^{n}X_i - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} = \frac{\frac{1}{n}\sum_{i=1}^{n}X_i - x}{\left(\frac{\sigma}{\sqrt{n}}\right)},$$

tends to the standard normal as $n \to \infty$. With this result we have that the probability that we desire to bound

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i - x\right|\right\}$$

is equivalent to

$$P\left\{\frac{\left|\frac{1}{n}\sum_{i=1}^{n}X_i - x\right|}{\left(\frac{\sigma}{\sqrt{n}}\right)} > \frac{\epsilon}{\left(\frac{\sigma}{\sqrt{n}}\right)}\right\}.$$

By the central limit this is equal to $2\Phi(-\frac{\epsilon}{\sigma}\sqrt{n})$. Since as $n \to \infty$ we have that $-\frac{\epsilon}{\sigma}\sqrt{n} \to -\infty$ so that

$$\Phi(-\frac{\epsilon}{\sigma}\sqrt{n}) \to 0,$$

and we have the condition required in problem number 4 and have proven the desired result.

Chapter 8: Theoretical Exercises

Problem 1 (an alternate Chebyshev inequality)

Now the Chebyshev inequality is given by

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}.$$

Defining $k = \sigma\kappa$ the above becomes

$$P\{|X - \mu| \geq \sigma\kappa\} \leq \frac{\sigma^2}{\sigma^2\kappa^2} = \frac{1}{\kappa^2},$$

which is the desired inequality.

Problem 10

Using the Chernoff bound of $P\{X \leq a\} \leq e^{-ta}M(t)$, we recall that if X is a Poisson random variable its moment generating function is given by $M(t) = e^{\lambda(e^t-1)}$ so

$$P\{X \leq i\} \leq e^{-ti}e^{\lambda(e^t-1)} \quad \text{for} \quad t < 0\,,$$

To minimized the right hand side of this expression is equivalent to minimizing $-ti+\lambda(e^t-1)$. Taking the derivative with respect to t and setting it equal to zero we have

$$-i + \lambda e^t = 0\,.$$

Solving for t gives $t = \ln(i/\lambda)$. Since $i < \lambda$ this t is negative as required. Putting this into the expression above gives

$$\begin{aligned}
P\{X \leq i\} &\leq e^{-i\ln(i/\lambda)}e^{\lambda(e^{\ln(i/\lambda)}-1)} \\
&= e^{\ln\left(\left(\frac{i}{\lambda}\right)^{-i}\right)}e^{\lambda(i/\lambda-1)} \\
&= e^{-\lambda}\frac{(\lambda e)^i}{i^i}\,.
\end{aligned}$$

Problem 12 (an upper bound on the complemetary error function)

From the definition of the normal density we have that

$$P\{X > a\} = \int_a^\infty \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx\,,$$

which we can simplify by the following change of variable. Let $v = x - a$ (then $dv = dx$) and the above becomes

$$\begin{aligned}
P\{X > a\} &= \int_0^\infty \frac{1}{\sqrt{2\pi}}e^{-(v+a)^2/2}dv \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi}}e^{-(v^2+2va+a^2)/2}dv \\
&= \frac{e^{-\frac{a^2}{2}}}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{v^2}{2}}\,e^{-va}dv \\
&\leq \frac{e^{-\frac{a^2}{2}}}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{v^2}{2}}dv\,,
\end{aligned}$$

since $e^{-va} \leq 1$ for all $v \in [0, \infty)$ and $a > 0$. Now because of the identity

$$\int_0^\infty e^{-\frac{v^2}{2}}dv = \sqrt{\frac{\pi}{2}}\,,$$

we see that the above becomes

$$P\{X > a\} \leq \frac{1}{2}e^{-\frac{a^2}{2}}\,.$$

Problem 13 (a problem with expectations)

We are assuming that if $E[X] < 0$ and $\theta \neq 0$ such that $E[e^{\theta X}] = 1$, and want to show that $\theta > 0$. To do this recall Jensen's inequality which for a convex function f and an arbitrary random variable Y is given by

$$E[f(Y)] \geq f(E[Y]) .$$

If we let the random variable $Y = e^{\theta X}$ and the function $f(y) = -\ln(y)$, then Jensen's inequality becomes (since this function f is convex)

$$-E[\theta X] \geq -\ln(E[e^{\theta X}]) ,$$

or using the information from the problem we have

$$\theta E[X] \leq \ln(1) = 0 .$$

Now since $E[X] < 0$ by dividing by this expression we have $\theta > 0$ as was to be shown.

Chapter 9 (Additional Topics in Probability)

Chapter 9: Problems

Problem 2 (helping Al cross the highway)

At the point where Al wants to cross the highway the number of cars that cross is a Poisson process with rate $\lambda = 3$, the probability that k cars appear in t time is given by

$$P\{N = k\} = \frac{e^{-\lambda t}(\lambda t)^k}{k!}.$$

Thus Al will have no problem in the case when *no* cars come during her crossing. If her crossing time takes s second this will happen with probability

$$P\{N = 0\} = e^{-\lambda s} = e^{-3s}.$$

Note that this is the density function for a Poisson random variable (or the cumulative distribution function of a Poisson random variable with $n = 0$). This expression is tabulated for $s = 2, 5, 10, 20$ seconds in `chap_9_prob_2.m`.

Problem 3 (helping a nimble Al cross the highway)

Following the results from Problem 2, Al will cross unhurt, with probability

$$P\{N = 0\} + P\{N = 1\} = e^{-\lambda s} + e^{-\lambda s}(\lambda s) = e^{-3s} + 3se^{-3s}.$$

Note that this is the cumulative distribution function for a Poisson random variable. This expression is tabulated for $s = 5, 10, 20, 30$ seconds in `chap_9_prob_3.m`.

Chapter 10 (Simulation)

Chapter 10: Problems

Problem 2 (simulating a specified random variable)

Assuming our random variable has a density given by

$$f(x) = \begin{cases} e^{2x} & -\infty < x < 0 \\ e^{-2x} & 0 < x < \infty \end{cases}$$

Lets compute the cumulative distribution $F(x)$ for this density function. This is needed if we simulate from f using the inverse transformation method. We find that

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} e^{2\xi} d\xi \quad \text{for} \quad -\infty < x < 0 \\ &= \left. \frac{e^{2\xi}}{2} \right|_{-\infty}^{x} = \frac{1}{2} e^{2x} . \end{aligned}$$

and that

$$\begin{aligned} F(x) &= \frac{1}{2} + \int_{0}^{x} e^{-2\xi} d\xi \quad \text{for} \quad 0 < x < \infty \\ &= \frac{1}{2} + \left. \frac{e^{-2\xi}}{(-2)} \right|_{0}^{x} = 1 - \frac{1}{2} e^{-2x} . \end{aligned}$$

Then to simulate from the density $f(\cdot)$ we require the inverse of this cumulative probability density function. Since our F is given in terms of two different domains we will compute this inverse function in the same way. If $0 < y < \frac{1}{2}$, then the equation we need to invert i.e. $y = F(x)$ is equivalent to

$$y = \frac{1}{2} e^{2x} \quad \text{or} \quad x = \frac{1}{2} \ln(2y) \quad \text{for} \quad 0 < y < \frac{1}{2}$$

While if $\frac{1}{2} < y < 1$ then $y = F(x)$ is equivalent to

$$y = 1 - \frac{1}{2} e^{-2x} ,$$

or by solving for x we find that

$$x = -\frac{1}{2} \ln(2(1-y)) \quad \text{for} \quad \frac{1}{2} < y < 1 .$$

Thus combining these two results we find that

$$F^{-1}(y) = \begin{cases} \frac{1}{2} \ln(2y) & 0 < y < \frac{1}{2} \\ -\frac{1}{2} \ln(2(1-y)) & \frac{1}{2} < y < 1 \end{cases}$$

Thus our simulation method would repeatedly generate uniform random variables $U \in (0, 1)$ and apply $F^{-1}(U)$ (defined above) to them computing the corresponding y's. These y's are guaranteed to be derived from our density function f.

References

[1] J. Bewersdorff. *Luck, Logic, and White Lies*. AK Peters, 2004.