PRACTICE PROBLEMS FOR THE FINAL

Math 3160Q – Spring 2015
Professor Hohn

Below is a list of practice questions for the Final Exam. I would suggest also going over the practice problems and exams for Exam 1 and Exam 2 to help you prepare. Any quiz, homework, or example problem has a chance of being on the exam. For more practice, I suggest you work through the review questions at the end of each chapter as well.

1. The amount of time a customer spends at a certain store is modeled by an exponential random variable with mean 10 minutes. If each customer's time at the store is independent, use the Central Limit Theorem to approximate the probability that 100 randomly selected customers spend between 950 and 1050 minutes at the store. Leave your answer in terms of the standard normal distribution $\Phi(x)$.

Solution: Let X_i be the time the ith customer spent in the store; we are given that $X_i \stackrel{d}{=} \text{Exp}(\lambda)$ where $\lambda = 1/10$ (since $\mathbb{E}[X_i] = 10 = 1/\lambda$). Let $Z \stackrel{D}{=} N(0,1)$. Then $X_1, X_2, ..., X_{100}$ are i.i.d. random variables and the central limit theorem tells us

$$\sum_{i=1}^{100} X_i \approx \sqrt{n \cdot \sigma^2}\, Z + n\mu = \sqrt{100(10^2)}\, Z + 100 \cdot 10 = 100Z + 1000$$

where we used the fact that $\sigma^2 = \text{Var}(X_i) = \left(\frac{1}{\lambda}\right)^2 = 10^2$ and $\mu = \mathbb{E}[X_i] = \frac{1}{\lambda} = 10$. We are looking for $P\big(950 \le \sum_{i=1}^{100} X_i \le 1050\big)$, so by the Central Limit Theorem

$$\begin{aligned} P\Big(950 \le \sum_{i=1}^{100} X_i \le 1050\Big) &\approx P\big(950 \le 100\,Z + 1000 \le 1050\big) = P\Big(-\frac{1}{2} \le Z \le \frac{1}{2}\Big) \\ &= \Phi\Big(\frac{1}{2}\Big) - \Phi\Big(-\frac{1}{2}\Big) = 2\Phi\Big(\frac{1}{2}\Big) - 1 \end{aligned}$$

2. Let X and Y be jointly continuous random variables with joint density function

$$f_{X,Y}(x,y) = \begin{cases} kx^2y & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find k.

Solution:

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = \int_{0}^{1}\int_{0}^{y} kx^2y\,dxdy = k\int_0^1 y\Big(\frac{1}{3}x^3\big|_0^y\Big)\,dy$$

$$= \frac{k}{3}\int_0^1 y^4\,dy = \frac{k}{15}y^5\big|_0^1 = \frac{k}{15}.$$

Solving for k, we get $k = 15$.

(b) Find $\mathrm{Cov}(X,Y)$.

Solution: We know that $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$. We have

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)\,dxdy = 15\int_0^1\int_0^y x^3y^2\,dxdy = \frac{15}{4}\int_0^1 y^6\,dy = \frac{15}{28}.$$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf_{X,Y}(x,y)\,dxdy = 15\int_0^1\int_0^y x^3y\,dxdy = \frac{15}{4}\int_0^1 y^5\,dy = \frac{15}{24}.$$

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf_{X,Y}(x,y)\,dxdy = 15\int_0^1\int_0^y x^2y^2\,dxdy = \frac{15}{3}\int_0^1 y^5\,dy = \frac{15}{18}.$$

Therefore $\mathrm{Cov}(X,Y) = \frac{15}{28} - \left(\frac{15}{24}\right)\left(\frac{15}{18}\right)$.

3. This problem shows that not all functions can be moment generating functions, even if they're very differentiable near zero! Let $f(t) = e^t + t$.

(a) Pretend that $f(t)$ is the moment generating function for some random variable X and use it to find $\mathrm{Var}(X)$.

Solution: If $f(t) = M_X(t)$, then $f'(0) = \mathbb{E}[X]$ and $f''(0) = \mathbb{E}[X^2]$. So, $\mathrm{Var}(X) = f''(0) - f'(0)^2$. We have that $f''(t) = e^t$ and $f'(t) = e^t + 1$. Therefore

$$\mathrm{Var}(X) = f''(0) - f'(0)^2 = e^0 - (e^0+1)^2 = 1 - (1+1)^2 = -3.$$

(b) Using the value of $\mathrm{Var}(X)$ from (a), how do you know that no such X can exist?

Solution: The variance of any random variable is non-negative. One way to convince yourself this is true is to remember that one way to define $\mathrm{Var}(X)$ is $\mathrm{Var}(X) = \mathbb{E}[(X-\mu)^2]$ where $\mu = \mathbb{E}[X]$. Since $(X-\mu)^2 \geq 0$, then $\mathrm{Var}(X) \geq 0$. We found that if $f(t)$ was the moment generating function then the variance is -3, which can't happen.

4. Suppose the number of meteors hitting earth is modeled by a Poisson process $(N_t)_{t\geq 0}$ with parameter $\lambda = 5$ (meteors / hour) where t is in hours.

(a) What is the probability that during the 1 hour you are in this class, 2 or more meteors hit the earth? I don't want you to leave me with an infinite sum, instead it should be something that you could calculate easily *if* you were allowed a calculator.

Solution: Let's suppose that you enter the class at time t_0. The number of meteors hitting the earth during the one hour you are in class is given by $N_{t_0+1} - N_{t_0}$ and we know that $N_{t_0+1} - N_{t_0} \stackrel{d}{=} \text{Pois}((t_0 + 1 - t_0)\lambda) = \text{Pois}(\lambda) = \text{Pois}(5)$. Therefore,

$$\begin{aligned} P(N_{t_0+1} - N_{t_0} \geq 2) = P(\text{Pois}(5) \geq 2) = 1 - P(\text{Pois}(5) = 0) - P(\text{Pois}(5) = 1) \\ = 1 - e^{-5} - 5e^{-5} = 1 - 6e^{-5}. \end{aligned}$$

(b) What is the conditional probability that during the 1 hour you are in this class, 2 or more meteors hit the earth given that 1 or more meteors hit the earth?

Solution: What we are looking for is $P(N_{t_0+1} - N_{t_0} \geq 2 \mid N_{t_0+1} - N_{t_0} \geq 1)$. Using the definition of conditional probability,

$$\begin{aligned} P(N_{t_0+1} - N_{t_0} \geq 2 \mid N_{t_0+1} - N_{t_0} \geq 1) &= \frac{P(N_{t_0+1} - N_{t_0} \geq 2,\, N_{t_0+1} - N_{t_0} \geq 1)}{P(N_{t_0+1} - N_{t_0} \geq 1)} \\ &= \frac{P(N_{t_0+1} - N_{t_0} \geq 2)}{P(N_{t_0+1} - N_{t_0} \geq 1)} = \frac{P(\text{Pois}(5) \geq 2)}{P(\text{Pois}(5) \geq 1)} = \frac{1 - 6e^{-5}}{1 - e^{-5}}. \end{aligned}$$

5. Suppose X is a continuous random variable with distribution function given by

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases}$$

(a) Find the density $f_X(x)$ of X.

Solution:

$$f_X(x) = F'(x) = \begin{cases} 0 & x < 0 \\ 2x & 0 \leq x < 1 \\ 0 & 1 \leq x. \end{cases}$$

(b) Find $\text{Var}(X)$.

Solution: $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. We have

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_0^1 x^2(2x)\, dx = 2\int_0^1 x^3\, dx = \frac{1}{2} \\ \mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^1 x(2x)\, dx = 2\int_0^1 x^2\, dx = \frac{2}{3}. \end{aligned}$$

So $\text{Var}(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$.

6. The following parts (a) and (b) refer to the letters MMMMPPIII.

(a) How many words can be created such that both P's occur to the left of the first I?

Solution: We have 9 total letters; we will want to arrange these letters into 9 vacancies. Imagine that we take out the Ps and Is and place the four Ms in the 9 vacancies first; we have $\binom{9}{4}$ ways of doing this. Once this is done we have 5 remaining vacancies for the Ps and Is to occupy, but since both Ps must come before the first I, there is only one way to fill the 5 remaining vacancies: the two Ps are in the first 2 and the three Is are in the last 3. Therefore, once we place the Ms we're done, and as we have mentioned, there are $\binom{9}{4}$ ways of doing this.

(b) Suppose each word is equally likely to occur. What is the probability that you randomly select a word where both P's occur to the left of the first I?

Solution: From the previous part, we have $\binom{9}{4}$ arrangements with both Ps before the first I, and using the word counting principle, $\binom{9}{4,2,3}$ distinguishable arrangements of MMMMPPIII. Since each outcome is equally likely, the probability is

$$\frac{\binom{9}{4}}{\binom{9}{4,2,3}}.$$

7. Suppose that there is a disease within a population. Each individual within the population independently and randomly has the disease with probability p. A certain test is created to screen individuals for the disease. If an individual has the disease, there is a probability q that the test correctly reports positive. If an individual does not have the disease, there is a probability r that the test falsely reports positive for the disease. If a randomly selected individual tested positive for the disease, what is the probability that they have the disease? You should assume that the test only reports a positive or negative result, and your answer should be in terms of p, q, and r.

Solution: Let A be the event that you have the disease and T be the event that you test positive for the disease. We want $P(A \mid T)$. We are told $P(A) = p$, $P(T \mid A) = q$, and $P(T \mid A^c) = r$. Using Bayes' formula,

$$P(A \mid T) = \frac{P(T \mid A)P(A)}{P(T)} = \frac{P(T \mid A)P(A)}{P(T \mid A)P(A) + P(T \mid A^c)P(A^c)} = \frac{qp}{qp + r(1-p)}.$$

8. Your "friend" wants to play a gambling game with you. The game goes as follows: You pay your friend \$$C$ to play. You then draw twice from an urn initially containing 4 red marbles and 5 black marbles. On the first draw you take a single marble and return it back into the urn with another marble of the same color. On the second draw you pull out a single marble: If it is a red marble you win \$5, whereas if it is a black marble you win nothing.

Question: What is the maximum amount of money \$$C$ you should pay your friend to play so that on average you don't lose money?

Solution: Let X represent your wealth (in dollars) during a play of the game. Then the possible values of X are $-C$ corresponding to pulling a black marble on the second draw, and $-C + 5$ corresponding to pulling a red marble on the second draw. We want to find the maximum C so that $\mathbb{E}[X] \geq 0$ (ie. we don't loose money). To do so, let R be the event that the first draw was red. Then

$$\begin{aligned} P(X = -C) &= P(X = -C|R)P(R) + P(X = -C|R^c)P(R^c) \\ &= \frac{5}{10} \cdot \frac{4}{9} + \frac{6}{10} \cdot \frac{5}{9} = \frac{5}{9} \end{aligned}$$

and

$$\begin{aligned} P(X = -C + 5) &= P(X = -C + 5|R)P(R) + P(X = -C + 5|R^c)P(R^c) \\ &= \frac{5}{10} \cdot \frac{4}{9} + \frac{4}{10} \cdot \frac{5}{9} = \frac{4}{9} \end{aligned}$$

We can then find the expected value of X,

$$\begin{aligned} \mathbb{E}[X] &= -C \cdot P(X = -C) + (-C + 5) \cdot P(X = -C + 5) \\ &= -C \cdot \frac{5}{9} + (-C + 5) \cdot \frac{4}{9} \\ &= -C + \frac{20}{9} \end{aligned}$$

Therefore, the maximum value we can pay to play such that we don't loose money is

$$\boxed{\$C = \$\left(\frac{20}{9}\right) \approx \$2.22}$$

9. You must select exactly one of two challenges: A or B. If you select challenge A you are forced to answer a question which you have a 1/2 probability of getting correct. If you answer correctly you will be rewarded silver. If you answer incorrectly, you will be forced to eat spoiled cheese. If you select challenge B, you will be forced to complete a physical task which you have 1/3 chance of completing. If you succeed you will rewarded gold. If you do not succeed you will be forced to eat spoiled cheese. You decide to flip a biased coin with probability 2/3 of getting heads to decide which challenge to take: heads means you take A, tails means you take B. Given that you are forced to eat spoiled cheese, what is the probability that you selected challenge A?

Solution: Let E_A be the event that you select challenge A, and C be the event that you eat spoiled cheese. Then what we want to find is $P(E_A|C)$. We will now use Bayes' Theorem. We are given $P(E_A) = 2/3$, $P(E_A^c) = 1 - P(E_A) = 1/3$, $P(C|E_A) = 1/2$, and $P(C|E_A^c) = 2/3$. Therefore,

$$\begin{aligned} P(E_A|C) = \frac{P(C|E_A)P(E_A)}{P(C)} &= \frac{P(C|E_A)P(E_A)}{P(C|E_A)P(E_A) + P(C|E_A^c)P(E_A^c)} \\ &= \frac{(1/2)(2/3)}{(1/2)(2/3) + (2/3)(1/3)} = \frac{3}{5} \end{aligned}$$

10. Each year, the cryptids known only as "tubes" independently kill Americans with an average

kill rate of 3 Americans per year.

(a) Using a Poisson distribution, approximate the probability that 2 or fewer Americans die next year from "tubes."

Solution: If X is the number of people killed by tubes next year, then we will assume $X \approx \text{Pois}(3)$. Therefore,

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = e^{-3} + 3e^{-3} + \frac{3^2}{2}e^{-3} = \frac{17}{2}e^{-3}.$$

(b) Briefly explain why a Poisson approximation here is a reasonable choice and why one might use it over the Central Limit Theorem (the normal approximation)?

Solution: If the population has size N, which we know is quite large, and p is the probability that a randomly selected person will die of tubes in this population. X acts like a $\text{Bin}(N, p)$ random variable with $Np = 3$. Since N is large and $Np \ll N$, the Poisson approximation is reasonable. Moreover, to answer the previous part, we are only needed to calculate a sum of few terms (i.e., we only needed to sum 3 terms in the previous part) making the Poisson approximation an attractive choice.

If, on the other hand, we wanted to approximate X by a normal distribution we would use

$$\frac{X - Np}{\sqrt{Np(1-p)}} \overset{d}{\approx} N(0, 1)$$

However, $Np(1 - p) < Np \ll N$. When using the normal approximation we hope that $Np(1 - p)$ is reasonably large (relative to N), which doesn't appear to be the case here.

11. Let X and Y be jointly continuous random variables with joint density

$$f_{X,Y}(x, y) = \begin{cases} c\,(1-y) & 0 < x < y,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find c.

Solution:

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)dxdy = c\int_0^1\int_0^y (1-y)dxdy = c\int_0^1 (y - y^2)dy = c(\frac{1}{2} - \frac{1}{3}) = \frac{c}{6}$$

Therefore $c = 6$.

(b) Given that $Y > 2/3$, what is the probability that $X < 1/2$?

Solution:

$$P(X < 1/2 | Y > 2/3) = \frac{P(X < 1/2, Y > 2/3)}{P(Y > 2/3)}.$$

Notice that if $X < 1/2$ and $Y > 2/3$, then we are immediately guaranteed that $X < Y$, so the bounds for $\{X < 1/2, Y > 2/3\}$ will look like $0 < X < 1/2$ and $2/3 < Y < 1$.

$$P(X < 1/2, Y > 2/3) = 6\int_{2/3}^1\int_0^{1/2} (1-y)dxdy = 3\int_{2/3}^1 (1-y)dy = \frac{1}{6}$$

We also have,

$$P(Y > 2/3) = 6 \int_{2/3}^{1} \int_{0}^{y} (1-y)dxdy = 6 \int_{2/3}^{1} (y - y^2)dy = \frac{7}{27}$$

Therefore,

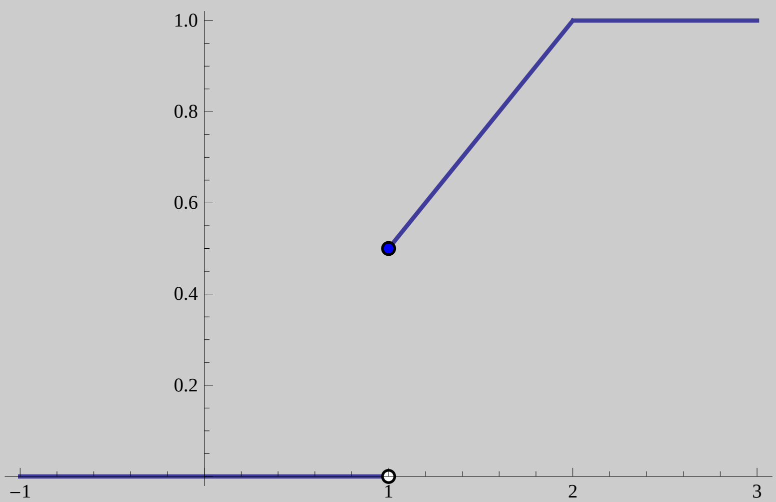
$$P(X < 1/2 | Y > 2/3) = \frac{1/6}{7/27} = \frac{9}{14}$$

12. Let X be a random variable with distribution

$$F_X(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{2}x & 1 \le x < 2 \\ 1 & 2 \le x \end{cases}$$

(a) Is X continuous, discrete, or neither? Remember to give me a brief justification.

Solution: Neither. The jump discontinuity at $x = 1$ prevents $F_X(x)$ from being a continuous function which in turn prevents X from being a continuous random variable. However, the positive sloped line of $F_X(x)$ on the bounds $1 \le x < 2$ prevents X from being discrete.



(b) Find $\mathbb{E}[X]$.

Solution:

$$\mathbb{E}[X] = \int_0^{\infty} (1 - F_X(x))dx - \overbrace{\int_{-\infty}^{0} F_X(x)dx}^{=0} = \int_0^1 1dx + \int_1^2 \left[1 - (1/2)x\right]dx$$
$$= 1 + \left[x - (1/4)x^2\right]_{x=1}^{2} = \frac{5}{4}$$

13. You roll two fair six-sided dice. Let X be the minimum value of the two rolls showing. For

example, if you roll a 3 and a 4, then X is 3, since $\min(3,4) = 3$. Similarly, if you roll a 5 and a 5, then X is 5, since $\min(5,5) = 5$.

(a) What is a reasonable sample space Ω for this experiment?

Solution: The most reasonable sample space for the experiment that I can think of is

$$\Omega = \{(1,1),(1,2),...,(6,6)\} = \{(i,j) : 1 \le i \le 6, 1 \le j \le 6\}$$

where each element (i,j) represents the first die rolling i and the second rolling j.

(b) What is the state space S_X of X?

Solution: Since any integer between 1 and 6 can be a value for X, we have

$$S_X = \{1,2,3,4,5,6\}.$$

(c) Find the probability mass function p_X of X.

Solution: If $k \in S_X$, then $p_X(k) = P(X = k)$ is the probability that one of the die lands on k and the other lands on a number $\ge k$. Therefore we have $(k,k),(k,k+1),...,(k,6),(k+1,k),...,(6,k)$ as the possible outcomes which make $X = k$. If we count these, there are $2(6-k)$ elements with one entry being k and the other being something bigger, and 1 element with both entries k. So, $\#\{X = k\} = 2(6-k)+1 = 13-2k$. Therefore, with 36 possible outcomes, $p_X(k) = (13-2k)/36$ for each $k \in S_X$.

Alternatively, we could have just listed each case out and counted:

$$p_X(1) = 11/36,\ p_X(2) = 9/36,$$
$$p_X(3) = 7/36,\ p_X(4) = 5/26,$$
$$p_X(5) = 3/36,\ p_X(6) = 1/36.$$

14. Carbon-14 has a half-life of approximately 5730 years. Organic life continually replenishes its stock of carbon-14 until death at which point the carbon-14 decays without replenishing. An organism with about 6000 carbon-14 particles just died. Let N_t count the number of decayed carbon-14 particles after a time t of the organism's death (where t is measured in years). Based on the information just described, it is reasonable to model $(N_t)_{t\ge 0}$ as a Poisson process with parameter $\lambda = .725$.

(a) Assuming that $(N_t)_{t\ge 0}$ is a Poisson process with parameter $\lambda = .725$, what is the probability that two or more carbon-14 particles decay during one year following the death of the organism? Please don't leave your answer as an infinite sum!

Solution: We want $P(N_1 \ge 2) = 1 - P(N_1 = 0) - P(N_1 = 1)$. Since $N(1) = \text{Pois}(.725)$,

$$P(N_1 \ge 2) = \underbrace{1 - e^{-.725} - (.725)e^{-.725}}_{\text{simplified enough}} = \boxed{1 - 1.725e^{-.725}}$$

(b) Assuming that $(N_t)_{t\ge 0}$ is a Poisson process with parameter $\lambda = .725$, what is the probability that no carbon-14 particles decay during the first year following the organism's death, *and* two carbon-14 particles decay during the following two years?

Solution: This time we want $P(N_1 = 0, N_3 - N_1 = 2) = P(N_1 = 0)P(N_3 - N_1 = 2)$ since the number of particles we count during non-overlapping periods of time are independent (by assumption that N_t is a Poisson process). Now, $N_1 = \text{Pois}(.725)$ and $N_3 - N_1 = \text{Pois}((3-1)(.725)) = \text{Pois}(1.45)$. Therefore,

$$P(N_1 = 0, N_3 - N_1 = 2) = P(N_1 = 0)P(N_3 - N_1 = 2)$$
$$= \underbrace{\left(e^{-.725}\right)\left(\frac{(1.45)^2}{2!}e^{-1.45}\right)}_{\text{simplified enough}} = \boxed{\frac{(1.45)^2}{2}e^{-2.175}}$$

15. Suppose that the moment generating function for a random variable X is given by

$$M_X(t) = \frac{1}{(1-t)^2}$$

(a) Find a formula for the general n^{th} moment of X. That is, in terms of n, find $\mathbb{E}[X^n]$.

Solution: We know that $\mathbb{E}[X^n] = \frac{d^n}{dt^n} M_X(t)\big|_{t=0}$. So let's write out the first few derivatives of $M_X(t)$:

$$M_X^{(1)}(t) = \frac{2}{(1-t)^3},\ M_X(t)^{(2)} = \frac{3\cdot 2}{(1-t)^4},\ M_X(t)^{(3)} = \frac{4\cdot 3\cdot 2}{(1-t)^5}$$

and we see that the following pattern emerges: $M_X^{(n)}(t) = \frac{(n+1)!}{(1-t)^{n+2}}$. Therefore,

$$\mathbb{E}[X^n] = M_X^{(n)}(0) = \frac{(n+1)!}{(1-0)^{n+2}} = (n+1)!$$

(b) Find $\text{Var}(X)$.

Solution: From the previous part we have $\mathbb{E}[X^n] = (n+1)!$. So $\mathbb{E}[X^2] = (2+1)! = 3! = 6$, and $\mathbb{E}[X] = (1+1)! = 2$. Therefore $\text{Var}(X) = \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2 = 6 - (2)^2 = 2$.

You could have alternatively found $\text{Var}(X)$ by $\text{Var}(X) = M_X^{(2)}(0) - \big(M_X^{(1)}(0)\big)^2$. This will of course lead to the same answer.

16. Suppose that $\{X_i\}_{i=1}^{\infty}$ are i.i.d. random variables with mean $\mu < \infty$ and variance $\sigma^2 < \infty$. Fix some $\varepsilon > 0$. Use the Central Limit Theorem to approximate

$$P\Big(-\varepsilon < \frac{1}{n}\sum_{i=1}^{n} X_i - \mu < \varepsilon\Big).$$

Leave your answer in terms of the standard normal distribution $\Phi(a) = P(N(0,1) \leq a)$. Your answer should depend on ε, n, and σ^2.

Hint: $\frac{1}{n}\sum_{i=1}^{n} X_i - \mu = \frac{1}{n}\left(\sum_{i=1}^{n} X_i - n\mu\right)$.

Solution: We want to manipulate $\frac{1}{n}\left(\sum_{i=1}^{n} X_i - n\mu\right)$ to look like $\frac{1}{\sqrt{n\sigma^2}}\left(\sum_{i=1}^{n} X_i - n\mu\right) \approx N(0,1)$. We therefore need to multiply the inequality by $\sqrt{(n/\sigma^2)}$. We have

$$\begin{aligned} P\Big(-\varepsilon < \frac{1}{n}\big(\sum_{i=1}^{n} X_i - n\mu\big) < \varepsilon\Big) &= P\Big(-\varepsilon\sqrt{\frac{n}{\sigma^2}} < \frac{1}{n}\big(\sum_{i=1}^{n} X_i - n\mu\big)\sqrt{\frac{n}{\sigma^2}} < \varepsilon\sqrt{\frac{n}{\sigma^2}}\Big) \\ &= P\Big(-\varepsilon\sqrt{\frac{n}{\sigma^2}} < \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} < \varepsilon\sqrt{\frac{n}{\sigma^2}}\Big) \approx P\Big(-\varepsilon\sqrt{\frac{n}{\sigma^2}} < N(0,1) < \varepsilon\sqrt{\frac{n}{\sigma^2}}\Big) \\ &= \Phi\Big(\varepsilon\sqrt{\frac{n}{\sigma^2}}\Big) - \Phi\Big(-\varepsilon\sqrt{\frac{n}{\sigma^2}}\Big) = 2\Phi\Big(\varepsilon\sqrt{\frac{n}{\sigma^2}}\Big) - 1 \end{aligned}$$

17. Let X be a random variable with cumulative distribution function

$$F_X(s) = \begin{cases} 0 & s < 1 \\ \frac{1}{2}s & 1 \leq s < 2 \\ 1 & 2 \leq s \end{cases}$$

(a) Let Y be the random variable defined by $Y = \ln(X)$ (here, ln is the natural log). Find the cumulative distribution function F_Y of Y.

Solution:

$$F_Y(a) = P(Y \leq a) = P(\ln(X) \leq a) = P(X \leq e^a)$$

Notice that we have the CDF of X. That is, we know $P(X \leq s)$. Then, if we plug in $s = e^a$ in our $F_X(s)$, we have

$$F_Y(a) = F_X(e^a) = \begin{cases} 0 & e^a < 1 \\ \frac{1}{2}e^a & 1 \leq e^a < 2 \\ 1 & 2 \leq e^a \end{cases}$$

and, with some clean up, we have

$$F_Y(a) = \begin{cases} 0 & a < 0 \\ \frac{1}{2}e^a & 0 \leq a < \ln(2) \\ 1 & \ln(2) \leq a \end{cases}$$

(b) Find $\mathbb{E}[\ln(X)]$.

Solution:

$$
\begin{aligned}
\mathbb{E}[\ln(X)] = \mathbb{E}[Y] &= \int_0^\infty (1 - F_Y(y))\, dy - \int_{-\infty}^0 F_Y(y)\, dy \\
&= \int_0^{\ln(2)} \left(1 - \frac{1}{2}e^y\right) dy + \int_{\ln(2)}^\infty (1-1)\, dy - \int_{-\infty}^0 0\, dy \\
&= \left[y - \frac{1}{2}e^y\right]_0^{\ln(2)} + 0 - 0 \\
&= \ln(2) - \frac{1}{2}e^{\ln(2)} - \left(0 - \frac{1}{2}e^0\right) \\
&= \ln(2) - 1 + \frac{1}{2} \\
&= \ln(2) - \frac{1}{2}
\end{aligned}
$$

18. Let X and Y be jointly discrete random variable with probability mass function $p_{X,Y}(s,t)$ described by the following table

X \ Y	0	1	2
1	$\frac{1}{2}$	0	$\frac{1}{4}$
4	0	$\frac{1}{8}$	0
9	$\frac{1}{16}$	0	$\frac{1}{16}$

(a) What is the state space S_X of X and the state space S_Y of Y?

Solution:

$$S_X = \{1, 4, 9\}$$

$$S_Y = \{0, 1, 2\}$$

(b) Calculate $\mathbb{E}[\sqrt{X}\, Y]$.

Solution:

$$
\begin{aligned}
\mathbb{E}[\sqrt{X}Y] &= \sqrt{1}\cdot 0\, P(X=1, Y=0) + \sqrt{1}\cdot 1 P(X=1, Y=1) + \sqrt{1}\cdot 2P(X=1, Y=2) \\
&\quad + \sqrt{4}\cdot 0P(X=4, Y=0) + \sqrt{4}\cdot 1P(X=4, Y=1) + \sqrt{4}\cdot 2P(X=4, Y=2) \\
&\quad + \sqrt{9}\cdot 0P(X=9, Y=0) + \sqrt{9}\cdot 1P(X=9, Y=1) + \sqrt{9}\cdot 2P(X=9, Y=2) \\
&= 0 + 0 + 2\cdot\frac{1}{4} + 0 + 2\cdot\frac{1}{8} + 0 + 0 + 6\cdot\frac{1}{16} \\
&= \frac{1}{2} + \frac{1}{4} + \frac{6}{16} \\
&= \frac{18}{16}
\end{aligned}
$$

(c) Are X and Y independent? Make sure to justify your answer.

Solution: For X and Y to be independent, we need to show that $P(X = i, Y = j) = P(X = i)P(Y = j)$ for all $i \in S_X$ and $j \in S_Y$. Let's start with $X = 1, Y = 0$ (the first entry of the table) and see if $P(X = 1, Y = 0) = P(X = 1)P(Y = 0)$.

$$P(X = 1) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

Similarly,

$$P(Y = 0) = \frac{1}{2} + \frac{1}{16} = \frac{9}{16}$$

Now, we check

$$P(X = 1)P(Y = 0) = \left(\frac{3}{4}\right)\left(\frac{9}{16}\right) = \frac{27}{64} \neq \frac{1}{2} = P(X = 1, Y = 0)$$

Since $P(X = 1, Y = 0) \neq P(X = 1)P(Y = 0)$, X and Y are not independent.

19. The microorganism E.coli living in ***your*** body has the average lifespan of 10 hours. Right at this moment you, the E.coli whisperer, befriend one of these living E.coli.

(a) If the lifespan of the E.coli is given by an exponential random variable, what is the probability that your E.coli friend is still alive in five hours from now? If you are unable to answer this question, make sure to give a good and brief explanation why you can't answer it and what more information you would need to be be able to answer.

Solution: Let $X \stackrel{d}{=} \text{Exp}(\lambda)$ represent the lifetime of the E.coli. Note that since $10 = \mathbb{E}[X] = 1/\lambda$, we have $\lambda = 1/10$. If t_0 is the time the E.Coli has already been alive until now, we want to know $P(X > 5 + t_0 \mid X > t_0)$, and by the memoryless property of exponential random variables

$$P(X > 5 + t_0 \mid X > t_0) = P(X > 5) = e^{-5/10} = e^{-1/2}.$$

(b) If the lifespan of the E.coli is uniformly distributed between 0 and 20 hours (i.e., has a Unif$(0, 20)$ distribution), what is the probability that your E.coli friend is still alive in five hours from now? If you are unable to answer this question, make sure to give a good and brief explanation why you can't answer it and what more information you would need to be be able to answer.

Solution: Let $X \stackrel{d}{=} \text{Unif}(0, 20)$ represent the lifetime of the E.coli since its birth (Are they born? More like binary fission...). Note that the ensure that $\mathbb{E}[X] = 10$, the midpoint. If t_0 is the time the E.coli has already been alive until now, we want to know $P(X > 5 + t_0 \mid X > t_0)$. We don't have the memoryless property anymore since we're using a uniform random variable.

$$P(X > 5 + t_0 \mid X > t_0) = \frac{P(X > 5 + t_0, X > t_0)}{P(X > t_0)} = \frac{P(X > 5 + t_0)}{P(X > t_0)} = \frac{(20 - 5 - t_0)/20}{(20 - t_0)/20} = ??$$

I can't finish my calculations without knowing t_0, the period of time the E.coli has already been living when befriended.

20. An urn initially contains 5 red marbles and 7 blue marbles. Bored out of your mind, you decide to play a "game" that goes as follows. During each round, you randomly pull a marble out of the urn. If the marble you chose was red, you return the marble back into the urn along with 1 more blue marble. If the marble you chose was blue, you put the marble back into the urn along with 2 more red marbles. What is the probability that on the first round you drew a red marble, and on the third round you drew a blue marble?

Solution: Let R_i be the event that you drew a red marble on the ith round and B_i be the event that you drew a blue marble on the ith round ($i = 1, 2, 3$). We are looking for $P(R_1 \cap B_3)$. We divide this into the cases which consider what you drew on the second round,

$$P(R_1 \cap B_3) = P(R_1 \cap R_2 \cap B_3) + P(R_1 \cap B_2 \cap B_3).$$

Now, using the multiplication rule

$$\begin{aligned} P(R_1 \cap B_3) &= P(R_1)P(R_2|R_1)P(B_3|R_2 \cap R_1) + P(R_1)P(B_2|R_1)P(B_3|B_2 \cap R_1) \\ &= \frac{5}{7+5} \cdot \frac{5}{8+5} \cdot \frac{9}{9+5} + \frac{5}{7+5} \cdot \frac{8}{8+5} \cdot \frac{8}{8+7} \\ &= \frac{5}{12} \cdot \frac{5}{13} \cdot \frac{9}{14} + \frac{5}{12} \cdot \frac{8}{13} \cdot \frac{8}{15} \end{aligned}$$

21. The (X, Y) location of the Mothman lands within an 2 mile radius, centered at downtown Point Pleasant, West Virginia (think of a giant 2 mile radius dartboard centered at Point Pleasant and Mothman as a dart landing at coordinate (X, Y) on this dartboard). It turns out that X and Y are jointly continuous with joint density

$$f_{X,Y}(s,t) = \begin{cases} k & s^2 + t^2 \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

where k is some constant.

(a) What is k? (This can be very quick with appropriate justification!)

Solution: Let R be the region where our joint density function is non-zero. Note that R is a circle of radius 2 centered at $(0, 0)$. Then,

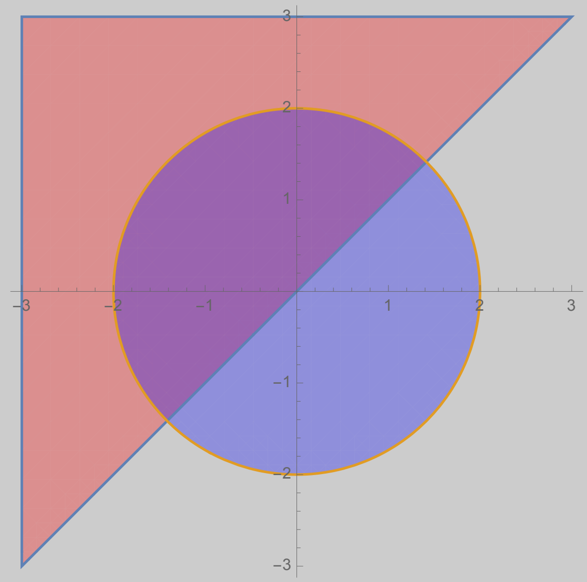
$$1 = \iint_R k\, DA = k \iint_R dA = k \cdot (\text{ area of R }) = k\pi 2^2 = 4k\pi \implies k = \frac{1}{4\pi}$$

Alternatively, we could have solved the integral using our multivariable skills (and polar coordinates).

$$1 = \iint_R k\, DA = k \iint_R dA = k \int_0^{2\pi} \int_0^2 r\, dr\, d\theta = k \int_0^{2\pi} \frac{r^2}{2}\Big|_0^2 d\theta = k \int_0^{2\pi} 2\, d\theta = 4k\pi$$

(b) What is $P(X < Y)$? (Again, this can be very quick with appropriate justification!)

Solution: Our region here is half of a circle.



Then,

$$P(X < Y) = \frac{\text{area of } 1/2 \text{ circle}}{\text{area of circle}} = \frac{1}{2}$$

Alternatively, we could also have found the probability by solving the integral (using polar coordinates again)

$$P(X < Y) = \frac{1}{4\pi} \int_{\pi/4}^{5\pi/4} \int_0^2 r \, dr \, d\theta = \frac{1}{4\pi} \int_{\pi/4}^{5\pi/4} \frac{r^2}{2}\Big|_0^2 \, d\theta = \frac{1}{4\pi} \int_{\pi/4}^{5\pi/4} 2 \, d\theta = \frac{1}{2}$$

(c) Your professor is a sucker. You invite her to play a game: if Mothman lands within 1 mile of downtown Point Pleasant, then you pay your professor \$10; if Mothman lands between 1 and 2 miles from downtown Point Pleasant, then your professor will pay you \$8. Your professor agrees to play. How much money do you expect to win?

Solution: Let W equal your wealth. Then,

$$S_W = \{-10, 8\}.$$

Since we want to find $\mathbb{E}[W]$, we need to find $P(W = -10)$ and $P(W = 8)$.

$$P(W = -10) = P(X^2 + Y^2 \leq 1) = \frac{\text{area of circle or radius 1}}{\text{area of circle of radius 2}} = \frac{\pi}{4\pi} = \frac{1}{4}$$

Alternatively, we could have computed the integral.

$$\frac{1}{4\pi} \int_0^{2\pi} \int_0^1 r \, dr \, d\theta = \frac{1}{4\pi} \int_0^{2\pi} \frac{r^2}{2}\Big|_0^1 \, d\theta = \frac{1}{4\pi} \int_0^{2\pi} \frac{1}{2} \, d\theta = \frac{1}{4}$$

Now, we know that

$$P(W = 8) = 1 - P(W = -10) = 1 - \frac{1}{4} = \frac{3}{4}.$$

Thus,

$$\mathbb{E}[W] = -10P(W = -10) + 8P(W = 8) = -10\left(\frac{1}{4}\right) + 8\left(\frac{3}{4}\right) = \frac{-10}{4} + \frac{24}{4} = \frac{7}{2}$$

You expect to win \$3.50.

22. Following her true dream, your professor opens a surf shop (and school) in Puerto Rico. Suppose that for each i, we let X_i be the number of customers that make a purchase at your professor's shop on day i. Assume that the collection $X_1, X_2, X_3, ...$ are independent and identically distributed such that $X_i \stackrel{d}{=} \text{Pois}(1)$. Use the Central Limit Theorem (i.e., a normal approximation) to approximate the probability that at most 7 customers made a purchase within the first 9 days. That is, approximate $P\big(\sum_{i=1}^{9} X_i \leq 7\big)$ using the Central Limit Theorem. You can leave your answer in terms of Φ, the CDF of a standard normal random variable.

Solution: Since $X_i = \text{Pois}(1)$, we know $\mu = \mathbb{E}[X_i] = 1$. We also know that $\sigma^2 = \text{Var}(X_i) = 1$. Then,

$$\begin{aligned} P\Big(\sum_{i=1}^{9} X_i \leq 7\Big) &= P\left(\frac{\sum_{i=1}^{9} X_i - 9 \cdot 1}{\sqrt{9 \cdot 1}} \leq \frac{7 - 9 \cdot 1}{\sqrt{9 \cdot 1}}\right) \\ &= P\left(Z \leq \frac{-2}{3}\right) \\ &= \Phi(-2/3) \\ &= 1 - \Phi(2/3) \end{aligned}$$