

MATH 20 – HOMEWORK 8 SOLUTIONS!

due Wednesday, August 23

Instructions: This assignment is due at the *beginning* of class. Staple your work together (do not just fold over the corner). Please write the questions in the correct order. If I cannot read your handwriting, you won't receive full credit.

You may use Wolfram Alpha or another similar tool to compute any necessary sums or integrals, and for your matrix calculations. If you have trouble with this, let me know.

If you're using facts about distributions to answer the questions, be very clear about which distribution you're using to model that problem and why that distribution is appropriate.

1. Suppose a stranger approaches you with two envelopes containing money, with one envelope containing n times as much as the other, where n is any positive integer greater than one. He gives you one at random and allows you to look inside the envelope. He then asks if you would like to switch. Suppose also that you know the probability distribution $p_a = P(\text{smaller envelope contains } a \text{ dollars})$.

Find an inequality in terms of p_a and $p_{a/n}$ that determines when you expect to benefit from switching. (Recall that we proved in class that when $n = 2$, we switch when $\frac{p_a}{p_a + p_{a/2}} > \frac{1}{3}$.)

Solution: There are two pieces of randomness here: first, the smaller dollar amount is chosen according to the distribution p_a ; second, you are randomly handed one of the two envelopes with equal probability.

Suppose that you open the envelope and see that it contains $\$X$. You do not know yet whether $\$X$ is the larger of the two amounts (in which case the other envelope contains $\$X/n$) or the smaller of the two amounts (in which case the other envelope contains $\$nX$). However, you know you must be in one of these two cases.

As in class, let us rigorously define events. Let S be the event that you have the envelope with the smaller amount of money, and let T be the event that your envelope contains $\$X$.

The decision of whether to switch comes down to determining the probability that you have the envelope with the smaller amount of money. This means we'd like to find

$$P(S|T).$$

In other words, we want to find the probability that you have the smaller dollar amount, given the knowledge that you have $\$X$. (This is subtly different from "the probability that $\$X$ is the smaller amount". This distinction is the key to the paradox.)

By the definition of conditional probability,

$$P(S|T) = \frac{P(S \cap T)}{P(T)}.$$

The numerator is the probability that (a) the smaller amount of money is chosen to be $\$X$, and (b) you are handed that smaller envelope. These events are independent, so

$$P(S \cap T) = P(S)P(T) = p_X \cdot \frac{1}{2}.$$

Next, $P(T)$ is the probability that you get an envelope containing $\$X$. Remember the trick where we split an event into parts based on the other event:

$$P(T) = P(T \cap S) + P(T \cap \bar{S}) = \frac{1}{2}p_X + \frac{1}{2}p_{X/n}.$$

Therefore, just as we found in class, the probability that you have the smaller amount of money, given the knowledge of the dollar amount your envelope contains, is

$$\frac{p_X/2}{p_X/2 + p_{X/n}/2} = \frac{p_X}{p_X + p_{X/n}},$$

and we will henceforth denote the quantity by p_X^* .

Now we can calculate the expected value of the dollar amount in the second envelope, given your knowledge that your envelope contains $\$X$. With probability p_X^* , you have the smaller dollar amount, and so the other envelope contains $\$nX$. With probability $1 - p_X^*$, you have the larger dollar amount, and so the other envelope contains $\$X/n$. Therefore,

$$\mathbb{E}[\text{other envelope}] = p_X^*(nX) + (1 - p_X^*)(X/n).$$

You should switch when the expected value exceeds the amount in your own envelope. Therefore, switching is beneficial when

$$\begin{aligned} p_X^*(nX) + (1 - p_X^*)(X/n) &> X \\ np_X^* + \frac{1 - p_X^*}{n} &> 1 \\ np_X^* + \frac{1}{n} - \frac{p_X^*}{n} &> 1 \\ p_X^* \left(n - \frac{1}{n} \right) &> 1 - \frac{1}{n} \\ p_X^* &> \frac{1 - \frac{1}{n}}{n - \frac{1}{n}} \\ p_X^* &> \frac{\frac{n-1}{n}}{\frac{n^2-1}{n}} \\ p_X^* &> \frac{n-1}{n^2-1} \\ p_X^* &> \frac{1}{n+1}. \end{aligned}$$

Therefore, switching is advantageous when

$$\boxed{\frac{p_X}{p_X + p_{X/n}} > \frac{1}{n+1}}.$$

Note that when $n = 2$, this matches the theorem proved in class.

2. Suppose that 10,000 random digits (between 0 and 9) are chosen uniformly at random. Find (a) exactly using a binomial distribution, and (b) approximately using the Central Limit Theorem and a normal distribution table, the probability that the digit 3 appears not more than 931 times. If you have trouble evaluating the sum in part (a) in Wolfram Alpha, tell me.

Solution to (a): This is easy to do exactly with a binomial distribution. Each digit is a Bernoulli trial with probability of success $1/10$. Let S_n be the number of 3s in n trials. Then,

$$P(S_{10000} \leq 931) = \sum_{k=0}^{931} \binom{10000}{k} \left(\frac{1}{10}\right)^k \left(\frac{9}{10}\right)^{10000-k}.$$

The Wolfram Alpha command

“sum from $k=0$ to $k=931$ of $(10000 \text{ choose } k) \cdot (1/10)^k \cdot (9/10)^{(10000-k)}$ ”.

gives the result

$$P(S_{10000} \leq 931) = \boxed{0.010647\dots}$$

Solution to (b): We will use the Central Limit Theorem to approximate the re-scaled binomial distribution. Our direct formula for this is:

$$b(n, p, j) \approx \frac{1}{\sqrt{np(1-p)}} \phi\left(\frac{j-np}{\sqrt{np(1-p)}}\right),$$

with ranges for j being approximated by integrals. We apply continuity correction to the upper bound, using 931.5 rather than 931. This leads to the estimate

$$P(S_{10000} \leq 931) \approx \int_{-\infty}^{\frac{931.5-np}{\sqrt{np(1-p)}}} \phi(x) dx = \int_{-\infty}^{-137/60} \phi(x) dx = \int_{-\infty}^{-2.28333} \phi(x) dx = \boxed{0.011205\dots}$$

This is an example of the case where the continuity correction actually gives a slightly worse estimate. If we had computed without the continuity correction, we'd get an estimate of

$$P(S_{10000} \leq 931) \approx 0.01072\dots$$

Nonetheless, the default is to always use continuity correct unless otherwise stated. Note that using 930.5 is a continuity correction in the *wrong direction*, and is considered an incorrect answer.

3. A random walker starts at 0 on the x -axis and at each time unit moves 1 step to the right or 1 step to the left with probability 1/2. Estimate the probability that, after 100 steps, the walker is more than 10 steps from the starting position. (Use the Central Limit Theorem and a normal distribution table.)

Solution: Let X be the random variable representing a single step, so that X takes on values 1 and -1 each with probability 1/2. One can then compute that $\mathbb{E}[X] = 0$ and $\text{Var}(X) = 1$. Let S_{100} be the sum of 100 trials of X . Then $\mathbb{E}[S_{100}] = 0$, $\text{Var}(S_{100}) = 100$, and the standard deviation of S_{100} is 10. The Central Limit Theorem allows us to estimate

$$P(-10 \leq S_{100} \leq 10)$$

and then the quantity that we want to find is

$$P(|S_{100}| > 10) = 1 - P(-10 \leq S_{100} \leq 10).$$

Because S_n is a discrete random variable, we apply a continuity correct. The estimate that we get is

$$\begin{aligned} P(-10.5 \leq S_{100} \leq 10.5) &= P\left(\frac{-10.5 - 0}{10} \leq \frac{S_{100}}{10} \leq \frac{10.5 - 0}{10}\right) \\ &\approx \int_{-1.05}^{1.05} \phi(x) dx \\ &\approx 0.7063. \end{aligned}$$

This gives

$$P(|S_{100}| > 10) \approx 1 - 0.7063 = \boxed{0.2937}.$$

Note: The true exact answer is about 0.2712. This can be found by reimagining the scenario to be a kind of binomial distribution. Do you see how?

Note 2: The estimate, even with the continuity correction, is off by a little more than we're used to with an n this large. This is essentially because the distance of the walker from 0 after 100 steps must be an *even number*. You cannot be 3 away after 100 steps, for example. So, when measuring the probability that you end K steps from the origin, you get 0 for all odd numbers. In cases like this, the estimate can be improved by using $S_n/2$ instead of S_n . Without the continuity correction, you get the same answer (check this!), but now the continuity correction is correcting twice as much. (It would be like adding/subtracting 1 instead of 1/2 in the original problem.) This leads to a much better estimate of 0.2713.

4. A surveyor is measuring the height of a cliff known to be about 1000 feet. He assumes his instrument is properly calibrated and that his measurement errors are independent, with mean $\mu = 0$ and variance $\sigma^2 = 10$. He plans to take n measurements and form the average. Estimate, using (a) Chebyshev's inequality and (b) the normal approximation, how large n should be if he wants to be 95 percent sure that his average falls within 1 foot of the true value. Now estimate, using (c) Chebyshev's inequality and (d) the normal approximation, how small σ^2 must be if he wants to take only 10 measurements with the same resulting confidence.

Solution: Let X_i be the random variable for the i th measurement and define

$$S_n = X_1 + X_2 + \cdots + X_n.$$

Define $A_n = S_n/n$, so that A_n represents the average of n measurements. By our usual formulas, we see that $\mathbb{E}[A_n] = 1000$ and $\text{Var}(A_n) = 10/n$.

(a): Chebyshev's Inequality says that

$$P(|A_n - 1000| \geq 1) \leq \frac{10}{n}.$$

We want this probability to be less than $1 - 0.95 = 0.05$. This holds when

$$\frac{10}{n} \leq \frac{1}{20} \implies n \geq 200.$$

Therefore, Chebyshev's Theorem tells us that 200 measurements are enough to be 95% certain that the measurement is correct within 1 foot.

(b): The Central Limit Theorem allows us to estimate

$$P(999 \leq A_n \leq 1001).$$

We find, after rescaling, that

$$\begin{aligned} P(999 \leq A_n \leq 1001) &= P\left(\frac{999 - 1000}{\sqrt{10/n}} \leq \frac{A_n - 1000}{\sqrt{10/n}} \leq \frac{1001 - 1000}{\sqrt{10/n}}\right) \\ &= P\left(-\frac{\sqrt{n}}{\sqrt{10}} \leq \frac{A_n - 1000}{\sqrt{10/n}} \leq \frac{\sqrt{n}}{\sqrt{10}}\right). \end{aligned}$$

We *want* this quantity to be at least 95%. We can use our 68-95-99.7 rule, which tells us here that we want $\sqrt{n/10}$ to be at least 2. (We can use a z-table to be more precise. It tells us that we really only need it to be at least 1.96. I'll use the 2 number here for simplicity.)

Hence,

$$\frac{\sqrt{n}}{\sqrt{10}} \geq 2 \implies n \geq 40.$$

(Using 1.96 instead of 2 would lead to $n \geq 38.4$. Since the number of measurements must be an integer, we'd conclude $n \geq 39$.)

NOTE: It is incorrect to use a continuity correction here. No correction is needed because the distribution we're approximating is itself continuous.

(c): This is the reverse question from part (a), and we can reuse most of our work. Now, n is fixed at 10 and we must find the largest variance that allows the confidence we seek. Set $\text{Var}(X) = \sigma^2$ and note that $\text{Var}(A_{10}) = \sigma^2/10$. Now,

$$P(|A_{10} - 1000| \geq 1) \leq \frac{\sigma^2}{10}.$$

We want this probability to be less than $1 - 0.95 = 0.05$. This holds when

$$\frac{\sigma^2}{10} \leq \frac{1}{20} \implies \sigma^2 \leq 0.5.$$

(This is a big jump! We'd have to reduce the variance of our measuring tool from 10 down to 0.5, a factor of 20, to get the required confidence in 10 measurements... at least according to Chebyshev's Theorem.)

(d): Reusing most of our work from (b) yields the inequality

$$\frac{\sqrt{10}}{\sigma} \geq 2 \implies \sigma^2 = 2.5.$$

You might fairly complain that $n = 10$ is not very large, and we shouldn't trust Central Limit Theorem estimates for such small n . I agree!

5. Assume that a student going to a certain four-year medical school in northern New England has, each year, a probability q of flunking out, a probability r of having to repeat the year, and a probability p of moving on to the next year (in the fourth year, this means graduating).

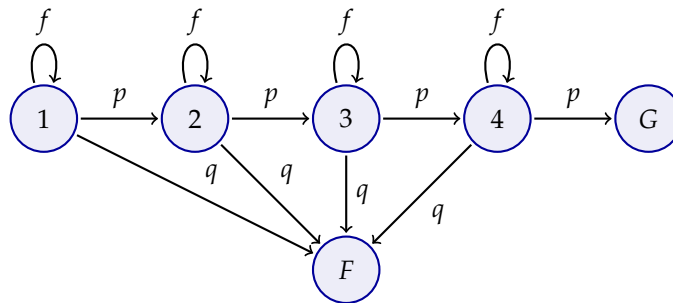
(a) Form a transition matrix for this process taking as states $F, 1, 2, 3, 4, G$, where F stands for flunking out, G stands for graduating, and the other states represent the year of study.

(b) For the case $q = 0.1, r = 0.2, p = 0.7$, find the amount of time a new student can expect to be in medical school?

(c) Find the probability that this beginning student will graduate.

Solution to (a):

This situation can be modeled with the Markov chain below.



The transition matrix is therefore

$$\begin{matrix} & 1 & 2 & 3 & 4 & F & G \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ F \\ G \end{matrix} & \begin{pmatrix} r & p & 0 & 0 & q & 0 \\ 0 & r & p & 0 & q & 0 \\ 0 & 0 & r & p & q & 0 \\ 0 & 0 & 0 & r & q & p \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

Solution to (b): The question essentially asks "If we start in State 1, what is the expected time until an absorbing state is reached." According to Theorem 11.5, the answer is found in the first entry of $N\vec{c}$, where $N = (I - Q)^{-1}$ for the identity matrix I and the upper-left 4×4 block of the transition Q , and \vec{c} is a column vector of all 1s.

First we calculate that

$$I - Q = \begin{pmatrix} 1-r & -p & 0 & 0 \\ 0 & 1-r & -p & 0 \\ 0 & 0 & 1-r & -p \\ 0 & 0 & 0 & 1-r \end{pmatrix}$$

Using Wolfram Alpha, we find the inverse:

$$(I - Q)^{-1} = \begin{pmatrix} \frac{1}{1-r} & \frac{p}{(1-r)^2} & \frac{p^2}{(1-r)^3} & \frac{p^3}{(1-r)^4} \\ 0 & \frac{1}{1-r} & \frac{p}{(1-r)^2} & \frac{p^2}{(1-r)^3} \\ 0 & 0 & \frac{1}{1-r} & \frac{p}{(1-r)^2} \\ 0 & 0 & 0 & \frac{1}{1-r} \end{pmatrix}$$

Thus, the first entry of $(I - Q)^{-1}\vec{c}$ is

$$\frac{1}{1-r} + \frac{p}{(1-r)^2} + \frac{p^2}{(1-r)^3} + \frac{p^3}{(1-r)^4}.$$

Upon substituting $r = 1/5$ and $p = 7/10$, we find that the average length of time until a student graduates or flunks out is $\boxed{8475/2048 = 4.13818359375}$ years.

Solution: In this question we need to find the probability that if starting in State 1, a path in the Markov chain gets absorbed into state G , not state F .

By Theorem 11.6, the answer we seek will be B_{12} where B is the matrix

$$B = NR.$$

The matrix N is the same one above, and the matrix R is the upper right 4×2 block of Q . Thus,

$$B = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{pmatrix} \frac{1}{1-r} & \frac{p}{(1-r)^2} & \frac{p^2}{(1-r)^3} & \frac{p^3}{(1-r)^4} \\ 0 & \frac{1}{1-r} & \frac{p}{(1-r)^2} & \frac{p^2}{(1-r)^3} \\ 0 & 0 & \frac{1}{1-r} & \frac{p}{(1-r)^2} \\ 0 & 0 & 0 & \frac{1}{1-r} \end{pmatrix} \begin{matrix} F & G \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{pmatrix} q & 0 \\ q & 0 \\ q & 0 \\ q & p \end{pmatrix} = \begin{matrix} F & G \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{pmatrix} - & \frac{p^4}{(1-r)^4} \\ - & - \\ - & - \\ - & - \end{pmatrix},$$

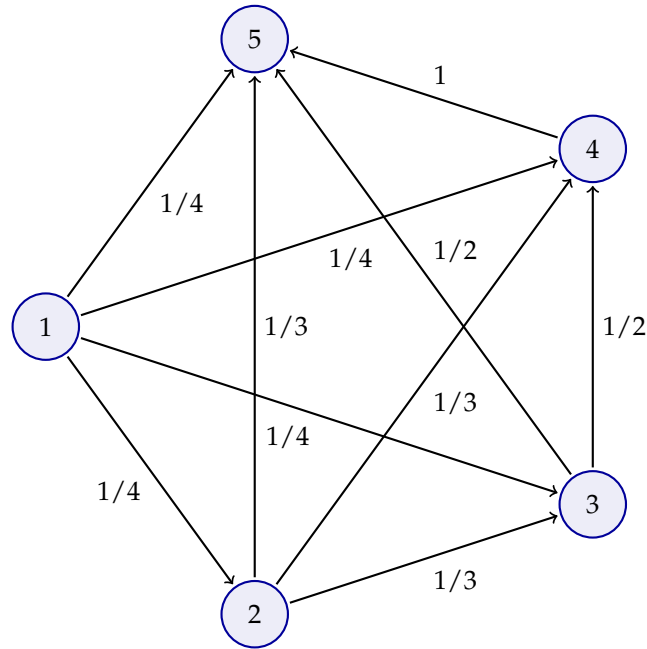
where I've omitted the unnecessary entries of the rightmost matrix to save space. Substituting $r = 1/5$ and $p = 7/10$ gives a probability of

$$\boxed{\frac{2401}{4096} = 58.6181640625\%}$$

of graduating.

6. **Bonus: 5 points. (Not required.)** A process moves on the integers 1, 2, 3, 4, 5. It starts at 1 and, on each successive step, moves to an integer greater than its present position, moving with equal probability to each of the remaining larger integers. State five is an absorbing state. Find the expected number of steps to reach state five. Explain how this question is related to the "Stubborn Candles" experiment in Lab 1, and how you could now find *exactly* the expected number of attempts to blow out 35 candles. (You don't need to actually perform this calculation.)

Solution: We follow the same basic steps as in Question 5. The Markov chain that models this situation is below.



The transition matrix is

$$\begin{array}{c}
 \begin{matrix} & 1 & 2 & 3 & 4 & 5 \\
 \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}
 \end{matrix}
 \end{array}$$

Following the same steps as in the previous question leads to the answer

$$\boxed{\frac{25}{12} = 2.08333\dots}$$

This is clearly analogous to the case of blowing out 4 candles, which would be like traversing the integers 4, 3, 2, 1, 0 in the same way that we're traversing 1, 2, 3, 4, 5 here. So, this method could be used to determine the expected number of attempts to blow out 35 candles. Doing this leads to the exact answer

$$\frac{54437269998109}{13127595717600} = 4.1467814\dots$$

7. **Bonus: 5 points. (Not required.)** A fair coin is flipped 400 times. Determine the number x such that the probability that the number of heads is between $200 - x$ and $200 + x$ is approximately 0.8. (Use the Central Limit Theorem and a normal distribution table.)

Solution: Let S_n be the number of successes after flipping the coin 400 times. Then, $\mathbb{E}[S_n] = 200$ and $\text{Var}(S_n) = 100$. By the Central Limit Theorem for Binomial Trials, (and using the continuity correction)

$$P(200 - x - 0.5 \leq S_n \leq 200 + x + 0.5) = P\left(-\frac{x - 0.5}{10} \leq \frac{S_n - 200}{10} \leq \frac{x + 0.5}{10}\right) \approx \int_{-(x-0.5)/10}^{(x+0.5)/10} \phi(z) dz.$$

(I've used the letter z in the integral so as to not confuse with x , which has already been claimed.)

Next we look at a z -table to find the number of standard deviations a for which 80% of the normal pdf lies between $-a$ and a . We find this to be approximately 1.28.

Therefore, we want

$$\frac{x + 0.5}{10} = 1.28$$

which implies

$$x = 12.3.$$

To round to the closest integer, we use $\boxed{x = 12}$. Note that if we had forgotten the continuity correction, we would have found $x = 12.8$ and rounded to $x = 13$. Which is more accurate?

Since this is really a binomial distribution, we can just compute the exact values. We find that

$$P(200 - 12 \leq S_n \leq 200 + 12) = 0.7887\dots,$$

while

$$P(200 - 13 \leq S_n \leq 200 + 13) = 0.8230\dots$$

Hence, the continuity correction brought us closest to the correct value.
