Math 10B with Professor Stankova
Worksheet, Midterm #2; Wednesday, 4/3/2019
GSI name: Roy Zhao

1 Concepts

1. **Independence**: What does it mean for two events to be independent?

   - What is the difference between independence and mutually disjoint/exclusive? Can two events be both? Neither?
   - What does it mean for three or more events to be independent?

   What does it mean for two random variables to be independent?

   - How do we show two random variables are independent? Dependent?
   - How can we use covariance to show that two random variables are independent? Dependent?

2. **Discrete Random Variables**: Can you draw the picture relating a random variable to a probability space and the PMF?

   - How are the PMF of a random variable and probability function P of a probability space related?
   - How do you calculate and draw a PMF?

   What is the definition of the expected value, variance, standard error, and covariance?

   - How would you explain the expected value, variance, standard error, and covariance in words to a 5 year old?
   - What are some properties of the expected value?
     - When can we split up $E[X+Y] = E[X] + E[Y]$? When can we split up $E[XY] = E[X]E[Y]$?
   - What is the "shortcut" formula for the variance? Can you prove it?
   - When can we split up $Var[X+Y] = Var[X] + Var[Y]$? When can we split up $Var[XY] = Var[X]Var[Y]$?
   - What is the expected value and variance of a constant?
   - What is the relationship between the variance and standard error? Why do we have both of them?
   - How can we recover the variance from the covariance?

Fill out the following table:

Distribution	Range(X)	PMF	$E(X)$	Variance	$SE(X)$	Example
Uniform						
Bernoulli Trial						
Binomial						
Geometric						
Hyper-Geometric						
Poisson						

3. **Limit Theorems**: What does i.i.d stand for?
   - What does $\bar{X}$ represent?
   - What do $\bar{\mu}, \bar{\sigma}$ represent?
   - How do $\mu, \sigma, \bar{\mu}, \bar{\sigma}, X, \bar{X}$ relate to each other?
   - What is the difference between $\mu, \bar{\mu}, \bar{X}$?
   - Can we say anything about whether σ or $\bar{\sigma}$ is bigger?

   What do the Central Limit Theorem (CLT) and Law of Large Numbers (LoLN) state?
   - Can we use CLT to prove LoLN or vice versa?
   - How does the CLT relate to z-scores? Is this relationship exact?
   - Pictorially, what are the CLT and LoLN saying?

4. **Continuous Random Variables:** How do the PMF and PDF differ?
   - What properties must a function satisfy to be a PDF? A CDF?
   - What is the relationship between a PDF and CDF?
   - How do we calculate the median and mean?
   - How do the median and mean relate? When are they equal? When is the mean bigger? When is the mean smaller?
   - How are histograms related to PDFs?
   - Find the PDF, CDF, mean, and median of the following distributions:
     - Continuous uniform
     - Dart-player
     - Normal distribution
     - Exponential
     - Pareto

2 Problems

2.1 Discrete Distributions

5. Suppose that we roll two die and let X be equal to the maximum of the two rolls. Find $P(X \in \{1, 3, 5\})$ and draw the PMF for X. What is the expected value and standard error of X?

Solution: First we draw the PMF. We calculate $P(X = x)$ by counting the number of ways we can roll two die so that the maximum is x and then dividing by the total number of possibilities, which is 36. So for instance, the only way to get $X = 1$ is if we roll $(1, 1)$ and hence $P(X = 1) = \frac{1}{36}$. Then $P(X = 2) = \{(1, 2), (2, 2), (2, 1)\}/36 = \frac{3}{36}$. Thus, we have that

$$f(1) = \frac{1}{36}, f(2) = \frac{3}{36}, f(3) = \frac{5}{36}, f(4) = \frac{7}{36}, f(5) = \frac{9}{36}, f(6) = \frac{11}{36}.$$

We draw the PMF with stalks at 1 through 6 of those respective heights. Then $P(X \in \{1, 3, 5\}) = P(X = 1) + P(X = 3) + P(X = 5) = \frac{1}{36} + \frac{5}{36} + \frac{9}{36} = \frac{15}{36} = \frac{5}{12}$.

The expected value is

$$E[X] = 1 \cdot f(1) + 2f(2) + \cdots + 6f(6) = \frac{161}{36}.$$

The variance is

$$Var(X) = E[X^2] - E[X]^2 = 1^2 \cdot f(1) + 2^2 f(2) + \cdots + 6^2 f(6) - \frac{161^2}{36^2} = \frac{791}{36} - \frac{161^2}{36^2}.$$

So the standard error is

$$SE(X) = \sqrt{Var(X)} = \sqrt{\frac{791}{36} - \frac{161^2}{36^2}}.$$

6. I draw 5 cards from a deck of cards. Let X be the number of hearts I draw. What is the range of X and draw the PMF of X. Use this to find the probability that I draw at least 2 hearts.

Solution: The range is $\{0, 1, 2, 3, 4, 5\}$. To calculate $f(x) = P(X = x)$, we count the number of good ways over the total number of ways. The number of good ways to draw x hearts is to first pick out x hearts out of the 13 hearts, and then fill out the rest of the hand and pick $5 - x$ non-heart cards from the remaining 39 cards. Thus

$f(x) = \frac{\binom{13}{x}\binom{39}{5-x}}{\binom{52}{5}}$. Thus, we have that $P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) = \frac{\binom{13}{2}\binom{39}{3}+\binom{13}{3}\binom{39}{2}+\binom{13}{4}\binom{39}{1}+\binom{13}{5}\binom{39}{0}}{\binom{52}{5}}$.

7. Suppose I have a weighted 4 sided die that lands on 1 with probability $\frac{1}{2}$ and lands on $2, 3, 4$ with equal probability. Let X be the value of the die when I roll it once. What is the PMF for X. What is $E[X]$ and $Var(X)$?

Solution: The PMF is

x	1	2	3	4
$f(x)$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

The expected value is

$$E[X] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} = 2.$$

The variance is

$$Var(X) = E[X^2] - E[X]^2 = 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} - 2^2 = \frac{32}{6} - 4 = \frac{4}{3}.$$

8. I am picking cards out of a deck. What is the probability that I pull out 2 kings out of 8 cards if I pull with replacement? What about without replacement?

Solution: With replacement is repeated Bernoulli trials which means binomial distribution. The probability of a success or pulling out a heart is $\frac{1}{13}$. Therefore, the probability of pulling 2 kings out of 8 is

$$\binom{8}{2}\left(\frac{1}{13}\right)^2 \cdot \left(\frac{12}{13}\right)^{8-2}.$$

If do not have replacement, then this is a hyper-geometric distribution with $N = 52, n = 8, m = 4$, the the answer is

$$\frac{\binom{4}{2}\binom{48}{6}}{\binom{52}{8}}.$$

9. What is the probability that first king is the third card I draw (with replacement)?

Solution: We want to know the probability of the first success, which is geometric since we are doing with replacement. The probability of a success is $p = \frac{1}{13}$ and we have two failures before we have a success so $k = 2$. Hence the answer is $f(2) = (12/13)^2(1/13)$.

10. In a class of 40 males and 60 females, I give out 4 awards randomly. What is the probability that females will win all 4 awards if the awards must go to different people? What about if the same person can win multiple awards?

Solution: This is like the probability of picking 4 females out of 4 people chosen. If the awards must go to different people, there is no replacement so it is the hypergeometric distribution where a success is picking a female. So we have $N = 100$ students total, there are $m = 60$ females, and I am picking $n = 4$ students and I want $k = 4$ females. This gives

$$f(4) = \frac{\binom{60}{4}\binom{40}{0}}{\binom{100}{4}} = \frac{\binom{60}{4}}{\binom{100}{4}}.$$

If the same person can will all the awards, then we are choosing with replacement. So, this is a binomial distribution where the probability of success is $p = \frac{60}{100}$. Thus, we have that the answer is

$$f(4) = \binom{4}{4}\left(\frac{60}{100}\right)^4\left(\frac{40}{100}\right)^0 = \frac{3^4}{5^4}.$$

11. I am playing the lottery and I have a 0.1% chance of winning. What is the probability that I need to play at least 100 times until winning?

Solution: This is a geometric distribution and since I am asking for the probability of playing at least 100 times before winning, I am asking for $P(X \geq 100)$. We have a formula for the geometric distribution which tells us that this is

$$P(X \geq 100) = (1-p)^{100} = (0.999)^{100}.$$

12. In a class of 50 males and 90 females, I give out 4 awards randomly. What is the probability that females will win 2 awards if the awards must go to different people? What about if the same person can win multiple awards?

Solution: This is like the probability of picking 2 females out of 4 people chosen. If the awards must go to different people, there is no replacement so it is the hypergeometric distribution where a success is picking a female. So we have $N = 140$ students total, there are $m = 90$ females, and I am picking $n = 4$ students and I want $k = 2$ females. This gives

$$f(2) = \frac{\binom{50}{2}\binom{90}{2}}{\binom{140}{4}}.$$

If the same person can will all the awards, then we are choosing with replacement. So, this is a binomial distribution where the probability of success is $p = \frac{90}{140}$. Thus, we have that the answer is

$$f(2) = \binom{4}{2}\left(\frac{90}{140}\right)^2\left(\frac{50}{140}\right)^2.$$

13. At Berkeley, 3/4 of the population is undergraduates. I cold call someone at random and ask for their age. What is the probability that I have to call 10 people before I call an undergraduate? What is the probability that I call 4 undergraduates out of 10 people I call (if I can call someone more than once)?

Solution: The probability of calling an undergrad is $\frac{3}{4}$. The probability that I call 10 people before the first undergrad is given by the geometric distribution since we are talking about times until a success. I have 10 failures before and so this is

$$f(10) = (1 - 3/4)^{10}(3/4) = \frac{3}{4^{11}}.$$

The probability that I call 4 undergraduates out of 10 people is given by a binomial distribution since I can call someone more than once and hence there is replacement. So plugging this into the binomial distribution gives

$$f(4) = \binom{10}{4}(3/4)^4(1/4)^6.$$

14. Suppose that the probability that a child is a boy is 0.51. What is the probability that a family of five children has at least one girl? What is the probability that the family of 5 has all children of the same sex?

Solution: The probability that they have at least one girl is 1 minus the probability they have no girls which is $1 - 0.51^5$. For the family to have children all of the same sex, they must have all boys or all girls. The probability of all boys is 0.51^5 while the probabilities of all girls is 0.49^5. So the probability of all boys or all girls is $0.51^5 + 0.49^5$.

15. In a dorm of 100 people, there are 20 people who are underage. I go to a party with 40 people. What is the probability that there is at least one underage person there?

Solution: The problem says at least which clues te fact that we should think about complimentary probability. The probability that there is at least one underage person is 1 minus the probability that there are no underage people there. In order to calculate this latter probability, we determine what kind of distribution this is. There are $m = 20$ "tagged" people who are underage and out of a total population of $N = 100$ people, we want to choose $n = 40$ people and want to get $k = 0$ minors. This is a hyper-geometric distribution because we are picking people without replacement and hence the probability of having at least one underage person is

$$1 - \frac{\binom{20}{0}\binom{80}{40}}{\binom{100}{40}} = 1 - \frac{\binom{80}{40}}{\binom{100}{40}}.$$

16. In a class of 30 students, I split them up into 6 groups of 5. What is the expected number of days of splitting them up randomly into new groups of 5 before I split them up into the same groups again (assume that the groups are indistinguishable)?

Solution: First we count the number of ways to split up 30 people into 6 groups of 5. First we choose the first group and that is $\binom{30}{5}$ ways, then the second is $\binom{25}{5}, \ldots,$ and the last group is $\binom{5}{5}$ ways. Finally, the 6 groups are indistinguishable which means it doesn't matter which way we order the groups so we need to divide by 6!. Thus, the total number of ways to make the groups is

$$N = \frac{\binom{30}{5}\binom{25}{5}\binom{20}{5}\binom{15}{5}\binom{10}{5}\binom{5}{5}}{6!}.$$

Since each of these group splitting is chosen randomly, the probability of splitting them up the same way is $p = \frac{1}{N}$. And the expected number of days I need before splitting them up in the same way is $\frac{1-p}{p} = N - 1$.

17. When a cell undergoes mitosis, the number of mutations that occurs is Poisson distributed and an average of 8 mutations occur. What is the probability that no more than 1 mutation occurs when two cells divide?

Solution: This is a Poisson process with $\lambda = 8$ for one cell, but if two cells divide, we expect an average of $\lambda = 16$ mutations to occur. We want to know $P(X \leq 1) = P(X = 0) + P(X = 1) = f(0) + f(1) = \frac{\lambda^0 e^{-\lambda}}{0!} + \frac{\lambda^1 e^{-\lambda}}{1!} = e^{-16} + 16e^{-16} = 17e^{-16}$.

18. I roll a fair 6-sided die over and over again until I roll a 6. What is the probability that it takes me more than 10 tries? What is the expected number of total rolls I need and what is the variance?

Solution: Let X be the number of failures before a success (rolling a 6). Then X is given by a geometric distribution with $p = \frac{1}{6}$. We want to know $P(X \geq 10)$ which we can calculate as

$$P(X \geq 10) = P(X = 10)+P(X = 11)+P(X = 12)+\cdots = (1-p)^{10}p+(1-p)^{11}p+\cdots.$$

This is a geometric series which is equal to $\frac{(1-p)^{10}p}{1-(1-p)} = (1-p)^{10}$. In fact, for a geometric series, we have $P(X \geq n) = (1-p)^n$ and look at Homework 15 for the justification. Now for a geometric distribution, we have $E[X] = \frac{1-p}{p} = 5$ but then we need to add 1 to include the final successful roll to get an answer of 6. The variance is $\frac{1-p}{p^2} = 30$.

19. I am throwing darts some number of times and suppose that I expect to hit the dart board 20 times with a standard error of 2 times. What is the probability that I hit the dart board on a single throw?

Solution: We are in a binomial distribution with the expected value being $np = 20$ and standard error $\sqrt{np(1-p)} = 2$ so $np(1-p) = 4$ and $1 - p = \frac{1}{5}$ so $p = \frac{4}{5}$. Therefore the probability of hitting the dart board is just $p = \frac{4}{5}$.

20. When creating this worksheet, the probability a particular problem has a typo is 1%. What is the probability that I had at most one typo if I created 100 problems?

Solution: This is a binomial distribution because each problem has a typo with probability 0.01. A success in this case is a problem having a typo so $p = 0.01$ and $n = 100$ because we have 100 problems. Then, we want to calculate the probability that I have at most one typo which is either 0 or 1 typo. So we want

$$f(0)+f(1) = \binom{100}{0}(1-0.01)^{100}(0.01)^0+\binom{100}{1}(1-0.01)^{99}(0.01)^1 = 0.99^{100}+100\cdot 0.99^{99}\cdot 0.01 = 0.99^{100}$$

21. The number of chocolate chips in a cookie is Poisson distributed with an average of 15 chocolate chips. What is the probability that you pick up a cookie with only 10 chocolate chips in it?

Solution: This is a Poisson distribution with $\lambda = 15$. We want to calculate $f(10) = \frac{\lambda^{10}e^{-\lambda}}{10!} = \frac{15^{10}e^{-15}}{10!}$.

22. The number of errors on a page is Poisson distributed with approximately 1 error per 50 pages of a book. What is the probability that a novel of 300 pages contains at most 1 error?

Solution: If we average 1 error per 50 pages, then over a novel of 300 pages, we should expect $300/50 \cdot 1 = 6$ errors. Thus, the probability of having at most one error with $\lambda = 6$ is $f(0) + f(1) = \frac{\lambda^0 e^{-\lambda}}{0!} + \frac{\lambda^1 e^{-\lambda}}{1!} = e^{-6} + 6e^{-6} = 7e^{-6}$.

23. This weekend I went to the marina and saw 20 boats at sea out of the 200 total docked at the marina. The next weekend, I go back and see 30 at sea this time. What is the probability that 6 of those boats were also at sea the previous weekend?

Solution: This is a hyper geometric distribution. I want to know how many boats were there the previous weekend and so a success is a boat being out on both days. There are 20 boats that were "tagged" on the first weekend that we care about so $m = 20$ and on the second weekend I see $n = 30$ boats total. Then $N = 200$ and $k = 6$ so the probability is

$$\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} = \frac{\binom{20}{6}\binom{180}{24}}{\binom{200}{30}}.$$

24. Suppose that I am trying to make a half court shot and I have a 1% chance of making it. I try 800 times total and assume the trials are independent (I don't get tired). What is the exact probability that I make exactly 6 shots? Using the Poisson distribution, approximate the probability that I make 6 shots.

Solution: The first distribution is the binomial distribution and using the formula, the probability of getting 6 shots is

$$f(6) = \binom{800}{6}(0.01)^6(0.99)^{794}.$$

Now to use the Poisson distribution, we need to know the expected value λ. The expected value of the binomial distribution was np so we set $\lambda = np = 800 \cdot 0.01 = 8$. Now the probability of getting 6 is

$$f(6) = \frac{\lambda^6 e^{-\lambda}}{6!} = \frac{8^6 e^{-8}}{6!}.$$

25. I roll two fair 6 sided die. What is the expected value of their product?

Solution: Let X be the first value I roll and Y be the second. The rolls are independent and so $E[XY] = E[X]E[Y] = 3.5 \cdot 3.5 = 12.25$.

26. In a class of 30 students, I split them up into 3 groups of 10 on Tuesday. Today, Thursday, I split them up into 6 groups of 5 randomly. What is the expected number of people in your new group were in your old group on Tuesday?

Solution: We can think of this as a hypergeometric distribution where day 1, we "tag" or mark the 9 students that were in your group. Then on day 2, out of the 29 other students, you want to select 4 of them without replacement to be in your group. This is a hyper-geometric distribution with $N = 29, n = 4$. Then $m = 9$ because there are 9 students that were in your group before. So, the expected number of people in your new group who were in your old group is $\frac{mn}{N} = \frac{9 \cdot 4}{29} = \frac{36}{29}$.

27. While pulling out of a box of cookies, what is the expected number of cookies I have to pull out before I pull out an oatmeal raisin if 25% of cookies are oatmeal raisin and I choose with replacement? What is the variance?

Solution: This is a geometric distribution because I am counting the number of cookies I have to pull out before a success. The probability of success is $25\% = p = 1/4$. So the expected number of cookies I have to pull out is $\frac{1-p}{p} + 1 = 3 + 1 = 4$. The variance is $\frac{1-p}{p^2} = 3/(1/4) = 12$.

28. What is the expected number of aces I have when I draw 5 cards out of a deck?

Solution: Drawing cards out of a deck without replacement is the hypergeometric distribution. There are $N = 52$ cards total and $m = 4$ aces total. Then, we pull out $n = 5$ cards and so the expected number of aces is $\frac{mn}{N} = \frac{20}{52}$.

29. Suppose that X is a binomial random variable with expected value 20 and variance 4. What is $P(X = 3)$?

Solution: We are told that $E[X] = np = 20$ and $Var(X) = np(1-p) = 4$ so $1 - p = \frac{1}{5}$ and $p = \frac{4}{5}$ so $n = 25$. Therefore

$$P(X = 3) = \binom{25}{3} p^3 (1-p)^{22} = \binom{25}{3} \left(\frac{4}{5}\right)^3 \frac{1}{5^{22}}.$$

30. In a safari, safari-keepers have caught and tagged 300 rhinos. On a safari, out of the 15 different rhinos you see, there are 5 of them expected to be tagged. How many rhinos are there at the safari?

Solution: This is a hyper-geometric distribution because out of the N rhinos total and $m = 300$ tagged rhinos, you see that $n = 15$ rhinos that you see, there are 5 of them expected to be tagged. So $5 = E(X) = \frac{mn}{N} = \frac{300 \cdot 15}{N}$. So $N = \frac{300 \cdot 15}{5} = 900$.

31. Suppose that I flip a fair coin 8 times. Let T be the number of tails I get and H the number of heads. Calculate $E[T], E[H], Var[T], Var[H], Var[T + H]$. Now calculate $E[T - H]$ and $Var[T - H]$.

Solution: Both T, H are binomial distributions with $T + H = 8$ because there are 8 coin flips total. Thus, using the formula for the binomial distribution with $n = 8, p = \frac{1}{2}$, we get that

$$E[T] = E[H] = np = 4.$$

Then $Var[T] = Var[H] = np(1-p) = 2$. Finally $Var[T + H] = Var[8] = 0$. And also $E[T - H] = E[T] - E[H] = 0$.

Now to calculate the variance, we cannot split it up since T, H are not independent. But, we know that $H + T = 8$ so $T - H = T - (8 - T) = 2T - 8$ and so $Var[T - H] = Var[2T - 8] = Var[2T] = 4Var[T] = 8$.

32. Prove the short cut formula for variance from the definition of variance.

Solution:

$$\begin{aligned} Var[X] &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2] \\ &= E[X^2] - 2E[XE[X]] + E[E[X]^2] \\ &= E[X^2] - 2E[X]E[X] + E[X]^2 \\ &= E[X^2] - E[X]^2. \end{aligned}$$

Where we use the fact that $E[X]$ is a constant so we can move it out of the expected value and the expected value of a constant is the constant itself.

2.2 Limit Theorems

33. Let $X_1, \ldots, X_4$ be i.i.d Bernoulli trials with $p = \frac{3}{4}$. Let $\bar{X}$ be the average of them. What is $Var[\bar{X}]$? Find $Cov(X_1, \bar{X})$ (Hint: Write $\bar{X} = \frac{1}{4}(X_1 + X_2 + X_3 + X_4)$).

Solution: Each of the X_i is Bernoulli so expected value of $p = \frac{3}{4}$ and variance of $p(1-p) = \frac{3}{16}$. Then the variance of $Var[\bar{X}] = Var[X_1]/n = \frac{3}{16 \cdot 4} = \frac{3}{64}$. Finally, we have that

$$\begin{aligned} Cov(X_1, \bar{X}) &= Cov(X_1, \frac{1}{4}(X_1 + X_2 + X_3 + X_4)) \\ &= \frac{1}{4}(Cov(X_1, X_1) + Cov(X_1, X_2) + Cov(X_1, X_3) + Cov(X_1, X_4)) \\ &= \frac{1}{4}(Var(X_1) + 0 + 0 + 0) \\ &= \frac{1}{4}\frac{3}{16} = \frac{3}{64}. \end{aligned}$$

34. Let f be normally distributed with mean 3 and standard error 5. Calculate the probability $P(X \geq 0)$.

Solution: We have $P(X \geq 0) = P(0 \leq X \leq 3) + P(3 \leq X)$ and we calculate the z scores. The second is just $\frac{1}{2}$ and the first is $\frac{|0-3|}{5} = 0.6$. Thus, the probability is $0.5 + z(0.6)$.

35. Let f be normally distributed with mean 2 and standard error 1. Calculate the probability $P(X \leq 0)$.

Solution: We have $P(X \leq 0) = P(X \leq 2) - P(0 \leq X \leq 2)$ and we calculate the z scores. The first is just $\frac{1}{2}$ and the second is $\frac{|0-2|}{1} = 2$. Thus, the probability is $0.5 - z(2)$.

36. Let f be normally distributed with mean 3 and standard error 5. Calculate the probability $P(X \geq 0)$.

Solution: We have $P(X \geq 0) = P(0 \leq X \leq 3) + P(3 \leq X)$ and we calculate the z scores. The second is just $\frac{1}{2}$ and the first is $\frac{|0-3|}{5} = 0.6$. Thus, the probability is $0.5 + z(0.6)$.

37. Let f be normally distributed with mean 2 and standard error 1. Calculate the probability $P(X \leq 0)$.

Solution: We have $P(X \leq 0) = P(X \leq 2) - P(0 \leq X \leq 2)$ and we calculate the z scores. The first is just $\frac{1}{2}$ and the second is $\frac{|0-2|}{1} = 2$. Thus, the probability is $0.5 - z(2)$.

38. Let $X_1, \ldots, X_4$ be i.i.d Bernoulli trials with $p = \frac{3}{4}$. Let $\bar{X}$ be the average of them. What is $Var[\bar{X}]$? Find $Cov(X_1, \bar{X})$ (Hint: Write $\bar{X} = \frac{1}{4}(X_1 + X_2 + X_3 + X_4)$).

Solution: Each of the X_i is Bernoulli so expected value of $p = \frac{3}{4}$ and variance of $p(1-p) = \frac{3}{16}$. Then the variance of $Var[\bar{X}] = Var[X_1]/n = \frac{3}{16 \cdot 4} = \frac{3}{64}$. Finally, we have that

$$\begin{aligned} Cov(X_1, \bar{X}) &= Cov(X_1, \frac{1}{4}(X_1 + X_2 + X_3 + X_4)) \\ &= \frac{1}{4}(Cov(X_1, X_1) + Cov(X_1, X_2) + Cov(X_1, X_3) + Cov(X_1, X_4)) \\ &= \frac{1}{4}(Var(X_1) + 0 + 0 + 0) \\ &= \frac{1}{4}\frac{3}{16} = \frac{3}{64}. \end{aligned}$$

39. Suppose the weight of newborns is distributed with an average weight of 8 ounces and a standard deviation of 1 ounce. Today, there were 25 babies born at the Berkeley hospital. What is the probability that the average weight of these newborns is less than 7.5 ounces?

> **Solution:** The average weight of these babies will be approximately normally distributed with mean 8 and standard deviation $1/\sqrt{25} = 0.2$. The probability is $P(\bar{X} \leq 7.5) = 0.5 - P(7.5 \leq \bar{X} \leq 8) = 0.5 - z(2.5)$.

40. Suppose that the average lifespan of a human is 75 years with a standard deviation of 10 years. What is the probability that in a class of 25 students, they will on average live longer than 80 years?

> **Solution:** The average lifespan of 25 students is approximately normally distributed with mean 75 and standard deviation $10/\sqrt{25} = 2$. Thus $P(\bar{X} \geq 80) = 0.5 - z(|80 - 75|/2) = 0.5 - z(2.5)$.

41. The newest Berkeley quarterback throws an average of 0.9 TDs/game with a standard deviation of 1. What is the probability that he averages at least 1 TD/game next season (16 total games)?

> **Solution:** In 16 games, he will average 0.9 TDs/game with a standard deviation of $1/\sqrt{16} = 0.25$. So the probability that he averages at least 1 TD/game is $P(\bar{X} \geq 1) = 0.5 - P(0.9 \leq \bar{X} \leq 1) = 0.5 - z(\frac{|1-0.9|}{0.25}) = 0.5 - z(0.4)$.

42. Suppose that the average shopper spends 100 dollars during Black Friday, with a standard deviation of 50 dollars. What is the probability that a random sample of 25 shoppers will have spent more than $3000?

> **Solution:** In a sample of 25 shoppers, the average shopper will spend 100 dollars with a standard deviation of $50/\sqrt{25} = 10$. Thus, the probability that a random sample will spend more than 3000 dollars is the probability that a random sample will average more than $3000/25 = 120$ dollars per person. This probability is $0.5 - z(|120 - 100|/10) = 0.5 - z(2)$.

43. Suppose that on the most recent midterm, the average was 60 and the standard deviation 20. What is the probability that a class of 25 had an average score of at least 66?

> **Solution:** In a class of 25, the average score will be distributed with mean 60 and standard deviation $20/\sqrt{25} = 4$. The probability that they had an average score of at least 66 is $0.5 - z(|66 - 60|/4) = 0.5 - z(1.5)$.

2.3 Continuous Random Variables

44. Let $f(x) = \begin{cases} ce^{-x} & -1 \le x \\ 0 & \text{otherwise} \end{cases}$. Find c such that $f(x)$ is a PDF. Graph f and the CDF F. Find the mean and median of $f(x)$.

Solution: First we calculate

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{\infty} ce^{-x}dx = -ce^{-x}|_{-1}^{\infty} = ce.$$

So $ec = 1$ and $c = 1/e$. The CDF is

$$F(x) = \begin{cases} 0 & x \le -1 \\ \int_{-1}^{x} 1/ee^{-t}dt & x \ge -1 \end{cases} = \begin{cases} 0 & x \le -1 \\ 1 - e^{-x}/e & x \ge -1 \end{cases}.$$

The median is when $F(x) = 1/2$ or when $1 - e^{-x}/e = \frac{1}{2}$ which is $x = \ln 2 - 1$. The mean is

$$\mu = \int_{-\infty}^{\infty} xf(x)dx = \int_{-1}^{\infty} xe^{-x}/edx = (-xe^{-x} - e^{-x})|_{-1}^{\infty}/e = (e-e)/e = 0.$$

45. Let $c(x) = \begin{cases} \frac{c}{x^4} & x \le -1 \\ 0 & \text{otherwise} \end{cases}$. Find c such that $f(x)$ is a PDF. Graph f and the CDF F. Find the mean and median of $f(x)$.

Solution: First we calculate

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{-1} \frac{c}{x^4}dx = \frac{-c}{3x^{-3}}|_{-\infty}^{-1} = \frac{c}{3}.$$

Therefore, $c(1/3) = 1$ so $c = 3$. The CDF is

$$F(x) = \begin{cases} \int_{-\infty}^{x} \frac{3}{t^4}dt & x \le -1 \\ 1 & x \ge -1 \end{cases} = \begin{cases} \frac{-1}{x^3} & x \le -1 \\ 0 & x \ge -1 \end{cases}.$$

The median is when $F(x) = 1/2$ or when $-1/x^3 = 1/2$ which is $x = -\sqrt[3]{2}$. The mean is

$$\mu = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{-1} \frac{3}{x^3}dx = \frac{-3}{2}.$$

46. Let $f(x) = \frac{c}{1+x^2}$ for $x \geq 0$ and 0 otherwise. Find c such that $f(x)$ is a PDF. Graph f and the CDF F. Find the mean and median of $f(x)$.

Solution: First we calculate

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} \frac{c}{1+x^2}dx = c\arctan(x)|_0^{\infty} = \frac{c\pi}{2}.$$

Therefore, $\pi/2c = 1$ and $c = 2/\pi$. The CDF is

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \int_0^x \frac{2}{\pi(1+t^2)}dt & x \geq 0 \end{cases} = \begin{cases} 0 & x \leq 0 \\ \frac{2}{\pi}\arctan(x) & x \geq 0 \end{cases}.$$

The median is when $F(x) = 1/2$ or when $2/\pi \arctan(x) = \frac{1}{2}$ which is $x = \tan(\pi/4) = 1$. The mean is

$$\mu = \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{\infty} \frac{2x}{\pi(1+x^2)}dx = 1/\pi \ln(1+x^2)|_0^{\infty} = \infty,$$

so the mean does not exist.

3 True/False

47. True **FALSE** To partition a set Ω into a disjoint union of subsets $B_1, B_2, \ldots, B_n$, means that the intersection of these sets is empty; i.e., $B_1 \cap B_2 \cap \cdots \cap B_n = \emptyset$.

Solution: We need the pairwise intersections $B_i \cap B_j$ to be empty as well.

48. True **FALSE** Two disjoint events could be independent, but two independent events can never be disjoint.

Solution: If one event is the empty set, then it is disjoint and independent with any other event.

49. True **FALSE** If a fair coin comes up Heads six times in a row, it is more likely that it will come up Tails than Heads on the 7th flip.

Solution: If it is fair, the flips are independent.

50. **TRUE** False Contrary to how we may use the word "dependent" in everyday life; e.g., event A could be dependent on event B, yet event B may not be dependent on event A; in math "dependent" is a symmetric relation; i.e., A is dependent with B if and only B is dependent with A.

51. **TRUE** False If A and B are independent events, their complements are also bound to be independent, and to prove this we need a general argument since an example is not sufficient here.

52. True **FALSE** If A and B are independent events, $\overline{A}$ and B may fail to be independent, but to prove this we need just one counterexample, not a general proof.

Solution: They are also independent.

53. True **FALSE** If any pair of events among $A_1, A_2, ..., A_n$ are independent, then all events are independent.

54. True **FALSE** A random variable (RV) on a probability space (Ω, P) is a function $X : \Omega \to \mathbb{R}$ that satisfies certain rules and is related to the probability function P.

Solution: A RV is **any** function to $\mathbb{R}$ and not at all related to the probability function P. The PMF of this RV is related to P though.

55. **TRUE** False A RV X could be the only source of data for an outcome space Ω and hence could be very useful in understanding better X 's domain.

Solution: One reason we introduced random variables is because it is sometimes hard to understand all of Ω and we can only look at it through X.

56. True **FALSE** The notation "$X \in E$" means that the RV X starts from event $E \subseteq \Omega$ and lands in $\mathbb{R}$.

Solution: The notation is for $E \subset \mathbb{R}$ and is used to denote the event the RV X lands in E.

57. True **FALSE** The notation "$X^{-1}(E)$" for a RV X and event $E \subseteq \Omega$ means to take set $B \subseteq \Omega$ of the reciprocals of all elements in E that are in the range of X .

Solution: The notation $X^{-1}(E) \subset \Omega$ is for $E \subset \mathbb{R}$ and denotes all outcomes that are in the preimage of E.

58. True **FALSE** The PMF of a RV X on probability space (Ω, P) is a third function $f : \mathbb{R} \to [0, 1]$ such that the composition of X followed by f on any $\omega \in \Omega$ is equal to P; i.e., such that $f(X(\omega)) = P(\omega)$.

Solution: First off, the probability function takes in subsets of Ω and hence $P(\omega)$ does not make sense. Even if we were to replace it with $P(\{\omega\})$, this would still potentially be wrong as seen in TF Question 9. If $x = X(\omega)$, by definition, we have $f(X(\omega)) = P(X^{-1}(x))$ and $X^{-1}(x)$ could contain more things than just ω.

59. True **FALSE** It is possible that $f(x) > P(x)$ for some $\omega \in \Omega$ and the corresponding $x = X(\omega) \in \mathbb{R}$ where X a discrete RV on (Ω, P) with PMF f .

Solution: If we replace $P(x)$ with $P(\{\omega\})$, then the result is true (note that $x \in \mathbb{R}$ and so $P(x)$ doesn't make sense because both $x \notin \Omega$ and it is a subset). We have $f(X(\omega)) = P(X = X(\omega))$ so the probability of all outcomes that are mapped to the same value as ω. For example, if we are counting the number of boys in a family, then $f(X(BBG)) = P(\{BBG, BGB, GBB\}) > P(\{BBG\})$.

60. True **FALSE** To show that two RV's $X, Y : \Omega \to \mathbb{R}$ are independent on (Ω, P) , we can find two subsets $E, F \subseteq \mathbb{R}$ for which $P(X \in E \text{ and } Y \in F) = P(X \in E) \cdot P(Y \in F)$.

Solution: To show that they are independent, we need to show that for all subset E, F, the equality holds. To show that they are not independent, we only need to find a single counterexample.

61. True **FALSE** Two Bernoulli trials are independent only if the probability of success and failure are each $\frac{1}{2}$.

Solution: Bernoulli trials can be independent regardless of what p is.

62. **TRUE** False The product X and the sum Y of the values of two flips of a fair coin (H=1, T=0) are dependent random variables.

Solution: If we know that $Y = 2$, then we know we got 2 heads and hence we know that $X = 1$, showing that they aren't independent.

63. **TRUE** False To turn the experiment of "rolling a die once" into a Bernoulli trial, we need to split its outcome space into two disjoint subsets and declare one of them a success.

64. **TRUE** False Several Bernoulli trials performed on one element at a time from a large outcome space Ω, without replacement, are approximately independent because what happens in one Bernoulli trial hardly affects the ratio of "successes" to "failures" in the remainder of the population.

65. True **FALSE** The probability of having 20 women within randomly selected 40 people is about 50%, assuming that there is an equal number of women and men on Earth.

Solution: This is a binomial distribution with the probability of choosing a woman as $p = \frac{1}{2}$. But the probability of getting exactly 20 women is $f(20) = \binom{40}{20}\frac{1}{2^{40}} \approx 12.5\%$.

66. **TRUE** False The hypergeometric distribution describes the probability of k "successes" in n random draws without replacement from a population of size N that contains exactly m "successful" objects.

67. True **FALSE** While the hypergeometric and binomial probabilities depend each on 3 parameters and 1 (input) variable, the Poisson probability depends only on 1 parameter and 1 (input) variable.

Solution: The binomial distribution depends on two variables (n and p)

68. True **FALSE** To approximate well the probability of k successes in a large number n of independent Bernoulli trials, each with individual probability of success p that is relatively small, we can use the formula $\frac{(np)^k e^{-np}}{n!}$.

Solution: The denominator should have $k!$ instead of $n!$.

69. **TRUE** False $E(X-Y) = E(X) - E(Y)$ for any R.V.s X and Y, regardless of whether they are independent or not.

70. True **FALSE** $Var(X-Y) = Var(X) - Var(Y)$ for any independent R.V.s X and Y.

Solution: We have that $Var(X-Y) = Var(X) + Var(-Y) = Var(X) + (-1)^2 Var(Y) = Var(X) + Var(Y)$ for independent X, Y.

71. **TRUE** False $Var(X) = E(X^2) - E^2(X)$ holds true because, essentially, the expected value has linearity properties.

Solution: By definition $Var(X) = E[(X - E[X])^2]$ and then we expand by using linearity to get the above result.

72. **TRUE** False Splitting a R.V. X as a sum of simpler R.V.'s X_i could be advantageous when we want to compute $Var(X)$, but we need to be careful that these X_i's are independent.

73. **TRUE** False If X is the geometric R.V., then $Var(X) = \frac{1-p}{p^2}$.

74. **TRUE** False If X is the hypergeometric R.V. in variable k and with parameters m, n, N, then $N = \frac{mn}{E(X)}$.

75. **TRUE** False If Santa Claus randomly throws n presents into n chimneys (one present per chimney), on the average one home will receive their intended present, regardless of how large or small n is.

Solution: This is similar to the hat example.

76. **TRUE** False For any independent R.V's X and Y, we have $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$, but $Var(XY) = Var(X) \cdot Var(Y)$ only if $E(X) = 0 = E(Y)$ or $X = 0$ or $Y = 0$ or both X and Y are constants.

> **Solution:** This comes from studying the equation $Var(XY) = Var(X)Var(Y) + E[X]^2Var(Y)+E[Y]^2Var(X)$ and asking when can $E[X]^2Var(Y) = E[Y]^2Var(X) = 0$.

77. **TRUE** False As long as several RV's $X_1, X_2, ..., X_n$ are identically distributed, their average RV $\overline{X}$ will have the same mean as each of them, but the standard error of $\overline{X}$ may or may not be equal to $\frac{SE(X_3)}{\sqrt{n}}$.

> **Solution:** The mean will always be the same but in order for the standard error to be equal to that, we need independence of the variables.

78. True **FALSE** The formula for the mean of the average $\overline{X} = \frac{X_1+X_2+\cdots+X_n}{n}$ of independent identically distributed RV's has $\sqrt{n}$ in the denominator.

> **Solution:** The mean does not have a $\sqrt{n}$ in the denominator but the standard error does.

79. True **FALSE** The formula for the variance $Var(X+Y) = Var(X) + Var(Y)$ works regardless of whether the RV's X and Y are independent or not.

80. **TRUE** False To approximate the height of a tall tree (using similar triangles and measurements along the ground – without climbing the tree!) it is better to ask several people to do it independently of each other and then to average their results, than to do it once just by yourself.

> **Solution:** As n gets higher, the standard error of $\bar{X}$ gets smaller so your error becomes less.

81. True **FALSE** If $\overline{X}$ is the average of n IIDRVs, each with mean μ and standard error σ, then for n large the normalized distribution $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ has a very small standard error.

> **Solution:** The standard error of the normalized distribution is always 1.

82. True **FALSE** According to the Central Limit Theorem, the normalized distribution $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ is the standard normal distribution for n large, where $\overline{X}$ is the average of n independent, identically distributed variables, each with mean μ and standard error $\sigma/\sqrt{n}$.

Solution: Each of the X_i has standard error σ, not $\sigma/\sqrt{n}$.

83. True **FALSE** As n grows the average $\overline{X}$ of IIDRV's $X_1, X_2, X_3..., X_n$ becomes more wide-spread, since we are taking larger and larger samples and incorporating more data into our calculations.

Solution: The spread becomes thinner and thinner because the standard error is $\bar{\sigma} = \sigma/\sqrt{n}$ so as $n \to \infty$, we have $\bar{\sigma} \to 0$.

84. True **FALSE** The Central Limit Theorem states that for large n, the normalized distribution $Z = \frac{\bar{X}-\mu}{\bar{\sigma}}$ is the standard normal distribution.

Solution: The CLT says that Z becomes closer and closer to one but it is not exactly a normal distribution.

85. True **FALSE** Histograms are always defined over intervals of the form $[a, b]$ with $a \geq 0$ because probabilities are always non-negative.

Solution: The height of the histogram is always ≥ 0 but the domain doesn't have to be.

86. **TRUE** False The height of a rectangle in a histogram equals $\frac{\text{the amount of the data corresponding to the subinterval}}{\text{the width of the subinterval}}$, because all rectangular areas in a histogram must sum up to 1.

87. True **FALSE** $P(x) = e^{-x^2}$ is a PDF on $\mathbb{R}$.

Solution: The area underneath the curve is $\sqrt{\pi}$ so it is not a PDF. We need to normalize it by $\frac{1}{\sqrt{\pi}}$ to make it one.

88. **TRUE** False We allow for a PDF to occasionally "peak" above 1 (e.g., $f(x) > 1$ for some $x \in \mathbb{R}$), but a PMF is forbidden to do that!

89. True **FALSE** A PDF $f(x)$ can have values above 1, but this can happen only at finitely many places $x_1, x_2, \ldots, x_n \in \mathbb{R}$

Solution: A PDF can be greater than 1 at infinitely many places (and often is).

90. True **FALSE** Uniform PDFs are defined by $f(x) = c$ for all $x \in \mathbb{R}$.

Solution: All uniform PDFs must be finite on a domain like $[a, b]$.

91. True **FALSE** Shifting the bell-shaped PDF $f(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}$ to the left by 5 units results in another PDF $g(x) = \frac{1}{\sqrt{\pi}}e^{-(x+5)^2}$ centered at $x = 5$.

Solution: Shifting to the left results in a PDF centered at $x = -5$.

92. True **FALSE** For a PDF to be centered at $x = a$, it means that a is the median of the CDF.

Solution: For it to be centered, it needs to be symmetric as well.

93. **TRUE** False CDFs behave in general like antiderivaties of their PDFs, but there are some situations where the CDF is not continuous and hence there is no way it can be an antiderivative of a PDF.

Solution: CDFs could arise from PMFs and hence they are not antiderivatives of a PDF.

94. **TRUE** False The formula $P(a \leq X \leq b) = F(b) - F(a)$ for a CDF $F(x)$ works because $F(x)$ can be essentially thoughout of as the "area-so-far" function for a PDF.

95. True **FALSE** To prove that a function $F(x)$ on $\mathbb{R}$ is a CDF, we need only to confirm that it is non-decreasing on $\mathbb{R}$, attains only non-negative values, and that its value tends to 1 and 0 as $x \to \infty$ and $x \to -\infty$ correspondingly.

Solution: We also need to show that it is right continuous.

96. **TRUE** False The CDF of the bell-shaped PDF $f(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}$ has a graph that is increasing and looks like a solution to the classic logistic model with carrying capacity $K = 1$, but it cannot be exactly equal to it because there is no elementary function of an antiderivative of e^{-x^2} yet we have derived the formula $g(x) = \frac{1}{1+Ae^{-kt}}$ for these logistic model solutions.

> **Solution:** The derivative of the CDF will yield the PDF and so $g(x)$ cannot possibly be an antiderivative of $f(x)$.

97. **TRUE** False There are at least three ways to compute the probability of a person to be born in May: using a discrete random variable, and using the PDF or the CDF of a continuous random variable.

98. **TRUE** False The formula for the mean of a continuous random variable is a limit version of the mean for a discrete random variable; but while the latter always exists for a finite amount of data, the former may not exist for certain continuous random variables.

> **Solution:** The former may diverge.

99. **TRUE** False If the mean is larger than the median, the distribution tends to be more spread away on the right and more clustered together on the left.

100. True **FALSE** If we make the area between a PDF and the x-axis out of uniform cardboard material and make an infinite seesaw out of the x-axis, the point on the x-axis where the seesaw will balance is the median of the distribution because there is an equal material to the left and to the right of the median.

> **Solution:** The balancing point is the mean.

101. **TRUE** False Exam distributions of large classes tend to have smaller means than medians when the medians are higher than 50% of the maximum possible score.

102. True **FALSE** The Pareto distribution fails to have a well-defined mean when the constant $a \geq 2$.

> **Solution:** It fails to have a well defined mean when $a \leq 2$.

103. True **FALSE** Improper integrals resurface when we want to compute probabilities of discrete random variables with finitely many values.

Solution: The resurface when we want to calculate the mean and standard deviation of continuous random variables.

104. True **FALSE** For a symmetric distribution, we do not have to calculate the mean because it will always equal the median.

Solution: The mean may fail to exist, but if it does, it will equal the median.

105. **TRUE** False $\frac{1}{\sqrt{2\pi}}$ ensures that the formula for the normal distribution indeed represents a valid PDF.

Solution: It normalizes the area to 1.

106. True **FALSE** z scores are not suitable for computing probabilities of the type $P(-\infty \leq X \leq a)$ or $P(b \leq X \leq \mu)$ for arbitrary normal distributions.

Solution: Because the normal distribution is symmetric around μ, we can flip these to the positive side to and do the calculation as before.

107. True **FALSE** Normal distributions are defined only for positive X; yet, when converted to the standard normal distribution, they may be defined for negative X too.

Solution: Normal distributions are defined for all X.

108. **TRUE** False PMFs replace PDFs when moving from continuous to discrete random variables.

Solution: PMFs are the discrete version of a PDF.

109. True **FALSE** CDFs can be defined by the same probability formula for both discrete and continuous variables; however, at the next step when actually computing the CDFs, one must be careful to use correspondingly PMFs with integrals and PDFs with appropriate summations.

Solution: PMFs are with a summation and PDFs are with integrals.

110. True **FALSE** The target spaces for PDFs, PMFs, and probability functions are all the same.

Solution: The target space for PDFs and PMFs are $[0, \infty)$ and for probability functions $[0, 1]$.

111. True **FALSE** The domains of PDFs and PMFs are the corresponding outcome spaces Ω.

Solution: The domains are $\mathbb{R}$.