

---

## Midterm 1

---

Last Name	First Name	SID
-----------	------------	-----

***Rules.***

- No form of collaboration between the students is allowed. If you are caught cheating, you may fail the course and face disciplinary consequences.
- You have 10 minutes to read the exam and 90 minutes to complete it.
- The exam is not open book. No calculators or phones allowed.
- Unless otherwise stated, all your answers need to be justified. Show your work to get partial credit.
- Maximum you can score is 120 but 100 points is considered perfect.
- Don't read too much into the size of the answer boxes.

Problem	points earned	out of
Problem 1		49
Problem 2		15
Problem 3		19
Problem 4		17
Problem 5		20
Total		120

# 1 Assorted Problems [7+7+7+7+7+7+7]

## (a) MGF Basics

Let  $X$  be a random variable whose MGF is given as:

$$M_X(t) = \frac{1}{6}e^{-2t} + \frac{1}{3}e^{-t} + \frac{1}{4}e^t + \frac{1}{4}e^{2t}$$

Compute the probability that  $|X| \leq 1$ .

Looking at the form of  $M_X$ , we see that  $X$  must be a discrete random variable. In particular, we see that  $X$  takes on the values  $-2, -1, 1, 2$  with probabilities  $\frac{1}{6}, \frac{1}{3}, \frac{1}{4}, \frac{1}{4}$  respectively. Thus the probability that  $|X| \leq 1$  is  $\frac{1}{3} + \frac{1}{4} = \frac{7}{12}$ .

## (b) Weather Entropy

Let  $W$  take on values in  $\{sunny, clear, rainy\}$ , denoting the weather for a specific day. It is *sunny* on 50% of days, *clear* on 25% of days, and *rainy* on 25% of days. Now, let  $C$  take on values in  $\{tshirt, jacket, sweater\}$ , denoting the clothing that Justin is wearing for a specific day. Suppose Justin chooses his clothes based on the weather  $W$  of the day, with probabilities given by the table below; for example, when it is *sunny*, he wears a *tshirt* 50% of the time. What is the entropy of  $C$ ?

Clothing	<i>sunny</i>	<i>clear</i>	<i>rainy</i>
<i>tshirt</i>	50%	0%	0%
<i>jacket</i>	50%	25%	75%
<i>sweater</i>	0%	75%	25%

We can compute the marginal probability  $p_C(c)$  and find the entropy  $H(C)$  from there.

$$\begin{aligned} p_C(tshirt) &= p_W(sunny)p_{C|W=sunny}(tshirt|sunny) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4} \\ p_C(jacket) &= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = \frac{1}{2} \\ p_C(sweater) &= \left(\frac{1}{4}\right)\left(\frac{3}{4}\right) + \left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \frac{1}{4} \end{aligned}$$

Then,

$$\begin{aligned} H(C) &= \sum_c p_C(c) \log_2\left(\frac{1}{p_C(c)}\right) \\ &= \frac{1}{4} \log_2 4 + \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 \\ &= \frac{3}{2} \end{aligned}$$

(c) **Gaussian Comparison**

In this problem we will consider normal RVs  $\mathcal{N}(\mu, \sigma^2)$ . Let  $X_1$  be  $\mathcal{N}(0, 1)$ ,  $X_2$  be  $\mathcal{N}(2, 25)$ , and  $X_3$  be  $\mathcal{N}(-1, 100)$ . Rank the following probabilities from least to greatest. Your answer should be in the form of  $? < ? < ?$ .

$$a = P(0 < X_1 < 1) \quad b = P(4 < X_2 < 7) \quad c = P(-6 < X_3 < 4)$$

Let  $Z \sim N(0, 1)$ . Then

- $a = P(\frac{0-0}{1} < Z < \frac{1-0}{1}) = P(0 < Z < 1)$ .
- $b = P(\frac{4-2}{5} < Z < \frac{7-2}{5}) = P(0.4 < Z < 1)$
- $c = P(\frac{-6-(-1)}{10} < Z < \frac{4-(-1)}{10}) = P(-0.5 < Z < 0.5)$

$$b < a < c.$$

(d) **Random Functions**

Let the set  $S_N$  denote the first  $N$  positive integers, i.e.  $\{1, 2, 3, \dots, N\}$ . We randomly generate a function  $f$  over  $S_N$  as follows. For each  $x \in S_N$ , let  $f(x)$  be uniform in  $S_N$ . Let  $X$  be the size of the range of  $f$ , i.e.  $X = |\{f(x) : x \in S_N\}|$ .

**Example:** for  $N = 2$ , the uniform sample space of  $f$  and the corresponding values of  $X$  would be:

$$f_1 : f_1(1) = 1, f_1(2) = 1, X = 1$$

$$f_2 : f_2(1) = 1, f_2(2) = 2, X = 2$$

$$f_3 : f_3(1) = 2, f_3(2) = 1, X = 2$$

$$f_4 : f_4(1) = 2, f_4(2) = 2, X = 1$$

Compute  $\mathbb{E}[X]$ .

Let  $X_i$  denote the event that there exists some input  $x$  such that  $f(x) = i$ . We can see that  $X = \sum_{i=1}^N X_i$ .  $\mathbb{E}[X_i]$  is the probability that  $X_i$  is true. The complement is the event that no element in the domain of  $f$  maps to  $i$ . This probability is  $(1 - \frac{1}{N})^N$ . Thus,  $\mathbb{E}[X_i] = 1 - (1 - \frac{1}{N})^N$ . It then follows that  $\mathbb{E}[X] = N\mathbb{E}[X_1] = N(1 - (1 - \frac{1}{N})^N)$ .

(e) **3x3x3 Black Cube**

Suppose you have a 3x3x3 inch cube of solid black marble. You paint all 6 faces white and chop it up into 1x1x1 inch subcubes. You then toss all of these in the bag, and with your eyes closed, you take one out and roll it. Opening your eyes, you notice that 5 faces are black. What's the probability that the face you can't see (i.e. on the bottom) is also black?

Define the following events.

- Let  $O$  be the event that you observe 5 black faces.
- Let  $B_5$  be the event that the subcube you drew has 5 black faces. There are 6 such cubes, so  $P(B_5) = \frac{6}{27}$ .
- Let  $B_6$  be the event that the subcube you drew has 6 black faces.  $P(B_6) = \frac{1}{27}$ .

$$P(B_6 | O) = \frac{P(O | B_6)P(B_6)}{P(O | B_6)P(B_6) + P(O | B_5)P(B_5)}$$

$P(O | B_6) = 1$  and  $P(O | B_5) = \frac{1}{6}$ . The answer is

$$\frac{1 \cdot \frac{1}{27}}{1 \cdot \frac{1}{27} + \frac{1}{6} \cdot \frac{6}{27}} = \frac{1}{2}$$

(f) **1s, 0s, and -1s**

Given a list containing exactly  $n$  1's,  $m$  -1's, and  $k$  0's, where  $n > m$ , what is the probability that the sum of each prefix is strictly positive (not including the empty list)?

**Example:** If our list is  $[1, 0, 1, -1, 0, 1, -1]$ , then we are looking for the event that  $[1]$  and  $[1, 0]$  and  $[1, 0, 1]$  and ... and  $[1, 0, 1, -1, 0, 1, -1]$  all are positive when you sum their elements (which in this case is true).

**Solution 1**

This is similar to the Captain America and Superman problem from HW1. In this case let a tie be when the prefix sum is 0.

$$\begin{aligned} P(\text{strictly positive}) &= 1 - P(\text{tie}) \\ &= 1 - P(\text{tie} \cap 1 \text{ first}) - P(\text{tie} \cap 0 \text{ first}) - P(\text{tie} \cap -1 \text{ first}) \end{aligned}$$

A few observations.

- $P(\text{tie} \cap 0 \text{ first}) = P(0 \text{ first}) = \frac{k}{n+m+k}$  since if there's a 0 first we immediately have a tie.
- $P(\text{tie} \cap -1 \text{ first}) = P(-1 \text{ first}) = \frac{m}{n+m+k}$  since if there's a -1 first we're bound to have a tie.
- $P(\text{tie} \cap 1 \text{ first}) = P(\text{tie} \cap -1 \text{ first}) = \frac{m}{n+m+k}$  since there is a 1:1 mapping between every outcome where there is a tie and 1 first and every outcome where there is a tie and -1 first (namely by swapping the 1s and -1s up to the first tie).

Our answer is:

$$1 - \frac{2m}{n+m+k} - \frac{k}{n+m+k} = \frac{n-m}{n+m+k}$$

**Solution 2**

The 0s don't matter, except that we can't start with a 0 first. The probability that we don't start with a 0 is  $\frac{n+m}{n+m+k}$ . From there, we know the probability that Captain America stays strictly ahead is  $\frac{n-m}{n+m}$  from homework. So the answer is

$$\frac{n+m}{n+m+k} \cdot \frac{n-m}{n+m} = \frac{n-m}{n+m+k}$$

**(g) Ants**

There are  $n$  ants uniformly and independently distributed on a number line from 0 to 1. Each ant, with equal probability and independent of other ants, moves left or right at a speed of 1 unit per second. When two ants hit each other, they bounce back in opposite directions. An ant finishes once it reaches either 0 or 1. How long does it take for the last ant to finish in expectation?

**Hint:** Think carefully about two ants colliding. Is it as complicated as it seems?

Because the ants are indistinguishable, two ants colliding with each other is equivalent to two ants passing through each other. With this realization, let  $X_i$  be the position of the  $i$ -th ant and  $T_i$  be the time it takes to reach the end. Since the ant either moves left or right, we know that

$$T_i = \begin{cases} X_i & \text{w.p. } 0.5 \\ 1 - X_i & \text{w.p. } 0.5 \end{cases}$$

It turns out  $T_i$  is also uniform on  $U[0, 1]$ , and since we're trying to find  $E[\max_i T_i]$ , our answer is  $\frac{n}{n+1}$ .

**Derivation of maximum of i.i.d.  $U[0, 1]$**  Let  $S = \max_i T_i$ .

$$\begin{aligned} P(S \leq s) &= P(\cap_{i=1}^n T_i \leq s) \\ &= \prod_{i=1}^n P(T_i \leq s) \\ &= s^n \end{aligned}$$

Differentiating, we get the PDF to be  $f_S(s) = ns^{n-1}$ . Then, the expectation is

$$\begin{aligned} E[S] &= \int_0^1 s \cdot ns^{n-1} ds \\ &= \int_0^1 ns^n ds \\ &= \frac{n}{n+1} \end{aligned}$$

## 2 L&S CS and EECS [7+8]

There are  $n$  L&S CS students and  $n$  EECS students sitting at a round table. Every seating permutation is equally likely.

- (a) What's the probability that all the L&S CS students sit next to each other?

There are  $\binom{2n}{n}$  ways to choose the seats the L&S CS students sit at. Of those, the ones where the L&S CS students all sit together are  $[1, 2, \dots, n]$ ,  $[2, 3, \dots, n + 1]$ ,  $\dots$ ,  $[2n, 1, \dots, n - 1]$ . So the answer is  $\frac{2n}{\binom{2n}{n}}$ .

- (b) Let  $X$  be the number of lonely EECS students (i.e. EECS students who are not sitting next to any other EECS students). What is  $\mathbb{E}[X]$ ?

An EECS student is lonely if he is sitting between two L&S CS students. Let  $X_i$  be an indicator if the  $i$ -th EECS student is sitting between two L&S CS students. Then  $X = \sum_{i=1}^n X_i$ .

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}\left[\sum_{i=1}^n X_i\right] \\ &= \sum_{i=1}^n \mathbb{E}[X_i] \\ &= n \cdot \mathbb{E}[X_1]\end{aligned}$$

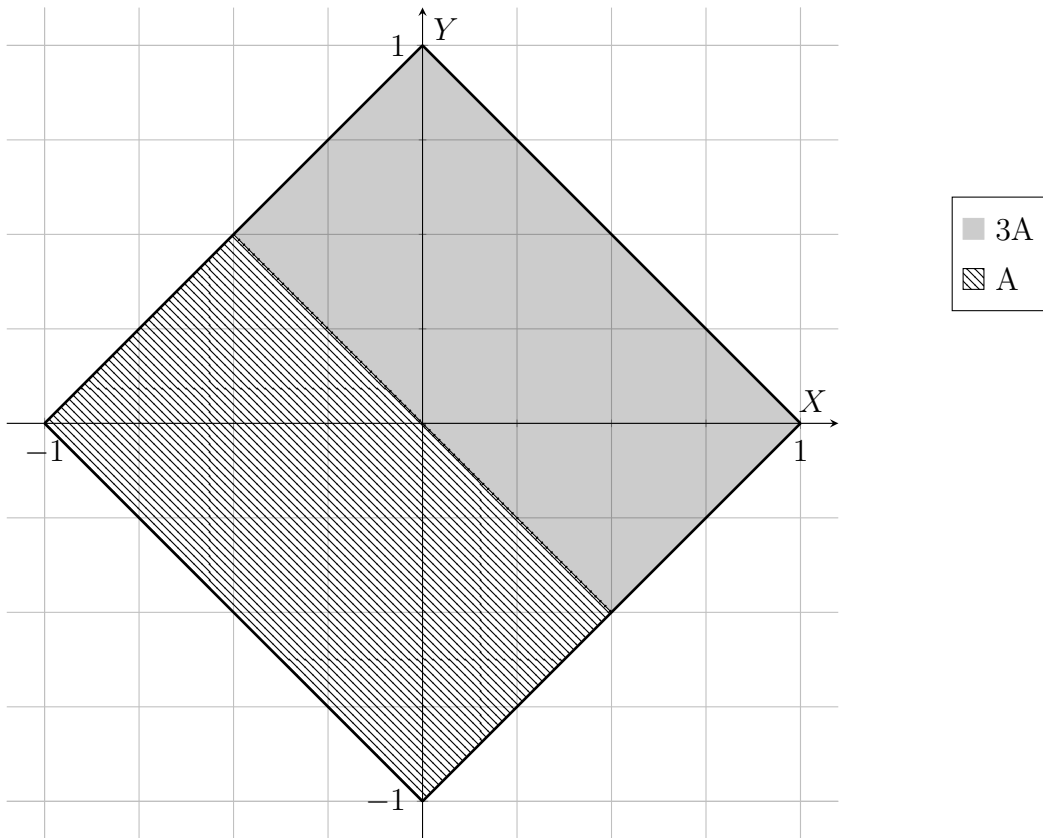
Let  $L$  be the event that a L&S CS student is sitting to the left of the EECS student 1 and  $R$  be the event that a L&S CS student is sitting to the right of the EECS student 1.

$$\begin{aligned}\mathbb{E}[X_1] &= P(L \cap R) \\ &= P(L) \cdot P(R \mid L) \\ &= \frac{n}{2n-1} \cdot \frac{n-1}{2n-2} \\ &= \frac{n}{2n-1} \cdot \frac{1}{2}\end{aligned}$$

So the answer is  $\frac{n^2}{2(2n-1)}$ .

### 3 Graphical Diamond [3+8+8]

Random variables  $X$  and  $Y$  have a joint density as pictured below:



- (a) Find the value of  $A$ .

The area of the square is 2, and each half has an area of 1. Thus, we need  $A+3A = 1$  for  $f_{X,Y}(x,y)$  to be a valid probability density. So,  $A = \frac{1}{4}$ .

- (b) Find  $f_Y(y)$  for  $y \in [0, 1]$  (Note: this is not a condition, just a way to reduce the amount of computation you have to do).

We can split this up into two intervals to make the marginal density easy to compute.

For  $y \in [1/2, 1]$ , we have  $f_Y(y) = \int_{y-1}^{1-y} \frac{3}{4} = \frac{3-3y}{2}$ .

For  $y \in [0, 1/2]$ , we have to split it up into two portions:

$$\begin{aligned} f_Y(y) &= \int_{y-1}^{-y} \frac{1}{4} + \int_{-y}^{1-y} \frac{3}{4} \\ &= \frac{1-2y}{4} + \frac{3}{4} \\ &= 1 - \frac{y}{2} \end{aligned}$$

We perform the analogous computations for the other two intervals to get the density:

$$f_Y(y) = \begin{cases} \frac{3-3y}{2} & \text{if } y \in [1/2, 1] \\ 1 - \frac{y}{2} & \text{if } y \in [0, 1/2] \end{cases}$$

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

If we split up the diamond into 4 quadrants, we have 4 sub-diamonds each with uniform density. The top and bottom quadrant have expected value 0 by inspection, and the left and right diamonds have expected value  $\frac{-1}{2}$  and  $\frac{1}{2}$  respectively. Now, we can take the total expected value, weighting each diamond accordingly.  $\mathbb{E}[X] = \sum_{S_i=\text{sub-diamond}} P(\text{being in diamond})E[S_i] = \frac{3}{8}\frac{1}{2} + \frac{1}{8}\frac{-1}{2} = \frac{1}{8}$ .

Alternate solution: a more interesting way to compute this is by noticing that the center of mass (or density) of the square is equidistant from the origin for all rotations of the square around the origin. So, we can rotate the square 45 degrees clockwise to make the problem a lot easier, computing  $\mathbb{E}[X^{45^\circ}] = \frac{\sqrt{2}}{8}$ . Then, we can find the x component of the vector rotated 45 degrees counter clockwise of magnitude  $\frac{\sqrt{2}}{8}$ , which is  $\cos(45^\circ)\sqrt{2}/8 = \frac{1}{8} = \mathbb{E}[X]$



## 4 Laplace [8+9]

Suppose  $X \sim \text{Laplace}(0, b)$ , where the PDF for the zero-centered Laplace distribution is

$$p_X(x) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right)$$

- (a) What's the conditional PDF for  $X$  if we know that  $X > k$ , where  $k > 0$ ?

**Hint:** The Laplace distribution is a composition of symmetric Exponentials.

The key observation here is that the Laplace is composed of two symmetric Exponential( $\frac{1}{b}$ ) distributions around 0. Then, letting  $Y \sim \text{Exponential}(\frac{1}{b})$ , the memoryless property implies that

$$p_X(x|X > k) = p_Y(x - k) = \frac{1}{b} \exp\left(-\frac{1}{b}(x - k)\right)$$

for  $x > k$ , otherwise the PDF should be 0.

- (b) Compute the variance of  $X \sim \text{Laplace}(0, b)$ . Your answer should depend on  $b$ .

Noting that the mean of  $X$  is 0, we know that  $\text{var}(X) = E[X^2]$ .

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right) dx \\ &= 2 \int_0^{\infty} x^2 \cdot \frac{1}{2b} \exp\left(-\frac{x}{b}\right) dx \\ &= \int_0^{\infty} x^2 \cdot \frac{1}{b} \exp\left(-\frac{x}{b}\right) dx \end{aligned}$$

This is the second moment of an exponential RV with  $\lambda = \frac{1}{b}$ . We know that for an exponential RV,  $E[X^2] = \frac{2}{\lambda^2}$ , so the answer is  $2b^2$ .

## 5 NumPy Confusion [5+7+8]

NumPy uses a method called inverse transform sampling to generate a number according to a random variable  $X$ . In particular, given the CDF  $F_X$  of  $X$ , they generate a seed  $U \sim \text{Uniform}[0, 1]$  and then return  $Z = F_X^{-1}(U)$ .

(a) Show that  $X$  and  $Z$  have the same CDF.

$F_Z(x) = \mathbb{P}(Z \leq x) = \mathbb{P}(F_X^{-1}(U) \leq x) = \mathbb{P}(U \leq F_X(x)) = F_X(x)$ . Note: since the CDFs are exactly the same, they also have the same PDF.

(b) Suppose  $X \sim \text{Exponential}(1)$ . What is the corresponding  $Z$  in terms of  $U$ ?

We know that  $Z = F_X^{-1}(U)$ , and thus that  $F_X(Z) = U$ .  $F_X(Z) = 1 - e^{-Z}$ , so letting  $1 - e^{-Z} = U$ , we get that  $Z = -\ln(1 - U)$ .  $Z = -\ln(U)$  is also valid since  $U$  and  $1 - U$  are the same.

(c) In NumPy, calling `np.random.exponential(0, λ)` actually generates a number according to the random variable  $\text{Exponential}(\frac{1}{\lambda})$ , not  $\text{Exponential}(\lambda)$ . Suppose we passed in the rate  $\lambda$  instead of the mean  $\frac{1}{\lambda}$ . We meant to generate according to  $X \sim \text{Exponential}(\lambda)$ , but instead did according to  $Y \sim \text{Exponential}(\frac{1}{\lambda})$ . What is  $E\left[\left(F_X^{-1}(U) - F_Y^{-1}(U)\right)^2\right]$ ?

**Hint:** Use part (b).

Extending part (b) a bit, we know that  $F_X^{-1}(U) = -\frac{\ln(1-U)}{\lambda}$  and  $F_Y^{-1}(U) = -\lambda \ln(1 - U)$ . So

$$\begin{aligned} E\left[\left(F_X^{-1}(U) - F_Y^{-1}(U)\right)^2\right] &= E\left[\left(-\frac{\ln(1-U)}{\lambda} - (-\lambda \ln(1-U))\right)^2\right] \\ &= E\left[\left(\frac{1}{\lambda} - \lambda\right)^2 \left(-\ln(1-U)\right)^2\right] \\ &= \left(\frac{1}{\lambda} - \lambda\right)^2 \cdot E\left[\left(-\ln(1-U)\right)^2\right] \\ &= \left(\frac{1}{\lambda} - \lambda\right)^2 \cdot E\left[\text{Exponential}(1)^2\right] \\ &= 2\left(\frac{1}{\lambda} - \lambda\right)^2 \end{aligned}$$

where the last line follows from the second moment of an  $\text{Exponential}(1)$  distribution.