

Student name: _____

Student Number: _____

CSE 312: Foundations of Computing II

Winter 2023

Final Exam

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.
- **No electronics are allowed** during the exam (no phones, no laptops, no smart watches, no pocket calculators, etc). Keep them stored in your bag.
- Write your solutions in the appropriate spaces only. We have provided you with scratch paper. Feel free to raise your hand and let us know if you need more scratch paper.
- **IMPORTANT:** Be sure to put 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, you can still 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 the scratch paper or your cheat sheet.)

Good luck!

- (a) (5 points) Suppose that X is a random variable with probability mass function

$$p_X(x) = \begin{cases} 1/3 & x = 3 \\ 2/3 & x = -6 \\ 0 & \text{otherwise.} \end{cases}$$

What is $\mathbb{E}[5^X]$? (You do not need to simplify your answer.)

$$\mathbb{E}[5^X] = 5^3 \cdot (1/3) + 5^{-6} \cdot (2/3)$$

- (b) (9 points) Suppose that n distinguishable balls are tossed uniformly at random into n distinguishable bins. What is the probability that bin 1 has 5 balls in it, given that bin 2 is empty?

$$\frac{\binom{n}{5} \left(\frac{1}{n}\right)^5 \left(\frac{n-2}{n}\right)^{n-5}}{\left(\frac{n-1}{n}\right)^n} \text{ or } \binom{n}{5} \frac{(n-2)^{n-5}}{(n-1)^n}$$

- (c) (5 points) Suppose that X and Y are two random variables such that $\mathbb{P}(X = 5) = 0.25$ and $\mathbb{P}(Y = 2|X = 5) = 0.5$. What is $p_{X,Y}(5,2)$?

$$0.5 \cdot 0.25 = 1/8$$

Task 2 – Basic continuous**[27 pts]**

Suppose that the time (in minutes) that it takes a single person to solve a probability problem has density

$$f_X(x) = \begin{cases} \frac{200}{x^3} & \text{for } x > 10 \\ 0 & \text{otherwise.} \end{cases}$$

Your answers to (a) and (b) will be in terms of an integral. You do not need to evaluate any of these integrals.

(a) (5 points) What is p , where $p = \mathbb{P}(X > 100)$?

$$\int_{100}^{\infty} \frac{200}{x^3} dx = -\frac{100}{x^2} \Big|_{100}^{\infty} = \frac{1}{100}$$

(b) (5 points) What is $F_X(200)$?

$$\int_{10}^{200} \frac{200}{x^3} dx = -\frac{100}{x^2} \Big|_{10}^{200} = \frac{399}{400}.$$

(c) (8 points) In terms of p , the answer to part (a) of this problem, what is the probability that of 150 people, exactly 15 take at least 100 minutes? (The time it takes each of these people has density f_X given above and the time it takes different people is independent.)

$$\binom{150}{15} p^{15} (1-p)^{135}$$

(d) (9 points) Suppose that Y is a normal random variable with mean 0 and variance 49. Find $a > 0$ such that aY is normal with variance 9.

If Y is normal, then aY is also normal for any constant a , so we just need to get the variance right. To this end:

$$\text{Var}(aY) = a^2 \text{Var}(Y) = a^2 \cdot 49 = 9.$$

$$\text{So } a = \sqrt{\frac{9}{49}} = \frac{3}{7}$$

Suppose that X and Y have the following joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} 4c & 0 < x \leq 1, 0 < y \leq 1 \\ c & 0 < x \leq 1, 1 < y < 2 \\ c & 1 < x < 2, 0 < y \leq 1 \\ 4c & 1 < x < 2, 1 < y < 2 \\ 0 & \text{otherwise,} \end{cases}$$

where c is a constant that makes $f_{X,Y}(x,y)$ a valid joint density.

Tip: Draw yourself a picture on your scratch paper!

1. (7 points) Find c such that this is a valid joint density.

We must have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = 1.$$

Plugging in the definition of $f_{X,Y}$, we need

$$\begin{aligned} \int_0^2 \int_0^2 f_{X,Y}(x,y) dx dy &= \int_0^1 \int_0^1 4c dx dy + \int_0^1 \int_1^2 c dx dy + \int_1^2 \int_0^1 c dx dy + \int_1^2 \int_1^2 4c dx dy \\ &= 4c + c + c + 4c = 10c = 1 \end{aligned}$$

So $c = 1/10$.

2. (5 points) What is $f_X(x)$ when $0 < x \leq 1$? Leave your answer in terms of c .

$$\int_0^1 4c dy + \int_1^2 c dy = 5c = 1/2$$

3. (5 points) What kind of random variable from our zoo is X ? Be sure to include the parameters of the distribution.

Uniform on $(0,2)$. Similar to part (b), we can also observe that $f_X(x) = 1/2$ when $1 < x < 2$. So $f_X(x) = 1/2$ for all $0 < x < 2$, which is the PDF of a uniform random variable on $(0,2)$.

4. (7 points) Are X and Y independent? Justify your answer. You may use without proof that the marginal density of Y is the same as the marginal density of X .

No, since for example $f_X(0.5) = 1/2$, $f_Y(0.5) = 1/2$ and $f_{X,Y}(0.5, 0.5) = 4c = 4/10 = 2/5$.

Task 4 – Law of total expectation**[12 pts]**

Let X be Binomial with parameters n, p . Let Y be a Geometric distribution with parameter equal to $p = X/n$.

(a) (4 points) What is $\mathbb{E}[Y|X = k]$?

Conditioned on $X = k$, Y is Geometric with parameter k/n . Therefore, its expectation is n/k .

(b) (8 points) Use the law of total expectation to compute $\mathbb{E}[Y]$, in terms of p and n . You can leave it as a sum (ranging over an integer parameter, say k).

$$\mathbb{E}[Y] = \sum_{k=0}^n \mathbb{E}[Y|X = k] \mathbb{P}[X = k] = \sum_{k=0}^n \frac{n}{k} \binom{n}{k} p^k (1-p)^{n-k}.$$

Task 5 – Swaps

[15 pts]

Suppose that the array $A[1 : n]$ contains the elements $1, 2, \dots, n$ in a *uniformly random permutation*. We say that the array contains a *swap* if there exist $i, j \in \{1, \dots, n\}$ where $A[i] = j$ and $A[j] = i$, where $i \neq j$. For example, if $n = 5$ and the array elements are

$$A[1] = 1, A[2] = 2, A[3] = 3, A[4] = 4, A[5] = 5$$

then there are no swaps, whereas if the array elements are

$$A[1] = 2, A[2] = 1, A[3] = 5, A[4] = 4, A[5] = 3$$

then there are two swaps (1,2) and (3,5).

- (a) (5 points) What is the probability that 1 and 2 are swapped, that is, that $A[1] = 2$ and $A[2] = 1$?

There are $n!$ possible permutations, and $(n-2)!$ of them have $A(1) = 2$ and $A(2) = 1$. This means

$$\mathbb{P}((1, 2) \text{ are a swap}) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}.$$

There was nothing special about 1 and 2 in this calculation.

- (b) (10 points) Suppose that q is the correct answer to part (a). What is the expected number of swaps in the array (as a function of n and q)?

For any $\{i, j\} \subset \{1, \dots, n\}$, the probability that i and j are swapped is the same as above. Let's write $I_{i,j}$ for the indicator that i and j are swapped, and N for the total number of swaps, so that

$$\begin{aligned} \mathbb{E}[N] &= \mathbb{E}\left[\sum_{\{i,j\} \subset \{1,\dots,n\}} I_{i,j}\right] \\ &= \sum_{\{i,j\} \subset \{1,\dots,n\}} \mathbb{P}((i, j) \text{ are a swap}) \\ &= \binom{n}{2} \frac{1}{n(n-1)} = \frac{1}{2}. \end{aligned}$$

This is equal to

$$\binom{n}{2} q.$$

Task 6 – No partial credit on this task...**[35 pts]**

Write any calculations for this problem on scratch paper. True/False parts worth 2 points each; **circle one of True or False**.

1. **True or False:** The variance of a random variable is always nonnegative (that is, ≥ 0).

True

2. **True or False:** For any random variable X , the quantity $\mathbb{E}[X^2]$ is always nonnegative (that is, ≥ 0).

True. By the definition of expectation, $\mathbb{E}[X^2] = \sum_x x^2 \mathbb{P}(X = x)$ (if X is discrete) or $\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx$ (if X is continuous). For the discrete case, each term in the sum is nonnegative since $x^2 \geq 0$ and $\mathbb{P}(X = x) \geq 0$. For the continuous case, the function $x^2 f_X(x)$ is nonnegative for all x since $x^2 \geq 0$ and $f_X(x) \geq 0$. Therefore, the sum/integral must be nonnegative, and $\mathbb{E}[X^2] \geq 0$.

3. **True or False:** If $X \sim \text{Binomial}(n, p)$, then $2X \sim \text{Binomial}(2n, p)$.

False: $\Omega_{2X} = \{0, 2, 4, \dots, 2n\}$, but $\Omega_{\text{Binomial}(2n, p)} = \{0, 1, 2, \dots, 2n\}$

4. **True or False:** If X is a discrete random variable that takes integer values that are all less than or equal to 100 and $\mathbb{E}[X] = 100$, then $\mathbb{P}(X = 100) = 1$.

True. We can write

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x \leq 100} x \cdot \mathbb{P}(X = x) \\ &= 100 \cdot \mathbb{P}(X = 100) + \sum_{x < 100} x \cdot \mathbb{P}(X = x) \\ &\leq 100 \cdot \mathbb{P}(X = 100) + 99 \cdot \sum_{x < 100} \mathbb{P}(X = x) \\ &= 100 \cdot \mathbb{P}(X = 100) + 99 \cdot (1 - \mathbb{P}(X = 100)) \\ &= \mathbb{P}(X = 100) + 99. \end{aligned}$$

Since $\mathbb{E}[X] = 100$, we must have $\mathbb{P}(X = 100) = 1$.

5. **True or False:** Suppose that x_1, \dots, x_n are independent samples from a Poisson distribution with unknown parameter θ . Consider an estimator $\hat{\theta}(x_1, \dots, x_n)$ that outputs $(x_1 + x_2 + x_3)/3$. Then $\hat{\theta}$ is an unbiased estimator for the true parameter of the given Poisson.

True: $\mathbb{E}[\hat{\theta}] = \frac{1}{3} (\mathbb{E}[x_1] + \mathbb{E}[x_2] + \mathbb{E}[x_3]) = \theta$

6. **True or False:** Suppose that x_1, \dots, x_n are independent samples from a geometric distribution with unknown parameter θ . Consider an estimator $\hat{\theta}(x_1, \dots, x_n)$ that outputs $(x_1 + x_2 + x_3)/3$. Then $\hat{\theta}$ is an unbiased estimator for the true parameter of the given Geometric distribution.

False: $\mathbb{E}[\hat{\theta}] = \frac{1}{3} (\mathbb{E}[x_1] + \mathbb{E}[x_2] + \mathbb{E}[x_3]) = 1/\theta$

7. **True or False:** Suppose that $X \sim N(\mu, \sigma^2)$ and $Y = cX$. Then $\mathbb{E}[X + Y] = (1 + c)\mu$ and $\text{Var}(X + Y) = (1 + c^2)\sigma^2$.

False: X and Y are not independent, so linearity of variance doesn't hold.

In fact, $\text{Var}(X + Y) = \text{Var}((1 + c)X) = (1 + c)^2 \text{Var}(X) = (1 + c)^2 \sigma^2 \neq (1 + c^2)\sigma^2$

8. (6 points) Multiple choice: Suppose that random variable X has a normal distribution with mean 5 and variance 36. What is the probability that X takes a value between 11 and 17? (Recall that $\Phi(z) = \mathbb{P}(Z \leq z)$ where $Z \sim N(0, 1)$.) **Circle the correct answer.**

- (a) $\Phi(36) - \Phi(22)$
 - (b) $\Phi(2) - (1 - \Phi(1))$.
 - (c) $\Phi(2) - \Phi(1)$.
 - (d) $\Phi(2) - \Phi(-1) - \Phi(1) + \Phi(-2)$
- (c)

$$\text{Pr}(11 \leq X \leq 17) = \text{Pr}\left(\frac{11 - 5}{6} \leq Z \leq \frac{17 - 5}{6}\right) = \text{Pr}(1 \leq Z \leq 2)$$

9. Suppose that T_1, T_2, T_3, \dots are independent exponential random variables each with *expected value* equal to 10. Let

$$\bar{T} = \frac{1}{100} \sum_{i=1}^{100} T_i$$

- (a) (4 points) What is $\mathbb{E}[\bar{T}]$?

By linearity of expectation, we have $\mathbb{E}[\bar{T}] = \frac{1}{100} \cdot 100 \cdot \mathbb{E}[T_1] = 10$.

- (b) (4 points) What is $\text{Var}(\bar{T})$?

Since T_i is an exponential random variable with expected value $1/\lambda = 10$, we know the variance of T_i is $1/\lambda^2 = 100$. Also, since the T_i 's are independent, by the linearity of variance, we have

$$\text{Var}[\bar{T}] = \frac{1}{100^2} \cdot 100 \cdot \text{Var}[T_1] = 1.$$

(c) (7 points) Multiple Choice: Use the Central Limit Theorem to estimate $\mathbb{P}(\bar{T} > 7)$. **Circle the correct answer.**

(a) $\Phi(3)$

(b) $\Phi(-3)$

(c) $1 - \Phi(-12)$

(d) $\Phi(-12)$

(e) $\Phi(-2)$

(a) $\mathbb{P}(\bar{T} > 7) = \mathbb{P}\left(\frac{\bar{T}-10}{1} > \frac{7-10}{1}\right) = \mathbb{P}(Z > -3) = \Phi(3).$

Task 7 – One more

[16 pts]

There are n companies, including Google, Facebook, and $n - 2$ others. Four people – Eleos, Melissa, Khanh and Zhihao – each independently pick two of these companies uniformly at random and apply for an internship at those two companies. Any applicant at any of these companies independently gets called for an interview with probability p . What is the probability that at least one of Eleos, Melissa, Khanh and Zhihao is interviewed at Google?

There are many ways to solve this problem. Here is one of them.

The probability that at least one of them is interviewed is $1 -$ probability none of the four is interviewed.

Probability Eleos is interviewed is the probability he applies multiplied by the probability he is interviewed given that he applies, which is $\frac{n-1}{\binom{n}{2}} \cdot p = 2p/n$.

So the probability that Eleos is not interviewed is $1 - 2p/n$. This is the same as the probability that Melissa is not interviewed, etc.

So the probability that none of the four is interviewed is $(1 - 2p/n)^4$ and the final answer is $1 - (1 - 2p/n)^4$.