

Student name: _____

Student Number: _____

CSE 312: Foundations of Computing II

Winter 2024

Final Exam

Important: Read the instructions on this page carefully, but do not turn the page until you are instructed to do so.

Instructions:

- **Write your name and student number on top of this page**, and write your name on top of every other page as well.
- This is a **closed-book exam, with the exception of a single double-sided cheat sheet**. You have 110 minutes to do it. It is worth 115 points.
- **No electronics are allowed** during the exam (no phones, no laptops, no smart watches, no pocket calculators, etc). Keep them stored in your bag.
- **IMPORTANT:** Be sure to put your **final** answers, and only your final answers, in the box provided for each part of each question. If the answer in the box is correct, you will get full credit for the problem regardless of what else you write.
- If the answer in the box is incorrect, there is some possibility you will get partial credit for any explanations you provide. Those explanations should be written in the space between the question and the box in which you put the final answer. If absolutely necessary, use the back side of the same page to finish writing. But you shouldn't need additional space.
- When you are finished, turn in your test only. (Do **not** turn in your cheat sheet.)

Good luck!

2 points each. For each of the following **circle one of True or False**. Note that "True" means "Always True".

1. **True or False:** Circle one. For any random variable X

$$\mathbb{E} \left[\binom{X}{2} \right] = \frac{(\mathbb{E}[X])^2 - \mathbb{E}(X)}{2}$$

2. **True or False:** A coin is flipped independently 1000 times, where the probability that the coin comes up heads is p . Then the probability that there are exactly 500 heads and 500 tails is $p^{500}(1-p)^{500}$.
3. **True or False:** Suppose that $Z = X + Y$, where X and Y are independent, discrete random variables with $\Omega_X = \{0, 1, 2, \dots, n\}$ and $\Omega_Y = \{0, 1, 2, \dots, n\}$. Then

$$\mathbb{P}(Z = z) = \sum_{k=0}^n \mathbb{P}(X = k) \cdot \mathbb{P}(Y = z - k).$$

4. **True or False:** If X and Y are independent random variables, then

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y).$$

5. **True or False:** The number of ways to select 7 cards from a standard deck of 52 cards if we want at least one card from each suit is $13^4 \binom{48}{3}$. (Order does not matter.)
6. **True or False:** Suppose X is a continuous random variable. The PDF $f_X(x)$ represents the probability that $X = x$ and the CDF $F_X(x)$ represents the probability that $X \leq x$.
7. **True or False:** Let $X \sim \text{Bin}(n, p)$ and $Y \sim \text{Bin}(m, p)$ be independent binomial R.V.s. Then $X + Y \sim \text{Bin}(n + m, p)$.
8. **True or False:** If A and B are independent events and $\mathbb{P}(A) > 0$ and $\mathbb{P}(B) > 0$, then it is impossible for A and B be mutually exclusive.
9. **True or False:** A curator wants to display 50 unique paintings in 7 different museums. The number of different ways to assign the paintings to the museums is $\binom{56}{50}$. (Any number of paintings can go into any of the 7 museums, and each painting goes to exactly one museum.)
10. **True or False:** Let X be a random variable with $\mathbb{E}[X] = \mu$ and $\text{Var}(X) = \sigma^2$. Then $\mathbb{P}(X \leq k) = \Phi\left(\frac{k-\mu}{\sigma}\right)$, where Φ is the CDF of the standard normal distribution.
11. **True or False:** If X_1, X_2, \dots, X_5 are i.i.d random variables, $\sum_{i=1}^5 X_i = 5 \cdot X_1$.
12. **True or False:** $\mathbb{E}[X_1 + X_2 + \dots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n]$ implies that X_1, \dots, X_n are all independent.
13. **True or False:** The following is a valid probability density function:

$$f_X(x) = \begin{cases} 2 & \text{if } 0 < x < 0.25 \\ 0 & \text{otherwise} . \end{cases}$$

Task 2 – Normals

[15 pts]

Let X_1 and X_2 be independent $N(\mu, \sigma^2)$ random variables.

(a) (3 points) What is $\mathbb{E}[X_1 + 2X_2]$?

$$\mathbb{E}[X_1 + 2X_2] =$$

(b) (4 points) What is $\text{Var}(X_1 + 2X_2)$?

$$\text{Var}(X_1 + 2X_2) =$$

(c) (4 points) Suppose that X has a normal distribution with mean μ and variance σ^2 . Find real numbers a and b such that $aX + b \sim N(0, 1)$.

$$a =$$

$$b =$$

(d) (4 points) Suppose that X has a normal distribution with mean μ and variance σ^2 . What is the probability that X takes a value between $\mu + 0.75\sigma$ and $\mu + 1.5\sigma$? (Recall that $\Phi(z) = P(Z \leq z)$ where $Z \sim N(0, 1)$.) **Put an X in the box next to the best answer. Choose only one answer!**

$\mathbb{P}(\mu + 0.75\sigma < X < \mu + 1.5\sigma) = \Phi(\mu + 1.5\sigma) - \Phi(\mu + 0.75\sigma)$.

$\mathbb{P}(\mu + 0.75\sigma < X < \mu + 1.5\sigma) = \Phi(1.5) - (1 - \Phi(0.75))$.

$\mathbb{P}(\mu + 0.75\sigma < X < \mu + 1.5\sigma) = \Phi(1.5) - \Phi(0.75)$.

$\mathbb{P}(\mu + 0.75\sigma < X < \mu + 1.5\sigma) = \Phi(1.5) - \Phi(-0.75) - \Phi(0.75) + \Phi(-1.5)$.

None of the above.

Task 3 – CLT

[15 pts]

The time between consecutive Instagram posts by Kylie Jenner is exponentially distributed with an **expected value** of 2 days. (The times between different posts are independent random variables.) At the stroke of midnight, she posts a picture on Instagram. Let T be the **time until she posts another 100 times**.

- (a) (3 points) What is the expected value of T ?

$$\mathbb{E}[T] =$$

- (b) (3 points) What is the variance of T ?

$$\text{Var}(T) =$$

- (c) (6 points) Let Z be a standard normal random variable. Use the Central Limit Theorem to fill in the blank in the box below. (Note that you are just filling in a number on the dotted line, you are not solving for the actual probability.)

$$\mathbb{P}(T > 205) \approx \mathbb{P}(Z > \dots\dots\dots).$$

- (d) (3 points) Suppose that a is the correct answer to the previous part of this problem (i.e., $\mathbb{P}(T > 205) \approx \mathbb{P}(Z > a)$). Write an expression for $\mathbb{P}(Z > a)$ in terms of $\Phi(a)$, where $\Phi(z) = \mathbb{P}(Z \leq z)$, is the CDF of the standard normal.

$$\mathbb{P}(Z > a) =$$

Task 4 – MLE

[12 pts]

Suppose that x_1, \dots, x_n are i.i.d. realizations from density

$$f_X(x; \theta) = \begin{cases} \frac{\theta x^{\theta-1}}{5^\theta}, & 0 < x < 5 \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown parameter we are trying to estimate.

- (a) (3 points) What is the likelihood function $\mathcal{L}(x_1, \dots, x_n; \theta)$, assuming that $0 < x_i < 5$, for $i = 1, \dots, n$?

$$\mathcal{L}(x_1, \dots, x_n; \theta) =$$

- (b) (4 points) What is the log likelihood function $\ln \mathcal{L}(x_1, \dots, x_n; \theta)$, assuming that $0 < x_i < 5$, for $i = 1, \dots, n$? (No credit for this part unless you simplify the expression using properties of the log function.)

$$\ln \mathcal{L}(x_1, \dots, x_n; \theta) =$$

(c) (5 points) Put an X in every box for which the corresponding statement is correct. There may be more than one correct answer.

Ignoring verification of second order conditions, the maximum likelihood estimate for θ is:

The solution to $\mathcal{L}(x_1, \dots, x_n; \theta) = 0$.

The solution to $\ln \mathcal{L}(x_1, \dots, x_n; \theta) = 0$.

The solution to $\frac{d}{d\theta} \mathcal{L}(x_1, \dots, x_n; \theta) = 0$.

The solution to $\frac{d}{d\theta} (\ln \mathcal{L}(x_1, \dots, x_n; \theta)) = 0$.

The solution to $\sum_{i=1}^n \frac{d}{dx_i} \mathcal{L}(x_1, \dots, x_n; \theta) = 0$.

Task 5 – Joint densities

[8 pts]

Let $f_{X,Y}(x, y)$ be the joint density function of two random variables X and Y . Give an expression for $\mathbb{P}(X > 1)$ in terms of $f_{X,Y}(x, y)$. (Your answer will involve integrals.)

$\mathbb{P}(X > 1) =$

Task 6 – Bayes Theorem

[10 pts]

You ask your neighbor to water a sickly plant while you are on vacation. Without water, it will die with probability 0.8. With water, it will live with probability 0.85. You are 90% sure that your neighbor will remember to water the plant. Your answers below should involve only numbers but you do not have to calculate them out.

It may be helpful to use the following notation:

- R is the event that the neighbor waters the plant (and \bar{R} is the complement event, i.e., that the neighbor forgets and doesn't water the plant).
- D is the event that the plant is dead upon your return.

(a) Fill in the blanks:

$$\begin{aligned}\mathbb{P}(D|\bar{R}) &= \dots\dots\dots \\ \mathbb{P}(D|R) &= \dots\dots\dots\end{aligned}$$

(b) What is the probability that the plant is dead when you return?

$$\mathbb{P}(D) =$$

(c) What is the probability your neighbor forgot to water the plant given that it is dead when you return? Write your answer in terms of q , the correct answer to part (b).

$$\mathbb{P}(\bar{R}|D) =$$

Task 7 – Lightbulbs

[18 pts]

Consider n lightbulbs and let X_i denote the lifetime of the i -th lightbulb. Suppose that the X_i 's are independent and each follow an exponential distribution with parameter λ .

- (a) (3 points) Suppose we use the lightbulbs **one at a time**, where we start using the i -th lightbulb as soon as the $(i - 1)$ -st lightbulb burns out. What is the expected time until all n lightbulbs have burned out?

Answer:

Suppose that instead we use all n lightbulbs **at the same time** (and this holds for all remaining parts of this problem).

- (b) (5 points) Let T_1 be the time until the first of these n lightbulbs burns out. (It could be any of them that is the first to go.) What is $\mathbb{P}(T_1 > t)$?

$\mathbb{P}(T_1 > t) =$

- (c) (4 points) What is $\mathbb{E}[T_1]$? (Hint: what type of random variable from our zoo is T_1 and what are its parameter(s)?)

$\mathbb{E}[T_1] =$

- (d) (6 points) Let T_2 be the total time from the start until the 2nd of the n light bulbs burns out. What is $\mathbb{E}[T_2]$?

$$\mathbb{E}[T_2] =$$

Task 8 – Bad sorting algorithm

[21 pts]

Consider the following bad (randomized) algorithm for sorting an array $A[0..n-1]$ of n distinct integers $(a_0, a_1, \dots, a_{n-1})$.

1. Randomly shuffle A (i.e. insert a_0, a_1, \dots, a_{n-1} into A according to a uniformly random permutation.)
2. For $i = 0$ to $n-2$, do the following
3. If $A[i] > A[i+1]$ then
4. exit the loop and go back to step 1.

In other words, the elements are shuffled at random (step 1) and then the array elements are inspected in order until the first time we see an element that is greater than the element after it (proving that the array is not actually sorted). As soon as that happens, we restart the whole process by going back to step 1.

- (a) (5 points) Let X be the total number of times that step 1 is executed. What kind of random variable from our zoo is X and what is $\mathbb{E}(X)$?

$\mathbb{E}(X) =$

- (b) (7 points) What is the probability that the first i elements of a random permutation of n distinct numbers are in sorted order? (For example, in the permutation $3, 5, 8, 1, 4, 6, 2, 7$, the first i elements are in sorted order for $i = 1$, for $i = 2$, and for $i = 3$, but not for any $i > 3$.)

Answer:

(c) (9 points) Again, consider the same bad algorithm for sorting an array, i.e.,

1. Randomly shuffle A (i.e. insert a_0, a_1, \dots, a_{n-1} into A according to a uniformly random permutation.)
2. For $i = 0$ to $n-2$, do the following
3. If $A[i] > A[i + 1]$ then
4. exit the loop and go back to step 1.

Let $Y_{k,i}$ be an indicator random variable for the event that $A[i]$ and $A[i + 1]$ are compared between the k -th time line 1 is executed and the $(k + 1)$ st time line 1 is executed (i.e., line 3 is reached for that i).

For example, suppose that $n = 4$. The following table gives an example of what might happen, where k is the number of times step 1 is executed.

k	Elements in shuffled A	Comparisons made in executions of step 3	$Y_{k,0}$	$Y_{k,1}$	$Y_{k,2}$
1	[1,3,2,4]	(1,3), (3,2)	1	1	0
2	[3,2,4,1]	(3,2)	1	0	0
3	[2,3,4,1]	(2,3), (3,4), (4,1)	1	1	1
4	[1,2,3,4]	(1,2), (2,3), (3,4)	1	1	1
>4	No reshuffle	None	0	0	0

Question: What is $\mathbb{E}(Y_{k,i})$? Your answer should work for any $k \geq 1$ and any $i \in \{0, \dots, n - 2\}$.

Hint: Use the law of total expectation conditioning on $X < k$ or not. (See part (a) for the definition of X .)

$$\mathbb{E}(Y_{k,i}) =$$