

Math/Stats 342: Solutions to Homework

Steven Miller (sjm1@williams.edu)

November 17, 2011

Abstract

Below are solutions / sketches of solutions to the homework problems from Math/Stats 342: Probability (Mount Holyoke College, Fall 2011, Professor Steven J. Miller, sjm1@williams.edu). The course homepage is

http://www.williams.edu/Mathematics/sjmiller/public_html/342.

Note to students: it's nice to include the statement of the problems, but I leave that up to you. **I am only skimming the solutions. I will occasionally add some comments or mention alternate solutions. If you find an error in these notes, let me know for extra credit.**

Contents

1	HW #1: Due Thursday, September 15, 2011	2
2	HW #2: Due Tuesday, September 20	4
3	HW #3: Due Tuesday, September 27	7
4	HW #4: Due Tuesday, October 4	9
5	HW #5	12
6	HW #6	14
7	HW #7: Due Thursday, October 27	16
8	HW #8: Due Tuesday, November 15	18

n	Prob(not a match)	Prob(no matches)
1	1.000	1.000
2	0.990	0.990
3	0.980	0.970
4	0.970	0.941
5	0.960	0.903
6	0.950	0.858
7	0.940	0.807
8	0.930	0.750
9	0.920	0.690
10	0.910	0.628
11	0.900	0.565
12	0.890	0.503
13	0.880	0.443

Table 1: Table of probabilities that there are not two people sharing a birthday in a 100 day year. The first column is n , the second is the probability that, given no matches among the first $n - 1$, the probability that n is not a match. The third column is the probability that there is no match among the first n .

1 HW #1: Due Thursday, September 15, 2011

Problem: (1) If you have 100 days in the year, how many people do you need in a room for there to be at least a 50% chance that two share a birthday?

Solution: Our formula from class approximates the number of people needed in a year with N days to be $1/2 + \sqrt{N^2 \log 2}$. Taking $N = 100$ gives approximately 12.27. Let's find the answer by brute force (the actual answer, not an approximation). We want the first n so that

$$\frac{100}{100} \cdot \frac{99}{100} \cdot \frac{98}{100} \cdots \frac{100 - (n - 1)}{100} \leq .5,$$

as this product is the probability we do not have a birthday in common with n people in a year with $N = 100$ birthdays. See Table 1.

Problem: (2) Generalize the argument from class of 20 blocks of 5 to blocks of size 6 (I can do 10 of size 10 without too much work). Note we still have 100 spins, and each spin has a 50% chance of being red, 50% chance of being black.

Solution: Let a_n be the probability we don't have 5 consecutive blacks in n spins, and b_n the probability that we do have 5 consecutive spins are black in n spins. We break the 100 spins into 16 blocks

of length 6 and one of length 4. Clearly the block of length 4 never has 5 consecutive blacks. For the 16 blocks of length 6, each block has $2^6 = 64$ possible strings of red and black. The only ones that have 5 consecutive blacks (or more) are BBBBRR, RBBBBB, and BBBBBB. Thus 61/64 do not have 5 consecutive blacks.

If none of the 16 blocks of length 6 have 5 consecutive blacks, by multiplying the probabilities we get an upper bound for a_n (it is an upper bound as we could have 5 in a row split between two blocks):

$$a_n \leq \left(\frac{61}{64}\right)^{16} \cdot 1,$$

so

$$b_n \geq 1 - \left(\frac{61}{64}\right)^{16} \approx .536.$$

2 HW #2: Due Tuesday, September 20

Problem: Page 9: #3: Sampling with replacement: Box contains tickets marked $1, 2, \dots, n$. Ticket drawn from random and then replaced and another drawn.

Solution: (a) Probability first ticket is a 1 and second is a 2 is just $1/n^2$. This is $1/n$ times $1/n$, as the outcome of the first event has no influence on the probability of the second.

(b) The numbers are two consecutive integers, with first one less than second. The answer is $(n-1)/n^2$. If the first ticket is an n can't happen; the probability the first is not n is $(n-1)/n$, and then for any number other than n , exactly one number works for the second.

(c) The second number is bigger than the first. Lots of ways to do this. If our first number is k (which happens with probability $1/n$), the probability the second number is larger is $(n-k)/n$. Thus our answer is the sum

$$\frac{1}{n} \frac{n-1}{n} + \frac{1}{n} \frac{n-2}{n} + \dots + \frac{1}{n} \frac{n-n}{n} = \frac{0+1+2+\dots+n-1}{n^2}.$$

We showed that $1+2+\dots+m = m(m+1)/2$ in class, and thus the probability is $\frac{1}{n^2} \frac{(n-1)n}{2} = \frac{n-1}{2n} = \frac{1}{2}(1 - \frac{1}{n})$. Another way to see this is let x be the probability the first number is larger, y the probability the second number is larger (note $x = y$ by symmetry) and z the probability they are the same. We can easily compute z , which is just $1/n$. As $x + y + z = 1$, we find $2y + \frac{1}{n} = 1$ or $y = \frac{1}{2}(1 - \frac{1}{n})$.

(d) Repeat (a) through (c) but now assume there is no replacement. (a) becomes $\frac{1}{n} \frac{1}{n-1}$, as now when we draw the second ticket there are only $n-1$ possibilities. Similarly (b) becomes $\frac{n-1}{n} \frac{1}{n-1}$, as again the second ticket now has $n-1$ options and not n options. For (c), by symmetry the probability the second is larger is the same as the probability the first is larger. The probability they are the same is now zero, and thus the probability the second is larger is just $1/2$.

Problem: Page 9: #5: Suppose a deck of 52 cards is shuffled and dealt.

Solution: (a) How many ordered pairs of two cards? There are $52 \cdot 51$ ordered pairs.

(b) What is the chance the first card is an ace? It's just $4/52 = 1/13$ (as there are four aces).

(c) What is the chance the second card is an ace? Two possibilities: first is an ace and the second is an ace, or first is not an ace and second is an ace. The probabilities of this are

$$\frac{4}{52} \frac{3}{51} + \frac{48}{52} \frac{4}{51} = \frac{12 + 192}{52 \cdot 51} = \frac{204}{52 \cdot 51} = \frac{1}{13}.$$

Very interesting that it comes out $1/13$. Why is this reasonable? If you see the two cards on the table, no way to know which one was picked first, and thus the $1/13$ should be expected.

(d) What is the chance both are aces? It's $\frac{4}{52} \frac{3}{51}$.

(e) What is the chance of at least one ace? We just did the chance of two aces; we could calculate the chance of exactly one ace and add. That's $\frac{4}{52} \frac{48}{51} + \frac{48}{52} \frac{4}{51}$. Alternatively, it's 1 minus the probability of no aces. The probability of no aces is $\frac{48}{52} \frac{47}{51}$, so the answer is $1 - \frac{48}{52} \frac{47}{51}$.

Problem: Page 10: #7: Roll two dice.

Solution: (a) What is the probability the maximum of the two numbers is less than or equal to 2? This means each die is at most 2. Each die is at most 2 with probability $2/6$, so the probability both rolls are at most 2 is $(2/6)^2 = 1/9$.

(b) What is the probability the maximum is at most 3? It's now $(3/6)^2 = 1/4$.

(c) The maximum of the two numbers is exactly a 3? Either first roll is a 3 and next is a 1 or 2, or the first is a 1 or 2 and the second is a 3, or both rolls are 3. We add the probabilities of each of these three possibilities, and get

$$\frac{1}{6} \frac{2}{6} + \frac{2}{6} \frac{1}{6} + \frac{1}{6} \frac{1}{6} = \frac{5}{36}.$$

Note what we're really doing is finding $\text{Prob}(A \cup B)$, where A is the first roll is a 3 and B is the probability the second roll is a 3. This probability is $\text{Prob}(A) + \text{Prob}(B) - \text{Prob}(A \cap B)$.

(d) Repeat (b) and (c) replacing 3 with x for all x from 1 to 6. For (b), it's just $(x/6)^2$. For (c), it's

$$\frac{1}{6} \frac{x-1}{6} + \frac{x-1}{6} \frac{1}{6} + \frac{1}{6} \frac{1}{6} = \frac{2x-1}{36}.$$

Summing this for $x = 1$ to 6 yields $36/36$, or 1. This is reasonable, as one of the numbers from 1 to 6 must be largest!

Problem: Page 30: #3: 500 tickets marked 1 thru 500 sold in a raffle. I have tickets 17, 93, 202 and friend has 4, 101, 102, 398. One ticket chosen at random. Make an outcome space for the situation, indicate how each of the events can be represented as a subset of outcome space.

Solution: (a) One of my tickets is the winner. The outcome space is the numbers 1 to 500, and this event is that one of tickets 17, 93 or 202 is drawn (for fun, note this has probability $3/500$).

(b) Neither I nor friend wins: The outcome space is as before, and now the event is all numbers from 1 to 500 except 4, 17, 93, 101, 102, 202, 398. For fun, note the probability of this is $(500 - 7)/500$.

(c) The winning number is one away from a number on my ticket. Outcome space as before, now event is the numbers 16, 18, 92, 94, 201, 203, and for fun the probability of this happening is $6/500$.

Problem: Page 30: #10: Find expression of events in terms of $\Pr((\) A)$, $\Pr((\) B)$, $\Pr((\) C)$, $\Pr((\) A \cap B)$, et cetera.

Solution: (a) The probability that exactly two of A, B, C occur. Need either A, B and not C , or A, C and not B , or B, C and not A . The probabilities of these are $\Pr((\) A \cap B \cap C^c)$, $\Pr((\) A \cap B^c \cap C)$, $\Pr((\) A^c \cap B \cap C)$. We need to break these probabilities into the nice intersections that are assumed known; sadly we aren't allowed to use complements directly. That said, we have a nice observation: $\Pr((\) A \cap B \cap C^c) + \Pr((\) A \cap B \cap C) = \Pr((\) A \cap B)$. Why? The first probability is the probability that A and B happen and C doesn't, while the second is the probability A and B happen and C does; their sum is thus the probability that A and B happen, irregardless of C . We thus find

$$\Pr((\) A \cap B \cap C^c) = \Pr((\) A \cap B) - \Pr((\) A \cap B \cap C).$$

Substituting, the probability that exactly two happens is just

$$\Pr((\) A \cap B) + \Pr((\) A \cap C) + \Pr((\) B \cap C) - 3\Pr((\) A \cap B \cap C).$$

(b) The probability that exactly one of A, B, C occurs. Well, the probability that all three occur is $\Pr((A \cap B \cap C))$. We calculated above the probability that exactly two occur. The probability that none occur is done in part (c). Our answer is just 1 minus these three probabilities. There must be a better way! Let's look at the inclusion / exclusion argument. Consider

$$\Pr((A)) + \Pr((B)) + \Pr((C)) - 2\Pr((A \cap B)) - 2\Pr((A \cap C)) - 2\Pr((B \cap C)) + 3\Pr((A \cap B \cap C)).$$

Where did this come from? If an x occurs in exactly one of A, B, C then it is counted just once above. If it occurs in exactly two of them, it is counted twice from $\Pr((A)) + \Pr((B)) + \Pr((C))$ and then it is counted -2 times from the intersection of two elements terms (only one of the three intersections of pairs is non-empty). If it occurs in all three of A, B, C then it is also in all three intersection of pairs, as well as the triple intersection. The numbers have been chosen so that this adds up to 0. Thus the only net contribution is from items occurring in exactly one of the three events A, B, C .

(c) The probability that none of A, B, C occurs. The probability at least one occurs is $\Pr((A \cup B \cup C))$, so our probability is $1 - \Pr((A \cup B \cup C))$. We expand the union using inclusion / exclusion, and find

$$1 - \Pr((A \cup B \cup C)) = 1 - [\Pr((A)) + \Pr((B)) + \Pr((C)) - \Pr((A \cap B)) - \Pr((A \cap C)) - \Pr((B \cap C)) + \Pr((A \cap B \cap C))].$$

Problem: Prove if A is a subset of B then $\Pr((A))$ is at most $\Pr((B))$.

Solution: We have $B = A \cup (A^c \cap B)$. Thus

$$\Pr((B)) = \Pr((A)) + \Pr((A^c \cap B)) \geq \Pr((A)),$$

as $\Pr((A^c \cap B)) \geq 0$.

3 HW #3: Due Tuesday, September 27

Problem: Section 1.4: #1. 92% women right handed, 88% men right handed. Determine if statements true, false or need more info.

Solution: (a) overall population right handed is 90%. Need more info. What if 100,000 and only 100 men? Or vice-versa? Only true if same number men and women.

(b) Overall proportion of right handed is between 88% and 92%. True: lowest when all men, highest when all women.

(c) If sex ratio is 1-1, then (a) is true. Correct. If P people, have fraction right handed is $(.92P + .88P)/P = .90$.

(d) If (a) is true then sex ratio is 1-1. True. If not, then by (c) wouldn't have 90% right handed.

(e) If at least 3 times as many women as men, then ratio of right handed is at least 91%. True. If have w women and m men, with $w \geq 3m$, then ratio is $(.92w + .88m)/(w + m) = .88 + .04 \frac{w}{w+m}$. If we consider $w = xm$, then the ratio $\frac{w}{w+m}$ equals $\frac{x}{x+1}$, with $x \geq 3$. Using calculus, we see that as x increases this ratio increases, and thus the *lowest* value is when $x = 3$, which gives $3/4$. Substituting this above we find the percentage is at least $.88 + .04 \cdot \frac{3}{4} = .91$.

Problem: Section 1.4: #3. Suppose $\Pr(\text{() rain today}) = .4$, $\Pr(\text{() rain tomorrow}) = .5$ and $\Pr(\text{() rain today and tomorrow}) = .3$. If it rains today what is the chance it rains tomorrow?

Solution: Let A be the event it rains today, B the event that it rains tomorrow. We're told $\Pr(\text{() } A) = .4$, $\Pr(\text{() } B) = .5$ and $\Pr(\text{() } A \cap B) = .3$. We are asked to find $\Pr(\text{() } A|B)$, which is $\Pr(\text{() } A \cap B)/\Pr(\text{() } B)$, or $.3/.5 = .6$ or 60%.

Problem: Section 1.4: #9. Three high schools have senior classes of size 100, 400, 500. Scheme A: make a list of all 1000 students and choose one randomly; Scheme B: pick a school at random then a student at random; Scheme C: pick school i with probability p_i such that $p_1 + p_2 + p_3 = 1$, and then pick a student at random in that school. Find choices for p_i so that Scheme C and A are equivalent.

Solution: (a) Show Schemes A and B are not probabilistically equivalent. This is similar to the problem we did in class with 50 white and 50 black marbles – if we had one jar with 50 white and 49 black and the other just 1 black, we have almost a 75% chance of picking a black marble. Same here. In Scheme A each person in first school is chosen with probability 1/1000; in Scheme B choose that school 1/3 of the time, and then each person is chosen 1/100 of the time, so a person in the first school is now chosen 1/300 of the time.

(b) Each person needs to be chosen 1/1000 of the time. Thus

$$p_1 \frac{1}{100} = p_2 \frac{1}{400} = p_3 \frac{1}{500}, \quad p_1 + p_2 + p_3 = 1.$$

We find $p_1 = \frac{p_2}{4} = \frac{p_3}{5}$, so $p_2 = 4p_1$, $p_3 = 5p_1$. Substituting gives $10p_1 = 1$ so $p_1 = 1/10$, and thence $p_2 = 4/10$ and $p_3 = 5/10$. Note p_i is the size of the high school divided by the sum of the sizes of the high schools.

Problem: Section 1.6: #5. You are one of n students in class.

Solution: (a) What is the chance another student has the same birthday as yours? Assuming all days equally likely and none on February 29, probability is $1/365$.

(b) How many students so that have at least a 50% chance someone shares your birthday? Probability n people don't share your birthday is $(364/365)^n$. If this equals $1/2$, then $n = \log(1/2)/\log(364/365) \approx 252.652$, or 253 days.

(c) Why is this different than the birthday problem? Difference is that it is not enough to have a pair involved in the same birthday, but one of the pair must be the 'special, pre-assigned birthday' (yours).

Problem: Section 1.6: #6. Roll a fair 6-sided die until roll a number previously rolled.

Solution: (a) For $r = 1, 2, \dots$ calculate the probability you rolled exactly r times. Probability is clearly zero whenever $r \geq 8$. This is the famous Pidgeon-hole principle; we must have our first repeat by the 7th roll. Similarly probability is zero when $r = 1$ as no chance for a repeat. When $r = 2$, probability is $1/6$ (first roll can be anything, second roll must equal first). When $r = 3$, first roll can be anything, second must be different, third must be the same as the first or second: $1 \cdot 5/6 \cdot 2/6 = 10/36$.

If $r = 4$ have $1 \cdot 5/6 \cdot 4/6 \cdot 3/6 = 60/6^3$.

For $r = 5$ have $1 \cdot 5/6 \cdot 4/6 \cdot 3/6 \cdot 4/6 = 240/6^4$.

For $r = 6$ have $1 \cdot 5/6 \cdot 4/6 \cdot 3/6 \cdot 2/6 \cdot 5/6 = 600/6^5$.

For $r = 7$, last roll must agree with one of first 6 rolls: $1 \cdot 5/6 \cdot 4/6 \cdot 3/6 \cdot 2/6 \cdot 1/6 \cdot 1 = 120/6^5$.

(b) What is $p_1 + \dots + p_{10}$? The sum must be 1; one of these events must occur.

(c) Calculated sum equals theoretical predictions.

4 HW #4: Due Tuesday, October 4

Problem: Section 2.1. Page 91: #2 (not the relative frequency part). In 4 child families, each child equally likely to be a boy or girl. Which would be more common among 4 child families: 2 boys and 2 girls, or different numbers of boys and girls?

Solution: There are $\binom{4}{2} = 6$ ways to have 2 boys and 2 girls; as there are $2^4 = 16$ ways to have four children, there are 16-6 or 10 ways to have a different number of boys and girls. Thus, it is more common not to have equal numbers of boys and girls with 4 children.

Problem: Page 91. #7. You and I roll a die, if I get a strictly larger number I win. If we play five times, what are the chances I win at least 4?

Solution: First figure out the probability I win in one attempt. If you roll a k then I have $6 - k$ numbers that I could roll that are larger (this makes sense; note I can't win if you roll a 6). As you roll k with probability $1/6$, the probability I win is

$$\sum_{k=1}^6 \text{Prob}(\text{first roll is } k) \frac{6-k}{6} = \frac{1}{6} \frac{5}{6} + \frac{1}{6} \frac{4}{6} + \frac{1}{6} \frac{3}{6} + \frac{1}{6} \frac{2}{6} + \frac{1}{6} \frac{1}{6} + \frac{1}{6} \frac{0}{6} = \frac{15}{36}.$$

There's another way to see this. There are 36 ways to choose a pair of numbers. Six of these will have the two numbers the same, the other 30 will have the two different. Of those 30, half the time I'll have the larger, and half the time you will.

We now turn to the question. It's a Binomial process with $n = 5$ and $p = 15/36$. Thus the probability I win at least 4 times is

$$\binom{5}{4} (15/36)^4 (21/36) + \binom{5}{5} (15/36)^5 = \frac{3125}{31104} \approx 10.0\%.$$

Problem: Page 91. #10. Toss a fair coin n times.

Solution: (a) What is probability have $k - 1$ heads given have $k - 1$ or k heads? Probability have $k - 1$ heads is $\binom{n}{k-1} (1/2)^{k-1} (1/2)^{n-(k-1)} = \binom{n}{k-1} (1/2)^n$, while the probability we have k heads is $\binom{n}{k} (1/2)^k (1/2)^{n-k} = \binom{n}{k} (1/2)^n$. As $\text{Prob}(A|B) = \text{Prob}(A \cap B) / \text{Prob}(B)$, we can find the desired answer. Let A be the event that we have $k - 1$ heads, B the event that we have $k - 1$ or k heads. Note $A \cap B = A$. Thus our answer is

$$\frac{\text{Prob}(A)}{\text{Prob}(B)} = \frac{\binom{n}{k-1} (1/2)^n}{\binom{n}{k-1} (1/2)^n + \binom{n}{k} (1/2)^n} = \frac{\binom{n}{k-1}}{\binom{n}{k-1} + \binom{n}{k}} = \frac{\binom{n}{k-1}}{\binom{n+1}{k}}.$$

We now do some algebra, using the expansion for the binomial coefficients:

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

(b) What is the probability that we have k heads given that we have $k - 1$ or k heads? The answer should just be 1 minus the probability above, but to check our work let's do it the long way. The argument is essentially the same as before; the only difference is that rather than have $\binom{n}{k-1}$ in the numerator at the end we have $\binom{n}{k}$, which leads to

$$\frac{\binom{n}{k}}{\binom{n+1}{k}} = \frac{n!/k!(n-k)!}{(n+1)!/k!(n-k+1)!} = \frac{n-k+1}{n+1};$$

note the sum of these probabilities is indeed 1.

Problem: Section 3.1. Page 158: #3. Roll a fair die twice. S is the sum of the rolls.

Solution: (a) What is the range of S ? S takes on the values $\{2, 3, \dots, 12\}$.

(b) Find the distribution of S . There are 36 rolls. Six of these have a sum of 7, one has a sum of 2, one has a sum of 12, and in fact the number that have a sum of $k \leq 6$ is k (and so $\text{Prob}(S = k) = k/36$), while the number that have a sum of $k \geq 6$ is $12 - k$ (and so $\text{Prob}(S = k) = (12 - k)/36$ here). Note these probabilities are non-negative and sum to 1.

Problem: Page 158: #10. Have $n + m$ independent Bernoulli trials with probability of success p . S_n is the number of successes in the first n trials, T_m in the last m .

Solution: (a) What is the distribution of S_n ? It is just Binomial with parameters n and p as sums of Bernoulli are Binomial and can just forget about the last m .

(b) What is the distribution of T_m ? Similarly, it is Binomial with parameters m and p .

(c) What is the distribution of $S_n + T_m$? It is just Binomial with parameters $n + m$ and p .

(d) Are S_n and T_m independent? Yes as we are assuming p is given. If p were unknown then, as in an example in the book, one could argue that observing S_n gives information on T_m .

Problem: Page 158: #16b. From #3 of this section, if S is the sum of two rolls of a die, we know $\text{Prob}(S = k)$ is $k/36$ if $k \leq 6$ and $(12 - k)/36$ if $k \geq 6$. Using the convolution formula, to find the probability the sum of four rolls of a die is 8 is

$$\sum_{\ell=2}^6 \text{Prob}(\text{first two rolls are } \ell) \text{Prob}(\text{second two rolls are } 8 - \ell);$$

the range of summation is from 2 to 6 as the sum of two rolls is at least 2 and at most 12. We find this probability is

$$\frac{2}{36} \frac{6}{36} + \frac{3}{36} \frac{5}{36} + \frac{4}{36} \frac{4}{36} + \frac{5}{36} \frac{3}{36} + \frac{6}{36} \frac{2}{36} = \frac{70}{1296} \approx 5.4\%.$$

Problem: If X is uniform on $[-1, 1]$, find the probability density function of Y if $Y = X^2$.

Solution: Let f_X, F_X be the pdf and cdf of the random variable X , and f_Y, F_Y the pdf and cdf of the random variable Y . We solve this problem by using the CDF method. We first compute the relevant

quantities for X . We have $f_X(x) = 1/2$ if $-1 \leq x \leq 1$ and 0 otherwise; it has to be $1/2$ as we need $f_X(x) = c$ in $[-1, 1]$ for some constant; as the length of this interval is 2 and the integral must be 1, the height c must be $1/2$. As $F_X(x) = \int_{-\infty}^x f_X(t)dt$, we have $F_X(x) = 0$ if $x \leq -1$, $(x+1)/2$ if $-1 \leq x \leq 1$, and 1 if $x \geq 1$.

We now turn to Y . We have

$$\begin{aligned}
 F_Y(y) &= \text{Prob}(Y \leq y) \\
 &= \text{Prob}(X^2 \leq y) \\
 &= \text{Prob}(-\sqrt{y} \leq X \leq \sqrt{y}) \\
 &= \text{Prob}(X \leq \sqrt{y}) - \text{Prob}(X \leq -\sqrt{y}) \\
 &= F_X(\sqrt{y}) - F_X(-\sqrt{y}).
 \end{aligned}$$

We might as well assume $y \in [0, 1]$; if $y < 0$ then $F_Y(y) = 0$ and if $y > 1$ then $F_Y(y) = 1$. For $y \in [0, 1]$, $\sqrt{y} \in [0, 1]$ and hence $F_Y(y) = \frac{\sqrt{y}+1}{2} - \frac{-\sqrt{y}+1}{2} = \sqrt{y}$. Thus $f_Y(y) = F_Y'(y) = \frac{1}{2}y^{-1/2}$. Note this is non-negative and does integrate to 1 on the interval $[0, 1]$.

5 HW #5

Problem: Page 182: #1: 10% numbers on a list are 15, 20% are 25, rest are 50. What is the average?

Solution: Let there be n numbers. The mean is $\frac{10n}{100n} \cdot 15 + \frac{20n}{100n} \cdot 25 + \frac{70n}{100n} \cdot 50 = \frac{150+500+3500}{100} = 41.5$. Note the answer is greater than 15 (the smallest number on our list) and smaller than 50 (the largest on our list). Also 70% of the numbers are 50, so we expect the mean to be close to 50.

Problem: Page 182, #4: All 100 numbers in list are non-negative and average is 2. Prove that at most 25 exceed 8.

Solution: Imagine there were 26 that were greater than 8. What would these contribute to the mean? Well, if the 26 numbers were 8, we would have a contribution to the mean of $\frac{26}{100} \cdot 8 = 2.08$; as the other numbers are non-negative, the mean would have to be at least 2.08, contradicting the fact that the mean is 2.

Problem: Page 182, #7: Have n switches, i -th is closed with probability p_i , X is the number of switches closed. What is $\mathbb{E}[X]$, or is there not enough information?

Solution: Let $X_i = 1$ if switch i is closed, and 0 if it is not. Note $\mathbb{E}[X_i] = p_i$ and $X = X_1 + \dots + X_n$. By linearity of expectation, $\mathbb{E}[X] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n] = p_1 + \dots + p_n$. Without knowing these values, cannot say more.

Problem: Page 182, #10: A and B independent events with indicator random variables I_A and I_B .

Solution: (a) What is the distribution of $(I_A + I_B)^2$? Squaring, it is $I_A^2 + 2I_A I_B + I_B^2 = I_A + 2I_A I_B + I_B$ as the square of an indicator random variable is just the indicator. It can only take on the values 0, 1, and 4. It is zero when A and B don't happen, or it is 0 with probability $(1 - \Pr(A))(1 - \Pr(B))$. It is 1 if exactly one of A and B happens, so it is 1 with probability $\Pr(A)(1 - \Pr(B)) + (1 - \Pr(A))\Pr(B)$. It is 2 if both happens, or it is 2 with probability $\Pr(A)\Pr(B)$.

(b) What is $\mathbb{E}[(I_A + I_B)^2]$? We use $I_A^2 + 2I_A I_B + I_B^2 = I_A + 2I_A I_B + I_B$ and the linearity of expectation to see that this is $\mathbb{E}[I_A] + 2\mathbb{E}[I_A I_B] + \mathbb{E}[I_B]$. The middle term is just $\mathbb{E}[I_A]\mathbb{E}[I_B] = \Pr(A)\Pr(B)$ as the random variables are independent, and so this answer is just $\Pr(A) + 2\Pr(A)\Pr(B) + \Pr(B)$. As an aside, if we had $(I_A + I_B)^n$, what do you think this will approximately equal for n large?

Problem: Page 182, #11: There are 100 prize tickets among 1000 in the lottery, what is the expected number of prize tickets you'll get if you buy 3 tickets? What is a simple upper bound for the probability you win at least one prize – why is your bound so close?

Solution: We can have either 0, 1, 2 or 3 prize tickets. We could try to do this by using binomials / various probability distributions, but linearity of expectation is a great way. Let p_i be the probability that our i -th ticket is a winner; let X_i be 1 if the i -th ticket is a winner and 0 otherwise. Set $X = X_1 + X_2 + X_3$. As each $p_i = 100/1000 = .1$, each $\mathbb{E}[X_i] = .1$, and by linearity of expectation $\mathbb{E}[X_1 + X_2 + X_3] = .3$. What is amazing is that it's okay that the random variables are dependent! Is this result surprising? No; we expect to need 10 tickets before we have one winning ticket, so having about .3 of a winning ticket is reasonable. The probability we do not win a prize is $\binom{900}{3} / \binom{1000}{3} = 403651/553890 \approx 72.8757\%$ (we have to choose 3 tickets from the 900 losing tickets). A quick approximation is to say each ticket wins

with probability .1, and then use Boole's inequality (page 32) to get about .7. The reason it's so close is that the probability of two or more winning tickets is so small.

Problem: Page 182, #13ad: Roll fair die 10 times. Find numerical values for expectations of (a) sum of numbers in the rolls and (d) the number of multiples of 3 in the ten rolls.

Solution: (a) We've shown before that the expected value from rolling a fair die is 3.5. As expectation is linear, if we roll 10 die it's just 10 times this, or 35. Equivalently, let X_i be the number on the i -th roll, $X = X_1 + \cdots + X_{10}$, and note each $\mathbb{E}[X_i] = 3.5$.

(b) The multiples of 3 are 3 and 6, so each roll has a $2/6$ or $1/3$ chance of being a 3. Let Y_i be 1 if the i -th roll is a multiple of 3 and 0 otherwise. We have $\mathbb{E}[Y_i] = \frac{2}{6} \cdot 1 + \frac{4}{6} \cdot 0 = 2/6$. Setting $Y = Y_1 + \cdots + Y_{10}$ and again using linearity of expectation, we find $\mathbb{E}[Y] = 10 \cdot \frac{2}{6} = 20/6 \approx 3.33$.

6 HW #6

Problem: Page 202, #2: Let Y be the number of heads from tossing a fair coin 3 times. Find the mean and variance of Y^2 .

Solution: First we find the density of Y . We have $\Pr(Y = 0) = 1/8$, $\Pr(Y = 1) = 3/8$, $\Pr(Y = 2) = 3/8$ and $\Pr(Y = 3) = 1/8$. Thus $\Pr(Y^2 = 0) = 1/8$, $\Pr(Y^2 = 1) = 3/8$, $\Pr(Y^2 = 4) = 3/8$ and $\Pr(Y^2 = 9) = 1/8$. The mean is thus $\frac{1}{8} \cdot 0 + \frac{3}{8} \cdot 1 + \frac{3}{8} \cdot 4 + \frac{1}{8} \cdot 9 = \frac{24}{8} = 3$. The variance of Y^2 is $\frac{1}{8} \cdot (0 - 3)^2 + \frac{3}{8} \cdot (1 - 3)^2 + \frac{3}{8} \cdot (4 - 3)^2 + \frac{1}{8} \cdot (9 - 3)^2 = 60/8 \approx 8.57143$.

Problem: Page 202, #3abcd: X, Y, Z independent, identically distributed random variables with mean 1 and variance 2.

Solution: (a) $\mathbb{E}[2X+3Y] = 2\mathbb{E}[X]+3\mathbb{E}[Y] = 2 \cdot 1 + 3 \cdot 1 = 5$. (b) $\text{Var}(2X+3Y) = 4\text{Var}(X)+9\text{Var}(Y) = 4 \cdot 2 + 9 \cdot 2 = 26$. (c) $\mathbb{E}[XYZ] = \mathbb{E}[X]\mathbb{E}[Y]\mathbb{E}[Z] = 1^3 = 1$. (d) $\text{Var}(XYZ) = \mathbb{E}[X^2Y^2Z^2] - \mathbb{E}[XYZ]^2 = \mathbb{E}[X^2]\mathbb{E}[Y^2]\mathbb{E}[Z^2] - \mathbb{E}[X]^2\mathbb{E}[Y]^2\mathbb{E}[Z]^2$. As $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 2$, so since $\mathbb{E}[X] = 1$ we have $\mathbb{E}[X^2] = \text{Var}(X) + \mathbb{E}[X]^2 = 3$. Thus $\text{Var}(XYZ) = 3^3 - 1^3 = 26$.

Problem: Page 202, #12: Random variable X with expectation 10 and standard deviation 5.

Solution: (a) Find the smallest upper bound you can for $\Pr(X \geq 20)$. Note that 20 is 2 standard deviations above the mean, and thus by Chebyshev's inequality the probability of being at least 20 is at most $1/4$. (b) Could X be a binomial random variable? If yes, it would have some n and a probability p . We would have to solve $\mathbb{E}[X] = np = 10$, $\text{Var}(X) = np(1-p) = 5$. There are many ways to do the algebra. Substitute for np , which must be 10, in the variance equation to find $10(1-p) = 5$, so $1-p = 1/2$ or $p = 1/2$. This then gives us $n = 20$, so yes, it is possible.

Problem: Page 202, #14: Suppose average family income is \$10,000.

Solution: (a) Find upper bound for percentage of families with income over \$50,000. Note that income is non-negative (we hope!), so let's try Markov's inequality. So $\Pr(I \geq \$50,000) \leq \mathbb{E}[I]/\$50,000 = 10000/50000 = 1/5$. (b) If we know the standard deviation is \$8,000, then we see that we are 4 standard deviations from the mean, so by Chebyshev the probability of being at least 4 standard deviations away from the mean is at most $1/4^2$. Not surprisingly, we can do much better when we know more.

Problem: Page 202, #27a: X is a random variable with $0 \leq X \leq 1$ and $\mathbb{E}[X] = \mu$. Show that $0 \leq \mu \leq 1$ and $0 \leq \text{Var}(X) \leq \mu(1-\mu) \leq 1/4$. (b) Generalize.

Solution: (a) As $0 \leq X \leq 1$, $0 \leq \mathbb{E}[X] = \mu \leq 1$. For the second claim, note $0 \leq X^2 \leq X \leq 1$ as $0 \leq X \leq 1$. As $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mu^2$, and $\mathbb{E}[X^2] \leq \mathbb{E}[X] \leq \mu$, we have $\text{Var}(X) \leq \mu - \mu^2 = \mu(1-\mu)$. Since $\mu \in [0, 1]$, a calculus exercise shows the maximum of the function $g(\mu) = \mu(1-\mu)$ occurs when $\mu = 1/2$, leading to the value $1/4$. Another way to see this is to note $\mu(1-\mu) = -\mu^2 - \mu = -(\mu - 1/2)^2 - 1/4$; as $(\mu - 1/2)^2 \geq 0$ the minimum value is $1/4$.

(b) The argument proceeds similarly. As $a \leq X \leq b$, $\int_a^b ap(x)dx \leq \int_a^b xp(x)dx \leq \int_a^b bp(x)dx$, so $a \leq \mathbb{E}[X] \leq b$. For the variance, we could use $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, but it's better to reduce to part (a). Let $Y = (X - a)/(b - a)$. Note $0 \leq Y \leq 1$ and $\mu_Y = (\mu_X - a)/(b - a)$. By part (a), the variance of

Y is at most $\mu_Y(1 - \mu_Y)$, which gives

$$\text{Var}(X) \leq \frac{\mu_X - a}{b - a} \left(1 - \frac{\mu_X - a}{b - a}\right) = \frac{(\mu_X - a)(b - \mu_X)}{(b - a)^2}.$$

Note that $\text{Var}(Y) = \text{Var}((X - a)/(b - a)) = \text{Var}(X)/(b - a)^2$. Thus $\text{Var}(X) \leq (\mu_X - a)(b - \mu_X)$. Using calculus, we see this is largest when $\mu_X = \frac{b+a}{2}$, which gives after some algebra $\text{Var}(X) \leq \frac{1}{4}(b - a)^2$. (To see this, let $f(u) = (u - a)(b - u) = -u^2 + (b - a)u - ab$, so $f'(u) = -2u + (b - a)$, so the critical point is where $u = (b - a)/2$.)

(c) If half the numbers are 9 and half are 0, then the mean is 4.5 and the variance is 4.5, as everything is 4.5 units from the mean. From part (b), the maximum the variance of X can be is $\frac{1}{4}(9 - 0)^2 = 20.25 = 4.5^2$. Thus the variance is as large as possible. This forces the mean to be 4.5, and then the variance is maximized when half are 0 and half are 9. It's not surprising that parts (a) and (b) are useful here.

7 HW #7: Due Thursday, October 27

Problem: Section 3.4, Page 217: #1abc. Coin lands with probability p of heads, tosses independent.

Solution: (a) Probability of exactly 5 heads in first 9 tosses: need 5 heads, 4 tails, order doesn't matter, so is just $\binom{9}{5} p^5(1-p)^4 = 126p^5(1-p)^4$. (b) Probability first head on the 7th toss: start with 6 tails followed by a head, only one way this can happen: $(1-p)^6p$. (c) Probability fifth head is on the 12th toss. Means the 12th toss is a head and exactly 4 heads in first 11 (and thus 7 tails in the first 11). Probability is $\binom{11}{4} p^4(1-p)^7 \cdot p$.

Problem: Section 3.4, Page 217: #7abc. A and B toss a coin that's heads with probability p .

Solution: (a) What is the probability A gets the first head? We did this in class. Let x be the probability A wins. Then $x = p + (1-p)^2x$, so $x = \frac{p}{1-(1-p)^2}$. (b) What is the probability B tosses the first head? It is $1-x = \frac{1-(1-p)^2-p}{1-(1-p)^2}$. (c) Now B gets to go twice. It can get hard if we try to think of the game stopping after B tosses once or twice. The right way is to regard it as B's turn. B could fail both times, which happens with probability $(1-p)^2$, or he could have at least one success, which happens with probability $1-(1-p)^2$. So now the probability A wins, say \tilde{x} , satisfies

$$\tilde{x} = p + (1-p)(1-p)^2x \Rightarrow x = \frac{p}{1-(1-p)^3}.$$

Similarly, B now wins with probability 1 minus this.

Problem: Page 217, #10: X is the number of Bernoulli trials (with parameter p) required to produce at least one success and at least one tail.

Solution: (a) What is the distribution of X ? Note $\Pr(X < 2) = 0$. We either start with a head and then wait for a tail, or start with a tail and then wait for a head. If we start with a head (happens with probability p), then the probability it takes $n-1$ tosses to get our first tail is $p^{n-2}(1-p)$; if we start with a tail (happens with probability $1-p$), then the probability it takes $n-1$ tosses to get our first head is $(1-p)^{n-2}p$. Thus $\Pr(X = n) = p^{n-1}(1-p) + (1-p)^{n-1}p$.

(b, c) We want to compute the mean and variance of X . Let X_h be the random variable on how long we have to wait for our first head with a coin with probability p of heads, and let X_t be how long we have to wait for our first tail with this coin. Then $X = (1-p)X_h + pX_t + 1$. Why? With probability $1-p$ we start with a tail and then need to wait for a head, with probability p we start with a head and need to wait for a tail. We then need to add 1 for the first toss. Thus $\mathbb{E}[X] = (1-p)\mathbb{E}[X_h] + p\mathbb{E}[X_t] + 1$ by linearity of expectation. From the book, we know $\mathbb{E}[X_h] = 1/p$, and $\text{Var}(X_h) = (1-p)/p^2$. Similarly $\mathbb{E}[X_t] = 1/q = 1/(1-p)$ and $\text{Var}(X_t) = (1-q)/q^2 = p/(1-p)^2$. Substituting gives $\mathbb{E}[X] = (1-p)\frac{1}{p} + p\frac{1}{1-p} + 1 = (1-p+p^2)/(p-p^2)$, while $\text{Var}(X) = (1-p)^2\frac{1-p}{p^2} + p^2\frac{p}{(1-p)^2} = ((1-p)^5 + p^5)/((-1+p)^2p^2)$. You could also do this problem by directly computing the mean and the variance, but it is *much* easier to observe that we can write X as a sum of two simpler random variables.


```

succandfail[num_, p_] := Module[{},
  wait = 0;
  For[n = 1, n <= num, n++,
    {
      numsucc = 0;
      numfail = 0;
      count = 0;
      While[numsucc * numfail == 0,
        {
          If[Random[] <= p, numsucc = numsucc + 1, numfail = numfail + 1];
          count = count + 1;
        }];
      wait = wait + count;
    }];
  Print["Observed average wait time: ", 1.0 wait / num];
  Print["Predicted average wait time: ",
    1 + (p/(1 - p)) + ((1 - p)/p)];
];

```

Problem: Section 3.5, Page 234: #4. Books have average of 1 misprints per page, what is the probability at least one page in a 300 page book has at least 5 misprints?

Solution: Assuming the number of misprints per page is a Poisson random variable with parameter $\lambda = 1$, and the pages are independent. The probability of 4 or fewer typos on a page is $\sum_{n=0}^4 e^{-1} 1^n / n! = 65/24e$. Thus the probability that none of the 300 pages has 5 or more typos is $(1 - \frac{65}{24e})^{300} \approx 1.09 \cdot 10^{-731}$ (so not that likely!).

Problem: Section 4.4, Page 310: #9a. Show if T is Weibull (λ, α) then T^α is exponential with parameter λ .

Solution: This is a change of variable problem. Let $Y = T^\alpha$. We do our CDF approach. We have

$$\begin{aligned}
 F_Y(y) = \text{Prob}(Y \leq y) &= \text{Prob}(T^\alpha \leq y) \\
 &= \text{Prob}(T \leq y^{1/\alpha}) = F_T(y^{1/\alpha}).
 \end{aligned}$$

We now use the Chain rule and find

$$f_Y(y) = F'_Y(y) = F'_T(y^{1/\alpha}) \frac{1}{\alpha} y^{1/\alpha - 1}.$$

Substituting gives $f_Y(y) = \lambda e^{-\lambda y}$ as desired.

8 HW #8: Due Tuesday, November 15

Problem: Section 5.3: #3a. W, X, Y, Z independent standard normal. What is the probability that $W + X > Y + Z + 1$?

Solution: If $A_i \sim N(\mu_i, \sigma_i^2)$ then $A_1 + A_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ (assuming of course they're independent). Thus $W + X, Y + Z \sim N(0, 2)$. This seems to require some integration, so let's do some additional thinking. Note our condition is equivalent to $W + X - Y - Z > 1$, and $-Y, -Z$ are also standard normals. Note $W + X - Y - Z \sim N(0, 4)$, and thus we're asking what is the probability a normal random variable with mean 0 and variance 4 is at least 1. It would be nicer to have a random variable with mean 0 and variance 1, so let's set $S = (W + X - Y - Z)/2$. Note the mean of S is 0 and the variance is 1, and our condition is now $S > 1/2$, where $S \sim N(0, 1)$. Letting Φ be the cumulative distribution function of the standard normal, the answer is just $1 - \Phi(.5)$. It's amazing how much easier these problems are if you look at them the right way and spend some time doing algebra.

Problem: Calculate the probability a chi-square distribution with 2 degrees of freedom is at least twice as large as its mean (so if the mean is μ , we want the probability it is 2μ or greater).

Solution: The density function is $f(x) = \frac{1}{2}e^{-x/2}$. The mean of a chi-square distribution with k degrees of freedom is k , so the mean is 2. We want the probability of being 4 or more, which is just

$$\int_4^{\infty} \frac{1}{2}e^{-x/2}dx = e^{-x/2} \Big|_4^{\infty} = e^{-2}.$$

Problem: Problems from my handout on Method of Least Squares notes: Page 9, #3.3. Consider the observed data $(0, 0), (1, 1), (2, 2)$. Show that if we use (2.10) from the Least Squares handout to measure error then the line $y = 1$ yields zero error, and clearly this should not be the best fit line!

Solution: We will use equation (2.10) to calculate the error of the line $y = 1$. This gives an error function $E_2(a, b) = \sum_{n=0}^N (y_i - (ax_i + b))$. Evaluating the sum with the line $y = 1$ (which means $a = 0$ and $b = 1$) gives an error of $E_2(0, 1) = (0 - 1) + (1 - 1) + (2 - 1) = 0$. The problem with (2.10) is that the errors are signed quantities, so during the calculation the positive errors cancel out the negative errors.

Problem: Problems from my handout on Method of Least Squares notes: Page 10, #3.9. Show that the Method of Least Squares predicts the period of orbits of planets in our system is proportional to the length of the semi-major axis to the $3/2$ power.

Solution: Using the numbers from the handout, namely We

$$a = \frac{\sum_{n=1}^N 1 \sum_{n=1}^N x_n y_n - \sum_{n=1}^N x_n \sum_{n=1}^N y_n}{\sum_{n=1}^N 1 \sum_{n=1}^N x_n^2 - \sum_{n=1}^N x_n \sum_{n=1}^N x_n}, \quad b = \frac{\sum_{n=1}^N x_n \sum_{n=1}^N x_n y_n - \sum_{n=1}^N x_n^2 \sum_{n=1}^N y_n}{\sum_{n=1}^N x_n \sum_{n=1}^N x_n - \sum_{n=1}^N x_n^2 \sum_{n=1}^N 1},$$

we find $N = 8$, $\sum_{n=1}^8 1 = 8$, $\sum_{n=1}^8 x_n = 9.409461$, $\sum_{n=1}^8 y_n = 14.1140384$, $\sum_{n=1}^8 x_n^2 = 29.29844102$ and $\sum_{n=1}^8 x_n y_n = 43.94486382$. Feeding these into the equations for a and b in the handout give best fit values of $a = 1.49985642$ and $b = 0.000149738$ (the reason b is so close to zero is we have chosen to measure distances in astronomical units, precisely to make the proportionality constant nice).