Student name: ______________ **Student Number:** ______________

CSE 312: Foundations of Computing II — Winter 2024

Final Exam A

Solutions

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 125 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!

Task 1 – True/False [26 pts]

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

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

   False. $\mathbb{E}\left[\binom{X}{2}\right] = \mathbb{E}\left[\frac{X^2-X}{2}\right] = \frac{\mathbb{E}[X^2]-\mathbb{E}[X]}{2}$. But $(\mathbb{E}[X])^2$ is not necessarily equal to $\mathbb{E}[X^2]$.

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 $p^{500}(1-p)^{500}$.

   False. Should be $\binom{1000}{500}p^{500}(1-p)^{500}$.

3. **True or False:** Suppose that $Z = X + Y$, where X and Y are independent, discrete random variables with $\Omega_X = \{0, 1, 2, \ldots, n\}$ and $\Omega_Y = \{0, 1, 2, \ldots, n\}$. Then

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

   True.

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

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

   True

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.)

   False. The counting is explained by picking one card from each suit, which gives 13^4 ways, and then picking the remaining 3 cards from the remaining 48 cards, which gives $\binom{48}{3}$ ways. But this overcounts.

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 $X \leqslant x$.

   False. For a continuous random variable, $P(X = x) = 0$ for all x. The CDF is correct.

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 $X + Y \sim \text{Bin}(n + m, p)$.

True.

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.

   True. If A and B are independent, then $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) > 0$. So A and B cannot 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.)

   False. The key is that each painting is *unique*. Each painting has 7 different choices, totalling 7^{50} total ways.

10. **True or False:** Let X be a random variable with $\mathbb{E}[X] = \mu$ and $\text{Var}(X) = \sigma^2$. Then $\mathbb{P}(X \leqslant k) = \Phi(\frac{k-\mu}{\sigma})$, where Φ is the CDF of the standard normal distribution.

    False. X is not necessarily normal.

11. **True or False:** If $X_1, X_2, \dots, X_5$ are i.i.d random variables, $\sum_{i=1}^{5} X_i = 5 \cdot X_1$.

    False.

12. **True or False:** $\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]$ implies that $X_1, \dots, X_n$ are all independent.

    False. Linearity of expectation holds regardless of whether the random variables are independent or not.

13. **True or False:** The following is a valid probability density function:

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

    False. We see that

$$\int_{-\infty}^{\infty} f_X(x)\, dx = \int_{0}^{0.25} 2\, dx = 2(0.25) = 0.5 \neq 1 \,,$$

    so f fails the normalization property.

Task 2 – Normals [15 pts]

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

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

3μ

(b) (4 points) What is Var$(X_1 + 2X_2)$?

$\sigma^2 + 4\sigma^2 = 5\sigma^2$

(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 $aX + b \sim N(0, 1)$.

Recall that if $X \sim N(\mu, \sigma^2)$, then $aX + b \sim N(a\mu + b, a^2\sigma^2)$.
We need $a\mu + b = 0$ and $a^2\sigma^2 = 1$. Therefore

$$a = 1/\sigma \text{ and } b = -\mu/\sigma \quad \text{OR} \quad a = -1/\sigma \text{ and } b = \mu/\sigma.$$

(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 $\Phi(z) = P(Z \leqslant z)$ 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.

$$Pr(\mu + 0.75\sigma < X < \mu + 1.5\sigma) = Pr\left(0.75 < \frac{X - \mu}{\sigma} < 1.5\right) = \Phi(1.5) - \Phi(0.75).$$

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?

We can write $T = X_1 + X_2 + \cdots + X_{100}$, where X_i's are i.i.d. exponential random variables representing the time between consecutive posts. Since the expected value of each X_i is 2, by linearity of expectation, the expected value of T is $100 \cdot 2 =$ 200 days.

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

Following the previous task, each X_i is an exponential random variable with some parameter λ. We know that the expected value of an exponential random variable is $1/\lambda$, so $\lambda = 1/2$. therefore the variance of each X_i is $1/\lambda^2 = 4$. Since each X_i is independent, by linearity of variance, the variance of T is $100 \cdot 4 = 400$.

(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.)

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

$$Pr(T > 205) = Pr\left(\frac{T - 200}{\sqrt{400}} > \frac{205 - 200}{\sqrt{400}}\right) = Pr(Z > 0.25).$$

(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 $\Phi(z) = \mathbb{P}(Z \leqslant z)$, is the CDF of the standard normal.

$$1 - \Phi(a)$$

Task 4 – MLE **[12 pts]**

Suppose that $x_1, \ldots, x_n$ 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 $\mathcal{L}(x_1, \ldots, x_n; \theta)$, assuming that $0 < x_i < 5$, for $i = 1, \ldots, n$?

$\mathcal{L}(x_1, \ldots, x_n; \theta) = \prod_{i=1}^n \frac{\theta x_i^{\theta-1}}{5^\theta}$

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

$\ln \mathcal{L}(x_1, \ldots, x_n; \theta) = n \ln(\theta) + (\theta - 1) \left(\sum_{i=1}^n \ln(x_i)\right) - \theta n \ln(5).$

(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 $\mathcal{L}(x_1, \ldots, x_n; \theta) = 0.$

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

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

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

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

The third and fourth statements are correct. The third statement finds the maximum of $\mathcal{L}$, while the fourth statement finds the maximum of $\ln \mathcal{L}$. Both of these will give the same answer.

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.)

$$\begin{aligned} Pr(X > 1) &= \int_1^{\infty} f_X(x)\, dx \\ &= \int_1^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy\, dx. \end{aligned}$$

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 $\overline{R}$ 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) (2 points) Fill in the blanks:

$\mathbb{P}\left(D|\overline{R}\right) =$ ....

0.8

$\mathbb{P}\left(D|R\right) =$ ....

0.15

(b) (5 points) What is the probability that the plant is dead when you return?

By the law of total probability,

$$Pr(D) = Pr(D|\overline{R}) \cdot Pr(\overline{R}) + Pr(D|R) \cdot Pr(R) = 0.8 \cdot 0.1 + 0.15 \cdot 0.9$$

(c) (3 points) 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).

By Bayes' theorem,

$$Pr(\overline{R}|D) = \frac{Pr(D|\overline{R}) \cdot Pr(\overline{R})}{Pr(D)} = \frac{0.8 \cdot 0.1}{q}.$$

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?

Each X_i has expected value $1/\lambda$. By linearity of expectation, the expected time until all n lightbulbs have burned out is $\frac{n}{\lambda}$.

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

$Pr(T_1 > t) = Pr(X_i > t \quad \forall i) = \prod_{i=1}^{n} Pr(X_i > t) = (e^{-\lambda t})^n = e^{-\lambda n t}$.
The term $Pr(X_i > t)$ is obtained by the CDF of the exponential distribution:

$$Pr(X_i > t) = 1 - Pr(X_i \leqslant t) = 1 - (1 - e^{-\lambda t}) = e^{-\lambda t}.$$

(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)?)

$\frac{1}{n\lambda}$.
Notice that $Pr(T_1 \leqslant t) = 1 - e^{-(n\lambda)t}$, which is exactly the CDF of an exponential random variable with parameter $n\lambda$. Therefore, T_1 is an exponential random variable with parameter $n\lambda$, and its expected value is $1/(n\lambda)$.

(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 $\mathbb{E}[T_2]$?

We can write $T_2 = T_1 + T_2'$, where T_2' is the time between the first and second lightbulbs burning out. After the first lightbulb burns out, there are $n-1$ lightbulbs left, and each of them has lifetime that is *memoryless* and follows an exponential distribution with parameter λ. Therefore, T_2' is the minimum of $n-1$ independent exponential random variables with parameter λ, which, following the previous arguments in part (b) and (c), is an exponential random variable with parameter $(n-1)\lambda$. Therefore, the expected value of T_2' is $1/((n-1)\lambda)$.
By linearity of expectation, we thus have

$$\mathbb{E}[T_2] = \frac{1}{n\lambda} + \frac{1}{(n-1)\lambda}.$$

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 $(a_0, a_1, \ldots, a_{n-1})$.

1. Randomly shuffle A (i.e. insert $a_0, a_1, \ldots, a_{n-1}$ 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 $\mathbb{E}[X]$?

X is a geometric random variable with parameter $1/n!$ as each shuffle is independent and has a $1/n!$ chance of being sorted. So $E(X) = n!$.

(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 $i > 3$.)

$$\frac{\binom{n}{i}(n-i)!}{n!} = \frac{1}{i!}.$$

There are $\binom{n}{i}$ ways to choose the first i elements, and for each of these choices, there is only one way to order them so that they are in sorted order. The remaining $n - i$ elements can be ordered in any way.

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

1. Randomly shuffle A (i.e. insert $a_0, a_1, \ldots, a_{n-1}$ 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.

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

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.

Question: What is $\mathbb{E}[Y_{k,i}]$? Your answer should work for any $k \geqslant 1$ and any $i \in \{0, \ldots, n-2\}$.

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

We have

$$E(Y_{k,i}) = E(Y_{k,i}|X \geqslant k)Pr(X \geqslant k) + E(Y_{k,i}|X < k)Pr(X < k) = \frac{1}{(i+1)!}\left(1 - \frac{1}{n!}\right)^{k-1}$$

where we used the fact that $Pr(Y_{k,i} = 1|X \geqslant k)$ is precisely the probability that in a random permutation, the first $i+1$ elements are in sorted order (since i runs from 0 to $n-2$). (See part (b).) Also

$$Pr(X \geqslant k) = \left(1 - \frac{1}{n!}\right)^{k-1},$$

as it is the probability that the first $k-1$ shuffles are not sorted.