Math3C: Homework 6

by Wesley Kerr (TA: Section 3a and 3b)

You should NOT look at these answers before you attempt the problems. If you do, then you're not going to learn as well. Note that these answers are not reviewed by Dr. Weisbert and/or may include some errors (as we figured out first week). If you find one or more, let me know and I'll edit the document. My email is wesleytk@ucla.edu.

Exercise 1. You toss a biased coin seven times. The probability the coin lands on heads is $\frac{1}{3}$. What is the probability that the coin lands on heads exactly four times?

This is exactly a binomial distribution because each coin flip is independent. $n = 7$, $p = \frac{1}{3}$, $k = 4$.

$$P(4H, 3T) = \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^3 \binom{7}{4}$$

Exercise 2. You toss a biased coin nine times. The probability the coin lands on heads is $\frac{1}{3}$. What is the probability that the coin lands on heads at least two times?

There's a hard way to do this and an easy way. Let's do the easy way. We know:

$$P(\#H \geq 2) = 1 - P(\#H < 2) = 1 - P(\#H = 0) - P(\#H = 1)$$

Now that's easier. For the two terms we do need to solve: $n = 9, p = \frac{1}{3}$. Giving us:

$$1 - \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^9 \binom{9}{0} - \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^8 \binom{9}{1}$$

Exercise 3. A disease has a prevalence of $\frac{1}{10}$ in a population. You come into contact with 10 people. What is the probability that you are exposed to the disease? What is the expected number of exposures?

Let's assume that you run into people independently, which is entirely not true, but we can't solve the problem if it's not true.

The probability that you are exposed to the disease is: $P(N \geq 1) = \sum_{k=1}^{10} P(N = k) = 1 - P(N = 0)$ when $n = 10$ and $p = \frac{1}{10}$. This gives us:

$$P(N \geq 1) = 1 - \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^1 0 \binom{10}{0}$$

Now for the expected number of people. We know for binomially distributed variables that:

$$E(N|n = 10) = n \cdot E(N|n = 1) = n \cdot p = 10 \cdot \frac{1}{10} = 1$$

Exercise 4. Disease A has a prevalence of $\frac{1}{4}$ in a population. Disease B has a prevalence of $\frac{1}{8}$ in a population. Disease C has a prevalence of $\frac{1}{10}$ in a population. Let A, B, and C be the events that a randomly selected person has disease A, B and C, respectively. Suppose that $\{A, B, C\}$ are an independent set of events. If you come into contact with 12 randomly selected people, what is the probability that you are exposed to at least one of the three diseases?

Let's write this out in probability speak. The probability of at least 1 is:

$$P(\geq 1 \text{ Disease}) = 1 - P(\text{no exposures})$$

The latter probability is MUCH easier to calculate (are you sensing a trend?). Each are independent events, and for each combination of number of exposures to A, you can have every combination of number of exposures to B (or C). That means we're going to multiply probability. That gives us:

$$\begin{aligned} P(\geq 1 \text{ Disease}) =& 1 - P\left(A = 0 | n = 12, p = \frac{1}{4}\right) \cdot P\left(B = 0 | n = 12, p = \frac{1}{8}\right) \cdot P\left(C = 0 | n = 12, p = \frac{1}{10}\right) \\ =& 1 - \left[\binom{12}{0}\right]^3 \left(\frac{3}{4} \cdot \frac{7}{8} \cdot \frac{11}{12}\right)^{12} \\ =& 1 - \left(\frac{3}{4} \cdot \frac{7}{8} \cdot \frac{11}{12}\right)^{12} \end{aligned}$$

Exercise 5. A DNA test is such that there is only a $\frac{1}{10{,}000}$ chance that a randomly selected invidivual will match a certain blood sample. Suppose a population of 50,000 inmates is screened using this test. What is the expected number of matches? What is the probability that there will be at least one match? What is the probability that there will be at least two matches? What does this tell us about screening a population in this way in order to identify the perpetrator of a crime?

Firstly, the expected number of matches. Let's assume the inmates are not related at all (not a safe assumption, but it's necessary). Then this is just a big binomial problem with $n = 50,000$ and $p = \frac{1}{10{,}000}$.

$$\begin{aligned} E(\text{matches}) =& np = 5 \\ P(\#\text{matches} \geq 1) =& 1 - P(\text{no matches}) = 1 - \left(\frac{9,999}{10,000}\right)^{50{,}000} \\ \approx& 99.3\% \\ P(\#\text{matches} \geq 2) =& 1 - P(\text{no matches}) - P(\#\text{matches} = 1) \\ =& 1 - \left(\frac{9,999}{10,000}\right)^{50{,}000} - \binom{50,000}{1} \frac{1}{10,000} \left(\frac{9,999}{10,000}\right)^{49{,}999} \\ \approx& 96\% \end{aligned}$$

This is not very effective to find the individual perpetrator, but it is a good way to split down to 5 from 50,000. (The latter part of that sentence is not what Dr. Weisbart was thinking of.)

Exercise 6. You toss a coin five times. Each time it lands on heads, you double your money. Each time it lands on tails you lose half of your money. You start with two dollars. What is the expected amount that you earn?

This time, you're not adding or subtracting a fixed amount. That means that we can't use our helpful $E(aX + b) = aE(X) + b$ formula. Multiplying is non-linear, so life becomes unhappy.

$$\begin{aligned} E(\$) =& 2 \cdot \left[P(HHHHH) \cdot 2^5 + P(HHHHT) \cdot 2^4 + P(HHHTT) \cdot 2^3 + P(HHTTT) \cdot 2^2 \right. \\ & \left. + P(HTTTT) \cdot 2 + P(TTTTT) \cdot 1\right] \end{aligned}$$

Note that the intial 2 is multiplied by everything, so we factored it out. Calculating this:

$$\begin{aligned}E(\$) =& 2 \cdot \left[\frac{1}{2^5} \cdot 2^5 + \binom{5}{4}\frac{1}{2^5} \cdot 2^4 + \binom{5}{3}\frac{1}{2^5} \cdot 2^3 + \binom{5}{2}\frac{1}{2^5}2^2\right. \\ &\left. + \binom{5}{1}\frac{1}{2^5}2 + \frac{1}{2^5} \cdot 1\right]\end{aligned}$$

Note that the probability of success and failure are the same, so we just combine them and are left with the binomial count term. This gives us:

$$\begin{aligned}E(\$) =& \frac{1}{2^4}\left[2^5 + \binom{5}{4}2^4 + \binom{5}{3}2^3 + \binom{5}{2}2^2 + \binom{5}{1}2 + 1\right] \\ \approx& 15.18\end{aligned}$$

Exercise 7. You toss a coin five times. Each time it lands on heads, you double your money. Each time it lands on tails you lose half of your money. You start with two dollars. What is the expected amount that you earn?

This time, you're not adding or subtracting a fixed amount. That means that we can't use our helpful $E(aX + b) = aE(X) + b$ formula. Multiplying is non-linear, so life becomes unhappy.

$$\begin{aligned}E(\$) =& 2 \cdot \left[P(HHHHH) \cdot 2^5 + P(HHHHT) \cdot 2^3 + P(HHHTT) \cdot 2 + P(HHTTT) \cdot \frac{1}{2}\right. \\ &\left. + P(HTTTT) \cdot \frac{1}{2^3} + P(TTTTT) \cdot \frac{1}{2^5}\right]\end{aligned}$$

Note that the intial 2 is multiplied by everything, so we factored it out. Also note that the power changes by 2 each time because you take away one power of 2 and add a power of $\frac{1}{2}$ at the same time. Calculating this:

$$\begin{aligned}E(\$) =& 2 \cdot \left[\frac{1}{2^5} \cdot 2^5 + \binom{5}{4}\frac{1}{2^5} \cdot 2^3 + \binom{5}{3}\frac{1}{2^5} \cdot 2 + \binom{5}{2}\frac{1}{2^5} \cdot \frac{1}{2}\right. \\ &\left. + \binom{5}{1}\frac{1}{2^5} \cdot \frac{1}{2^3} + \frac{1}{2^5} \cdot \frac{1}{2^5}\right]\end{aligned}$$

Note that the probability of success and failure are the same, so we just combine them and are left with the binomial count term. This gives us:

$$\begin{aligned}E(\$) =& \frac{1}{2^4}\left[2^5 + \binom{5}{4}2^3 + \binom{5}{3}2 + \binom{5}{2}\frac{1}{2} + \binom{5}{1}\frac{1}{2^3} + \frac{1}{2^5}\right] \\ \approx& 6.10\end{aligned}$$

Exercise 8. Roll an eight sided die and let X be the value of the side on which the die lands. Compute $E(X)$ and $Var(X)$.

First the easy one: $E(X)$:

$$E(X) = \frac{1 + 2 + 3 + 4 + 5 + 6 + 7 + 8}{8} = 4.5$$

Now the harder one:

$$\begin{aligned}
Var(X) =& \frac{(1-4.5)^2}{8} + \frac{(2-4.5)^2}{8} + \frac{(3-4.5)^2}{8} + \frac{(4-4.5)^2}{8} \\
&+ \frac{(5-4.5)^2}{8} + \frac{(6-4.5)^2}{8} + \frac{(7-4.5)^2}{8} + \frac{(8-4.5)^2}{8} \\
=& 2\left[\frac{3.5^2+2.5^2+1.5^2+0.5^2}{8}\right] \\
=& \left[\frac{3.5^2+2.5^2+1.5^2+0.5^2}{4}\right] \\
=& 5.25
\end{aligned}$$

Exercise 9. Let S_n measure the number of heads on n tosses of a fair coin. Computer $E[S_4]$ and $Var[S_4]$.

We know the mean already (see above):

$$E(S_4) = np = 4\frac{1}{2} = 2$$

There is a handy formula for the Variance of a binomial random variable:

$$Var(S_4) = np(1-p) = 4\frac{1}{4} = 1$$

Exercise 10. Suppose that a random variable X takes on only positive values and is such that $E(X) = 5$. Find an upper bound on the probability that X takes a value of greater than 40.

This should set off all your Markov formula bells.

$$\begin{aligned}
E(X \geq a) \leq& \frac{E(X)}{a} \\
E(X \geq 40) \leq& \frac{5}{40} = \frac{1}{8}
\end{aligned}$$

Exercise 11. Suppose that a random variable X measures the height of individual in a population. Suppose further that $Var[X] = 0.25\text{inches}^2$. Find an upper bound on the probability that a randomly selected individual is more than 3 inches taller or shorter than the mean height in the population.

Chebychev bells now (height cannot be negative in humans, animals, bacteria, or even aliens).

$$\begin{aligned}
P(|X - E(X)| \geq c) \leq& \frac{Var(X)}{c^2} \\
P(|X - E(X)| \geq 3) \leq& \frac{0.25}{3^2} \approx 2.7\%
\end{aligned}$$

Exercise 12. Each time you roll a fair six sided die, you earn double the value of the roll in dollars. What is the expected amount that you earn on a given game? If X measures the amount of money you have at the end of the game, what is the variance of X?

This can be done pretty directly:

$$\begin{aligned}E(X) =&\frac{2+4+6+8+10+12}{6} = 7\\Var(X) =&\frac{(2-7)^2+(4-7)^2+(6-7)^2+(8-7)^2+(10-7)^2+(12-7)^2}{6}\\=&2\frac{5^2+3^2+1}{6} = \frac{35}{3} \approx 11.67\end{aligned}$$

Exercise 13. You toss a coin 12 times. Each time the coin lands on heads, you earn five dollars. What is the expected amount that you earn? If X measures the amount of money you have at the end of the game, what is the variance of X?

This is just a transformation of X. Let the number of heads be Y, then $X = 5Y$. Then:

$$\begin{aligned}E(X) =&E(5Y) = 5E(Y) = 5np = 5 \cdot 12 \cdot \frac{1}{2} = 5 \cdot 6 = 30\\Var(X) =&Var(5Y) = 5^2 Var(Y) = 5^2 np(1-p) = 5^2 \cdot 12 \cdot \frac{1}{2^2} = 5^2 \cdot 3 = 75\end{aligned}$$

Exercise 14. Suppose that you toss a fair coin 10,000 times. Estimate the probability that the coin lands on heads between 4,000 and 6,000 times.

There's actually way that you can figure this out more directly, but we'll use what you guys know. If you toss a coin 10,000 times, then $E(X) = 5,000$. We also know that $Var(X) = 10,000 \cdot \frac{1}{2^2} = 2,500$. Then we think about Chebychev:

$$\begin{aligned}P(|X - E(X)| \geq c) \leq&\frac{Var(X)}{c^2}\\P(|X - E(X)| \geq 1,000) \leq&\frac{2,500}{1,000^2} = 0.25\%\end{aligned}$$

Exercise 15. Suppose that you toss a coin 20,000 times. The coins land on heads 5,000 times. How small can c be if you say with at least 99% certainty that the probability of heads is within c of 25%?

This is the Chebychev with a trick. Let's assume that $p = \frac{1}{4}$ (note that we're not accounting for the variation in this assumption, so this answer is not perfectly correct, but it's as correct as we can make with the tools that we have so far). We then know that, for X being the number of heads:

$$\begin{aligned}E(X) =&np = 5,000\\Var(X) =&np(1-p) = 5,000\frac{3}{4}\end{aligned}$$

Then Chebychev says:

$$\begin{aligned}
P(|X-E(X)| \geq c) \leq & \frac{Var(X)}{c^2} \\
P\left(\frac{|X-E(X)|}{20,000} \geq c\right) \leq & \frac{Var(X)}{c^2 20,000^2} \\
P\left(|p-25\%| \geq c\right) \leq & \frac{5,000\frac{3}{4}}{c^2 20,000^2} \\
1-P\left(|p-25\%| < c\right) \geq & \frac{5,000\frac{3}{4}}{c^2 20,000^2} \\
1-99\% \geq & \frac{5,000\frac{3}{4}}{c^2 20,000^2} \\
0.01 \geq & \frac{5,000\frac{3}{4}}{c^2 20,000^2} \\
c^2 \geq & \frac{5,000\frac{3}{4}}{0.01 \cdot 20,000^2} \\
c \geq & \sqrt{\frac{5,000\frac{3}{4}}{0.01 \cdot 20,000^2}} \approx 0.03
\end{aligned}$$

Exercise 16. Suppose that you toss a coin n times and experimentally determine that the probability of heads is 40%. How many times must you toss the coin in order to be 95% certain that the theoretical probability for heads is between 38% and 42%?

This is the same as before, except with variables now.

$$\begin{aligned}
E(H) = & n\frac{4}{10} \\
Var(H) = & n\frac{4}{10}\frac{6}{10} = n\frac{24}{100}
\end{aligned}$$

Then Chebychev says:

$$\begin{aligned}
P(|X-E(X)| \geq c) \leq & \frac{Var(X)}{c^2} \\
P\left(\frac{|X-E(X)|}{n} \geq x\right) \leq & \frac{Var(X)}{n^2 c^2} \\
P\left(\left|p-\frac{1}{4}\right| \geq 0.02\%\right) \leq & \frac{n\frac{24}{100}}{n^2(0.02)^2} \\
1-P\left(\left|p-\frac{1}{4}\right| < 0.02\%\right) \geq & \frac{\frac{24}{100}}{n(0.02)^2} \\
1-0.95 \geq & \frac{1}{n(0.02)^2}\frac{24}{100} \\
0.05 \geq & \frac{1}{n(0.02)^2}\frac{24}{100} \\
n \geq & \frac{1}{0.05 \cdot (0.02)^2}\frac{24}{100} \\
n \geq & 12,000
\end{aligned}$$

Exercise 17. Suppose that you have a revolver that holds up to give bullets. Suppose that exactly one bullet is loaded. You spin the cylinder and then pull the trigger. You repeat this process until the gun fires. What is the probability that the gun fires on the fourth trial? What is the probability that it fires before the fourth trial? What is the probability that it fires after the third trial? What is the probability that it fires on the seventh trial given that it does not fire on the first five trials? What is the probability that it fires on the fifth trial, given that it fires on an odd trial? What is the expected number of trials?

Welcome to the geometric distribution where

$$P(X = k) = (1 - p)^{k-1} p$$

In this case, $p = \frac{1}{5}$. So let's go through these one by one:

$$P(X = 4) = \left(\frac{4}{5}\right)^3 \frac{1}{5}$$

Next:

$$\begin{aligned} P(X < 4) =& P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3) \\ =& \frac{1}{5} + \frac{4}{5}\frac{1}{5} + \frac{4^2}{5^3} \\ =& \frac{1}{5}\left[1 + \frac{4}{5} + \frac{4^2}{5^2}\right] \end{aligned}$$

Next:

$$\begin{aligned} P(X > 3) =& 1 - P(X \leq 2) = 1 - P(X = 1) - P(X = 2) \\ =& 1 - \frac{1}{5} - \frac{4}{5}\frac{1}{5} \\ =& 1 - \frac{1}{5}\left[1 + \frac{4}{5}\right] = 1 - \frac{1}{5}\frac{9}{5} = \frac{16}{25} \end{aligned}$$

Next:

$$P(X = 7|X > 5) = P(X = 2) = \frac{4}{5}\frac{1}{5}$$

Next:

$$\begin{aligned} P(X = 5|X = \text{odd}) =& \frac{P(X = 5 \cap X = \text{odd})}{P(X = \text{odd})} = \frac{P(X = 5)}{P(X = \text{odd})} = \frac{\left(\frac{4}{5}\right)^4 \frac{1}{5}}{\sum_{k=1}^{\infty} \left(\frac{4}{5}\right)^{(2k-1)-1} \frac{1}{5}} \\ =& \frac{\left(\frac{4}{5}\right)^4}{\sum_{k=1}^{\infty} \left(\frac{4}{5}\right)^{(2k-1)-1}} \\ & \text{See the next question} \\ =& \frac{\left(\frac{4}{5}\right)^4}{\frac{5^2}{9}} = \frac{3^2 4^4}{5^6} \end{aligned}$$

Lastly,

$$E(X) = \frac{1}{p} = \frac{5}{1} = 5.$$

Exercise 18. Suppose that you have a revolver that holds up to five bullets and you use this gun to play a friendly game of Russian roulette with your friend. You randomize the position of the bullet between trials. You have your friend go first and you will go second. What is the probability that you lose the game and what is the probability that your friend loses the game?

The probability that your friend loses is the probability it ends on an odd trial:

$$P(X \text{ odd}) = \sum_{k=1}^{\infty}(1-p)^{(2k-1)-1}p = p\sum_{k=1}^{\infty}(1-p)^{2k-2} = \frac{1}{5}\sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-2}$$

The probability that you lose is the probability it ends on an even trial:

$$P(X \text{ even}) = \sum_{k=1}^{\infty}(1-p)^{2k-1}p = \frac{1}{5}\sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-1}$$

But wait, we see something interesting here:

$$\begin{aligned}
1 =& \frac{1}{5}\sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-2} + \frac{1}{5}\sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-1} \\
5 =& \sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-2} + \frac{4}{5}\sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-2} \\
5 =& \left[1+\frac{4}{5}\right]\sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-2} \\
\frac{5}{\frac{9}{5}} =& \sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-2} \\
\frac{5^2}{9} =& \sum_{k=1}^{\infty}\left(\frac{4}{5}\right)^{2k-2}
\end{aligned}$$

Therefore:

$$\begin{aligned}
P(X \text{ odd}) =& \frac{1}{5}\frac{5^2}{9} = \frac{5}{9} \\
P(X \text{ even}) =& \frac{1}{5}\frac{4}{5}\frac{5^2}{9} = \frac{4}{9}
\end{aligned}$$

Exercise 19. Suppose that you have a revolver that holds up to five bullets and you use this gun to play a friendly game of Russian roulette with your friend. You do NOT randomize the position of the bullet between trials. You have your friend go first and you will go second. What is the probability that you lose the game and what is the probability that your friend loses the game?

Now the events are NOT independent. Your friend seems to be getting the short end of these deals. Let's figure out:

$$\begin{aligned}
P(X \text{ odd}) =& P(X=1)+P(X=3)+P(X=5)\\
=& \frac{1}{5}+\frac{4}{5}\frac{3}{4}\frac{1}{3}+\frac{4}{5}\frac{3}{4}\frac{2}{3}\frac{1}{2}\frac{1}{1}\\
=& \frac{1}{5}+\frac{1}{5}+\frac{1}{5}\\
=& \frac{3}{5}
\end{aligned}$$

Note that the order matters. Once you find out that one chamber is empty (the first has probability $\frac{4}{5}$) the bullet is one of the remaining chambers, which are then equally likely to hold the bullet. The probability that you lose 1- that.

$$P(X \text{ even}) = \frac{2}{5}$$

You'd rather play this way than the way that it was played in #18.

Exercise 20. You have a revolver that holds up to give bullets and a revolver that holds up to eight bullets. You simultaneously pull each gun's trigger and then randomize the position of the bullets by spinning the cylinders. You repeat this until each gun fires at least one time. What is the expected number of trials for both guns to fire? Suppose instead that you simultaneously pull the triggers until a gun fires. Now, what is the expected number of trials?

First off, this question is HARD (mostly because it's messy). I wouldn't expect you to be able to answer it within the limits of a midterm, quiz or final. This is "how deep the rabbit hole goes." Let's do the second part first ($Y=\min\{X_5, X_8\}$), because it seems easier. We know that the events are independent so:

$$P(X_5 \cap X_8) = P(X_5)P(X_8)$$

Then the expectation is:

$$E(Y) = 1\cdot P(Y=1) + 2\cdot P(Y=2) + 3\cdot P(Y=3) + \sum_{k=4}^{\infty} P(Y=k)$$

Now let's think about $P(Y=k)$:

$$\begin{aligned}
P(Y=k) =& P(X_5=k, X_8\geq k) + P(X_8=k, X_5>k)\\
=& P(X_5=k)P(X_8\geq k) + P(X_8=k)P(X_5\geq k+1)\\
=& \left(1-\frac{1}{5}\right)^{k-1}\frac{1}{5}\frac{1}{8}\sum_{i=k}^{\infty}\left(1-\frac{1}{8}\right)^{i-1} + \left(1-\frac{1}{8}\right)^{k-1}\frac{1}{8}\frac{1}{5}\sum_{i=k+1}^{\infty}\left(1-\frac{1}{5}\right)^{i-1}
\end{aligned}$$

Now we have a throw back to calculus, where we talked about series. Remember that:

$$\begin{aligned}\sum_{i=0}^{\infty} q^i =& \frac{1}{1-q}\\ \sum_{i=0}^{k-1} q^i + \sum_{i=k}^{\infty} q^i =& \frac{1}{1-q}\\ \frac{1-q^k}{1-q} + \sum_{i=k}^{\infty} q^i =& \frac{1}{1-q}\\ \sum_{i=k}^{\infty} q^i =& \frac{1-1+q^k}{1-q} = \frac{q^k}{1-q}\end{aligned}$$

We also can show this another way:

$$\sum_{i=k}^{\infty} q^i = q^k \sum_{i=0}^{\infty} q^i = q^k \frac{1}{1-q} = \frac{q^k}{1-q}$$

Now let's figure out how to plug that into the formulae we have above:

$$\begin{aligned}\sum_{i=k}^{\infty} \left(1-\frac{1}{8}\right)^{i-1} &= \sum_{i=k+1}^{\infty} \left(\frac{7}{8}\right)^i = \frac{\left(\frac{7}{8}\right)^{k+1}}{1-\frac{7}{8}} = 8\left(\frac{7}{8}\right)^{k+1}\\ \sum_{i=k+1}^{\infty} \left(1-\frac{1}{5}\right)^{i-1} &= \sum_{i=k+2}^{\infty} \left(\frac{4}{5}\right)^i = \frac{\left(\frac{4}{5}\right)^{k+2}}{1-\frac{4}{5}} = 5\left(\frac{4}{5}\right)^{k+2}\end{aligned}$$

Plugging that in, we get:

$$\begin{aligned}P(Y=k) &= \left(\frac{4}{5}\right)^{k-1} \frac{1}{5\cdot 8} 8 \left(\frac{7}{8}\right)^{k+1} + \left(\frac{7}{8}\right)^{k-1} \frac{1}{8\cdot 5} 5 \left(\frac{4}{5}\right)^{k+2}\\ &= \left(\frac{4}{5}\cdot\frac{7}{8}\right)^{k-1} \left[\frac{1}{5}\left(\frac{7}{8}\right)^2 + \frac{1}{8}\left(\frac{4}{5}\right)^3\right]\end{aligned}$$

So the expectation is now:

$$\begin{aligned}E(Y) &= \sum_{k=1}^{\infty} k \left(\frac{4}{5}\cdot\frac{7}{8}\right)^{k-1} \left[\frac{1}{5}\left(\frac{7}{8}\right)^2 + \frac{1}{8}\left(\frac{4}{5}\right)^3\right]\\ &= \left[\frac{1}{5}\left(\frac{7}{8}\right)^2 + \frac{1}{8}\left(\frac{4}{5}\right)^3\right] \left(\frac{4}{5}\cdot\frac{7}{8}\right)^{-1} \sum_{k=1}^{\infty} k \left(\frac{4}{5}\cdot\frac{7}{8}\right)^k\end{aligned}$$

Let's try to figure out what that sum is. Using the internet, we know that:

$$\begin{aligned}\sum_{k=1}^{\infty} kp^k =& \frac{p}{(p-1)^2}\\ p =& \frac{4}{5}\cdot\frac{7}{8}\\ \sum_{k=1}^{\infty} k\left(\frac{4}{5}\cdot\frac{7}{8}\right)^k =& \frac{\frac{4}{5}\cdot\frac{7}{8}}{\left(\frac{4}{5}\cdot\frac{7}{8}-1\right)^2}\end{aligned}$$

Plugging that in:

$$E(Y) = \left[\frac{1}{5}\left(\frac{7}{8}\right)^2 + \frac{1}{8}\left(\frac{4}{5}\right)^3\right]\left(\frac{4}{5}\cdot\frac{7}{8}\right)^{-1} \frac{\frac{4}{5}\cdot\frac{7}{8}}{\left(\frac{4}{5}\cdot\frac{7}{8}-1\right)^2}$$
$$= \left[\frac{1}{5}\left(\frac{7}{8}\right)^2 + \frac{1}{8}\left(\frac{4}{5}\right)^3\right]\left(\frac{4}{5}\cdot\frac{7}{8}-1\right)^{-2}$$

Now for the first part ($Z = \max\{X_5, X_8\}$). Now let's think about $P(Z = k)$:

$$\begin{aligned} P(Z=k) =& P(X_5 = k, X_8 \le k) + P(X_8 = k, X_5 < k) \\ =& P(X_5 = k)P(X_8 \le k) + P(X_8 = k)P(X_5 \le k-1) \\ =& \left(1-\frac{1}{5}\right)^{k-1} \frac{1}{5}\frac{1}{8}\sum_{i=0}^{k}\left(1-\frac{1}{8}\right)^{i-1} + \left(1-\frac{1}{8}\right)^{k-1}\frac{1}{8}\frac{1}{5}\sum_{i=0}^{k-1}\left(1-\frac{1}{5}\right)^{i-1} \end{aligned}$$

Using the formula that we know:

$$\sum_{i=0}^{k} q^i = \frac{1-q^{k-1}}{1-q}$$

Plugging that in:

$$\begin{aligned} \sum_{i=0}^{k}\left(1-\frac{1}{8}\right)^{i-1} =& \frac{8}{7}\sum_{i=0}^{k}\left(\frac{7}{8}\right)^i = \frac{8}{7}\frac{1-\left(\frac{7}{8}\right)^{k-1}}{1-\frac{7}{8}} = \frac{8^2}{7}\left[1-\left(\frac{7}{8}\right)^{k-1}\right] \\ \sum_{i=0}^{k-1}\left(1-\frac{1}{5}\right)^{i-1} =& \frac{5}{4}\sum_{i=0}^{k-1}\left(\frac{4}{5}\right)^i = \frac{5}{4}\frac{1-\left(\frac{4}{5}\right)^{k-2}}{1-\frac{4}{5}} = \frac{5^2}{4}\left[1-\left(\frac{4}{5}\right)^{k-2}\right] \end{aligned}$$

Inserting those terms into our probability, we get:

$$\begin{aligned} P(Z=k) &= \left(1-\frac{1}{5}\right)^{k-1}\frac{1}{5}\frac{1}{8}\frac{8^2}{7}\left[1-\left(\frac{7}{8}\right)^{k-1}\right] + \left(1-\frac{1}{8}\right)^{k-1}\frac{1}{8}\frac{1}{5}\frac{5^2}{4}\left[1-\left(\frac{4}{5}\right)^{k-2}\right] \\ &= \left(\frac{4}{5}\right)^{k-1}\frac{1}{5}\frac{8}{7}\left[1-\left(\frac{7}{8}\right)^{k-1}\right] + \left(\frac{7}{8}\right)^{k-1}\frac{1}{8}\frac{5}{4}\left[1-\left(\frac{4}{5}\right)^{k-2}\right] \end{aligned}$$

The expectation is then:

$$\begin{aligned} E(Z) =& \sum_{k=1}^{\infty} k\left[\left(\frac{4}{5}\right)^{k-1}\frac{1}{5}\frac{8}{7}\left[1-\left(\frac{7}{8}\right)^{k-1}\right] + \left(\frac{7}{8}\right)^{k-1}\frac{1}{8}\frac{5}{4}\left[1-\left(\frac{4}{5}\right)^{k-2}\right]\right] \\ =& \frac{1}{5}\frac{8}{7}\left[\sum_{k=1}^{\infty}k\left(\frac{4}{5}\right)^{k-1} - \sum_{k=1}^{\infty}k\left(\frac{4}{5}\right)^{k-1}\left(\frac{7}{8}\right)^{k-1}\right] + \frac{1}{8}\frac{5}{4}\left[\sum_{k=1}^{\infty}k\left(\frac{7}{8}\right)^{k-1} - \sum_{k=1}^{\infty}k\left(\frac{7}{8}\right)^{k-1}\left(\frac{4}{5}\right)^{k-2}\right] \end{aligned}$$

Remember from the previous sum we solved for the easier part of the problem:

$$\begin{aligned}
\sum_{k=1}^{\infty} k\left(\frac{4}{5}\right)^{k-1} =& \frac{5}{4}\sum_{k=1}^{\infty} k\left(\frac{4}{5}\right)^{k} = \frac{5}{4}\frac{\frac{4}{5}}{\left(\frac{4}{5}-1\right)^2} = \frac{1}{\left(\frac{1}{5}\right)^2} = 5^2 \\
\sum_{k=1}^{\infty} k\left(\frac{4}{5}\right)^{k-1}\left(\frac{7}{8}\right)^{k-1} =& \frac{8\cdot 5}{4\cdot 7}\sum_{k=1}^{\infty} k\left(\frac{4\cdot 7}{5\cdot 8}\right)^{k} = \left(1-\frac{4\cdot 7}{5\cdot 8}\right)^{-2} \\
\sum_{k=1}^{\infty} k\left(\frac{7}{8}\right)^{k-1} =& 8^2 \\
\sum_{k=1}^{\infty} k\left(\frac{7}{8}\right)^{k-1}\left(\frac{4}{5}\right)^{k-2} =& \left(1-\frac{7\cdot 4}{5\cdot 8}\right)^{-2}
\end{aligned}$$

Plugging that in, we get:

$$E(Z) = \frac{1}{5}\frac{8}{7}\left[5^2 - \left(1-\frac{4\cdot 7}{5\cdot 8}\right)^{-2}\right] + \frac{1}{8}\frac{5}{4}\left[8^2 - \left(1-\frac{4\cdot 7}{5\cdot 8}\right)^{-2}\right]$$

This could be simplified more, but the solution is probably not any more interpretable. We therefore conclude that calculating the expectation of a minimum or maximum of two random variables is hard and messy.

Exercise 21. A disease has a prevalence of $\frac{1}{8}$. You do not have the disease. The probability that you contract the disease on beign exposed is $\frac{1}{20}$. You have two choices: You may come into contact one time each with n people or arbitrarily many times with one person. How large may n be so that the first is safer? What is the probability that you will contract the disease in the second case if you come into contact with the same person 100 times?

Let's do the second question about the second case first because it's easier. This is:

$$\begin{aligned}
P(D|\text{contacts}=100) =& P(\text{Person has disease})P(\text{You contract it at least once in 100 tries}) \\
&+ P(\text{Person doesn't have the disease})\cdot 0 \\
=& \frac{1}{8}\left(1-P(\text{don't contract in 100 tries})\right) + \frac{7}{8}\cdot 0 \\
=& \frac{1}{8}\left[1-\left(\frac{19}{20}\right)^{100}\right]
\end{aligned}$$

Now let's think about the first question regarding the second case. If you can come in contact with that one person an arbitrary number of times, then you can come into contact with them an infinite number of times. That means:

$$P(D|\text{contacts}=\infty) = \frac{1}{8}\left[1-\left(\frac{19}{20}\right)^{\infty}\right] = \frac{1}{8}[1-0] = \frac{1}{8}$$

Now let's figure out the probability of being exposed after contact with n separate people. To do that, we think about one person in particular, which gives us:

$$P(D|\text{contacts}=1) = \frac{1}{8}\frac{1}{20}$$

Let's assume that each person that you come into contact with is independent, then we get the probability of disease after coming into contact with n people as:

$$\begin{aligned}P(D|\text{contacts} = 1 \text{ with } n \text{ people}) =& 1 - P(D^c|\text{contacts} = 1 \text{ with } n \text{ people})\\ =& 1 - \left(1 - \frac{1}{8}\frac{1}{20}\right)^n\end{aligned}$$

Now we make an equality:

$$\begin{aligned}P(D|\text{contacts} = \infty) >& P(D|\text{contacts} = 1 \text{ with } n \text{ people})\\ \frac{1}{8} >& 1 - \left(1 - \frac{1}{8}\frac{1}{20}\right)^n\\ -\frac{7}{8} >& -\left(1 - \frac{1}{8}\frac{1}{20}\right)^n\\ \frac{7}{8} <& \left(1 - \frac{1}{8}\frac{1}{20}\right)^n\\ & \text{This must be solved using logarithms, remember those?}\\ n <& \frac{\log \frac{7}{8}}{\log\left(1 - \frac{1}{8}\frac{1}{20}\right)}\end{aligned}$$

Exercise 22. Suppose that a female of a certain species of mammal is fertile precisely one day out of every 10 days. This is to say that a female is fertile on day one of her cycle, infertile for nine days and then is fertile once again and the cycle repeats. Neither the male nor the female of this species has any ability to know which day the female is fertile. A female who is fertile will always become pregnant if mating occurs. If the male mates with the female for 10 consecutive days, what is the probability that the female will not get pregnant? If instead, the male mates with 10 different randomly selected females during this 10 day period, what is the probability that not one will get pregnant? What is the expected number of pregnancies in each situtation? In the second situation, what is the expected number of matings for one pregnancy to occur? In the first situtation, what is the expected number of matings for one pregnancy to occur?

In the first case, the female is mated with on 10 days, and is thereby mated with on every possible day of cycle, making:

$$P(P = 1|\text{method 1}) = 1$$

The expected number of matings is also easy to calculate:

$$\begin{aligned}E(\text{matings till one P}|\text{method 1}) =& 1 \cdot \frac{1}{10} + 2\frac{9}{10}\frac{1}{9} + 3\frac{9}{10}\frac{8}{9}\frac{1}{8} + \cdots\\ =& 1 \cdot \frac{1}{10} + 2 \cdot \cdot \frac{1}{10} + 3 \cdot \cdot \frac{1}{10} + \cdots\\ =& \frac{1+2+3+4+5+6+7+8+9+10}{10} = 5.5\end{aligned}$$

Now let's talk about the second way. First, the expected number of matings till one pregnancy in the second case:

$$E(\text{matings till one P}|\text{method 2}) = 1 \cdot \frac{1}{10} + 2 \cdot \cdot \frac{1}{10} + 3 \cdot \cdot \frac{1}{10} + \cdots = 5.5$$

Now let's think about:

$$P(P = 0|\text{method 2}) = \left(\frac{1}{10}\right)^{10}$$

Note that this is a binomial distribution (females and matings are independent), so the expected number of pregnancies is:

$$E(P|\text{method 1}) = n \cdot p = 10 \cdot \frac{1}{10} = 1$$

Note, however, that even though the second method has a small probability of no pregnancies, that is balanced by the ability to create multiple pregnancies. Both strategies of mating are equally beneficial, but the second has more variation in pregnancy amount.

Exercise 23. You have four types of lightbulbs: Type A, Type B, Type C, and Type D. The lightbulbs look identical except that Type A lightbulbs explode when they burn out and the others simply burn out without making a show. It takes four hours for a Type A lightbulb to burn out (and explode). It takes two, three and six hours for a Type B, C and D lightbulb to burn out, respectively. As soon as a lightbulb burns out, you replace it with the hope that you will eventually choose a Type A and see a cool explosion. The likelihood of choosing a Type A, B, C, and D lightbulb is $\frac{1}{10}$, $\frac{4}{10}$, $\frac{2}{10}$, and $\frac{3}{10}$ respectively. What is the expected length of time you must wait until you see an explosion.

This is a conditional expectation problem.

$$\begin{aligned} E(T) =& P(A)E(T|A) + P(B)E(T|B) + P(C)E(T|C) + P(D)E(T|D) \\ =& \frac{1}{10} \cdot 4 + \frac{4}{10}\left[2 + E(T)\right] + \frac{2}{10}\left[3 + E(T)\right] + \frac{3}{10}\left[6 + E(T)\right] \end{aligned}$$

If you choose a 'wrong' lightbulb, you wait the given amount of hours then restart the process of waiting.

$$\begin{aligned} \left[1 - \frac{4+2+3}{10}\right] E(T) =& \frac{4+8+6+18}{10} \\ E(T) =& \frac{36}{10} \cdot \frac{10}{1} = 36 \end{aligned}$$

Exercise 24. You toss a biased coin until it lands on heads on three consecutive tosses. The probability that it lands on heads is $\frac{1}{5}$. What is the expected number of tosses?

Another conditional expectation problem.

$$E(T) = E(T|HHH)P(HHH) + E(T|HHT)P(HHT) + E(T|HT)P(HT) + E(T|T)P(T)$$

Note that this is built so that I'm worried about when the process of flipping starts occuring again, not using the law of total probability. This turns out to:

$$\begin{aligned} E(T) =& 3\left(\frac{1}{5}\right)^3 + \left[3 + E(T)\right]\frac{4}{5^3} + \left[2 + E(T)\right]\frac{4}{5^2} + \left[1 + E(T)\right]\frac{4}{5} \\ \left[1 - \frac{4}{5^3} - \frac{4}{5^2} - \frac{4}{5}\right] E(T) =& 3\left(\frac{1}{5}\right)^3 + 3\frac{4}{5^3} + 2\frac{4}{5^2} + 1\frac{4}{5} \\ \frac{5^3 - 4 - 4 \cdot 5 - 4 \cdot 5^2}{5^3} E(T) =& \frac{3 + 12 + 2 \cdot 4 \cdot 5 + 4 \cdot 5^2}{5^3} \\ E(T) =& \frac{3 + 12 + 40 + 100}{125 - 4 - 20 - 100} \\ E(T) =& \frac{155}{1} = 155 \end{aligned}$$

Exercise 25. You roll a fair four sided die until it lands on 1 twice in a row. What is the expected number of rolls?

Yay. Let's do it again for $X = \{2, 3, 4\}$:

$$\begin{aligned}
E(R) =& E(R|11)P(11) + E(R|1X)P(1X) + E(R|X)P(X) \\
=& \frac{2}{4^2} + [2 + E(R)]\frac{3}{4^2} + [1 + E(R)]\frac{3}{4} \\
\left[1 - \frac{3}{4^2} - \frac{3}{4}\right] E(R) =& \frac{2}{4^2} + 2\frac{3}{4^2} + 1\frac{3}{4} \\
\frac{16 - 3 - 12}{4^2} E(R) =& \frac{2 + 2 \cdot 3 + 1 \cdot 3 \cdot 4}{4^2} \\
E(R) =& \frac{2 + 6 + 12}{1} = 20
\end{aligned}$$

Warning

Just a warning, for everything except the $E(T_{HHH})$ problem, there's a problem here of where part of the incorrect sequence of events overlaps with the beginning of the correct sequence. This is how I was taught how to do it. I'm checking with Dr. Weisbart to make sure this is right, so no gaurantees on 100% correctness other than the $E(T_{HHH})$ problem.

Exercise 26. You roll a fair three sided die until it lands on the same side twice in a row. What is the expected number of rolls?

This one is different because we don't care what side happens twice. That means we get, for S means same and D means different:

$$\begin{aligned}
E(R) =& E(R|SS)P(SS) + E(R|SD)P(SD) \\
=& 2\frac{1}{3} + [1 + E(R)]\frac{2}{3}
\end{aligned}$$

Note that you only pay a one roll penalty after misrolling (instead of 2) because you've already rolled the first roll that you want to then match.

$$\begin{aligned}
\left[1 - \frac{2}{3}\right] E(R) =& \frac{2 + 2}{3} \\
E(R) =& 4
\end{aligned}$$

Exercise 27. You toss a fair coin until it lands on HT. What is the expected number of tosses?

Yay again!

$$\begin{aligned}
E(T) =& E(T|HT) + E(T|HH) + E(T|T) \\
=& 2\frac{1}{2^2} + [1 + E(T)]\frac{1}{2^2} + [1 + E(T)]\frac{1}{2}
\end{aligned}$$

Notice again that we're not using the law of total probability. We're thinking about when we need to restart each the process of trying again. Let's solve it now:

$$\begin{aligned}\left[1-\frac{1}{2^2}-\frac{1}{2}\right]E(T) =& \frac{2}{2^2}+\frac{1}{2^2}+\frac{1}{2}\\ \frac{4-1-2}{2^2}E(T) =& \frac{2+1+2}{2^2}\\ E(T) =& \frac{5}{1}=5\end{aligned}$$

Exercise 28. You toss a fair coin until it lands on the sequence THH shows up. What is the expected number of tosses? What is the answer for HTH, HHT, and HHH?

Yet again. The difference between each of these is when you have to entirely restart the process. Let's do them in reverse order.

$$\begin{aligned}E(T_{HHH}) =& E(T_{HHH}|HHH)P(HHH)+E(T_{HHH}|HHT)P(HHT)\\ &+E(T_{HHH}|HT)P(HT)+E(T_{HHH}|T)P(T)\\ =& \frac{3}{2^3}+[3+E(T_{HHH})]\frac{1}{2^3}+[2+E(T_{HHH})]\frac{1}{2^2}+[1+E(T_{HHH})]\frac{1}{2}\\ \left[1-\frac{1}{2^3}-\frac{1}{2^2}-\frac{1}{2}\right]E(T_{HHH}) =& \frac{3+3+2\cdot 2+1\cdot 2^2}{2^3}\\ \frac{8-1-2-4}{2^3}E(T_{HHH}) =& 6+4+42^3\\ E(T_{HHH}) =& \frac{14}{1}=14\end{aligned}$$

Now the next one:

$$\begin{aligned}E(T_{HHT}) =& E(T_{HHT}|HHT)P(HHT)+E(T_{HHT}|HHH)P(HHH)\\ &+E(T_{HHT}|HT)P(HT)+E(T_{HHT}|T)P(T)\\ =& \frac{3}{2^3}+[1+E(T_{HHT})]\frac{1}{2^3}+[2+E(T_{HHT})]\frac{1}{2^2}+[1+E(T_{HHT})]\frac{1}{2}\end{aligned}$$

That second term is $[1+E(T_{HHT})]$ because you only pay a one roll penalty for rolling HHH because you already have 2 heads of the 3 in sequence. Solving:

$$\begin{aligned}\left[1-\frac{1}{2^3}-\frac{1}{2^2}-\frac{1}{2}\right]E(T_{HHT}) =& \frac{3+1+2\cdot 2+1\cdot 2^2}{2^3}\\ \frac{8-1-2-4}{2^3}E(T_{HHT}) =& \frac{12}{2^3}\\ E(T_{HHT}) =& \frac{12}{1}=12\end{aligned}$$

Now the next one:

$$\begin{aligned}
E(T_{HTH}) =& E(T_{HTH}|HTH)P(HTH) + E(T_{HTH}|HTT)P(HTT) \\
&+ E(T_{HTH}|HH)P(HH) + E(T_{HTH}|T)P(T) \\
=& \frac{3}{2^3} + [3 + E(T_{HTH})]\frac{1}{2^3} + [1 + E(T_{HTH})]\frac{1}{2^2} + [1 + E(T_{HTH})]\frac{1}{2} \\
\left[1 - \frac{1}{2^3} - \frac{1}{2^2} - \frac{1}{2}\right] E(T_{HTH}) =& \frac{3 + 3 + 1 \cdot 2 + 2^2}{2^3} \\
\frac{E(T_{HTH})}{2^3} =& \frac{12}{2^3} \\
E(T_{HTH}) =& 12
\end{aligned}$$

And the original one:

$$\begin{aligned}
E(T_{THH}) =& E(T_{THH}|THH)P(THH) + E(T_{THH}|THT)P(THT) \\
&+ E(T_{THH}|TT)P(TT) + E(T_{THH}|H)P(H) \\
=& \frac{3}{2^3} + [2 + E(T_{THH})]\frac{1}{2^3} + [1 + E(T_{THH})]\frac{1}{2^2} + [1 + E(T_{THH})]\frac{1}{2} \\
\frac{E(T_{THH})}{2^3} =& \frac{3 + 2 + 2 + 2^2}{2^3} = \frac{12}{2^3} \\
E(T_{THH}) =& 12
\end{aligned}$$

Exercise 29. You toss a fair coin until the side the coin lands on changes. What is the expected number of tosses?

This is very similar to the previous question, only it's easier:

$$\begin{aligned}
E(T) =& E(T|HT)P(HT) + E(T|TH)P(TH) + E(T|HH)P(HH) + E(T|TT)P(TT) \\
=& 2\frac{2}{2^2} + 2\,[1 + E(T)]\frac{1}{2^2} \\
\left[1 - \frac{2}{2^2}\right] E(T) =& 1 + \frac{2}{2^2} \\
E(T) =& 2 + 1 = 3
\end{aligned}$$