Review for final

slido.com/2167587

Distns (for your cheat sheet)

Discrete	Continuous
Uniform	Uniform (cont)
Bernoulli	Exponential
Binomial	Normal
Geometric	
Poisson	

- I will add solutions to pset 8 to google drive tomorrow
- Office hours Sunday – Tuesday

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

CSE 312: Foundations of Computing II Winter 2023

Final Exam - This is what it will look like!

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

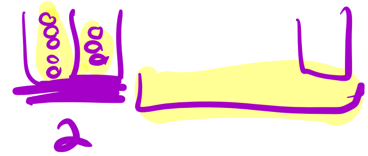
Instructions:

- **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 additinoal space.

Good luck!

Task 1 –

1. Suppose that 100 distinct balls are thrown independently and uniformly at random into 100 distinct bins. What is the probability that bin 1 has 5 balls in it given that bin 2 has 3 balls in it? (Use the following notation: Let B_i be the number of balls in bin i.)



$$P(B_1=5 \mid B_2=3) = \frac{P(B_1=5 \cap B_2=3)}{P(B_2=3)}$$

$$= \frac{\binom{100}{5}\binom{95}{3}\left(\frac{1}{100}\right)^8\left(\frac{98}{100}\right)^{92}}{\binom{100}{3}\left(\frac{1}{100}\right)^3\left(\frac{99}{100}\right)^{97}}$$

Answer: $$\frac{\binom{100}{5}\binom{95}{3}(98)^{92}}{\binom{100}{3}(99)^{97}}$$

2. Every minute, a random word generator spits out one word uniformly at random from the 3-word set {I, love, to}. The word spit out is independent of words spit out at other times. Let X be the number of times that the phrase "I love to love" appears if we let the generator run for n minutes. What is $\mathbb{E}[X]$?

$\frac{\text{I}}{n-3}\ \frac{\text{love}}{n-2}\ \frac{\text{to}}{n-1}\ \frac{\text{love}}{n}$

$$X_i = \begin{cases} 1 & \text{if "I love to love" appears starting at word } i \\ 0 & \text{o.w.} \end{cases}$$

$$X = \sum_{i=1}^{n-3} X_i$$

$$E(X) \underset{\text{LOE}}{=} \sum_{i=1}^{n-3} E(X_i)$$

$$E(X_i) = \left(\frac{1}{3}\right)^4$$

$\frac{\text{I}}{1/3}\ \frac{\text{love}}{1/3}\ \frac{\text{to}}{1/3}\ \frac{\text{love}}{1/3}$

Answer: $(n-3)\left(\frac{1}{3}\right)^4$

0 1 2 S=s 100

Task 2 – Law of total expectation

Sinho and Vrettos are playing a game where they each choose an integer uniformly at random from $[0, 100]$, then whoever has the larger number wins (and it's a draw if they tie). However, Vrettos doesn't like losing, so he's rigged his random number generator such that it instead picks uniformly at random from the integers that are between Sinho's number and 100, inclusive. Let S be Sinho's number and V be Vretto's number.

1. What is $\mathbb{E}[S]$?

   $S \sim \text{Unif}\{0,1,\dots,100\}$

   $\mathbb{E}[S] = \frac{100}{2}$

2. What is $\mathbb{E}[V|S = s]$, where s is any integer such that $0 \leqslant s \leqslant 100$?

   $\text{Unif}\{s,\dots,100\}$

   $\mathbb{E}[V|S = s] = \frac{s+100}{2}$

3. What is $\mathbb{E}[V]$? (You can leave your answer as a sum.)

   $\sum_{s=0}^{100} s = \frac{100 \cdot 101}{2}$

   $E(V) = \sum_{s=0}^{100} \underbrace{E(V|S=s)}_{\frac{s+100}{2}} \underbrace{P(S=s)}_{\frac{1}{101}}$

   $= 75$

   $\mathbb{E}[V] =$

4. What is the probability that the game ends in a draw?(You can leave your answer as a sum)

   $P(\text{draw}) = \sum_{s=0}^{100} \underbrace{P(\text{draw}|S=s)}_{\frac{1}{100-s+1}} \underbrace{P(S=s)}_{\frac{1}{101}}$

   s ⋯ 100

   $\mathbb{P}(\text{draw}) =$

X 1 p

-1 $(1-p)$

$E(X^2) = 1^2 p + (-1)^2 \cdot (1-p)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \cap B) = P(A) + P(B) - P(A \cup B)$

Task 3 – No partial credit on this task...

Write any calculations for this problem on scratch paper. **Circle one of True or False.**

$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$, $a = -4$, $b = 1$

1. **True or False:**

$$\sum_{k=0}^{n} \binom{n}{k} (-4)^k = (-3)^n.$$

$(1)^{n-k}$

$(-4+1)^n$

2. **True or False:** Let A and B be events such that $\mathbb{P}(A) = 0.5$, $\mathbb{P}(B) = 0.6$ and $\mathbb{P}(A \cup B) = 0.8$. Then the events A and B are independent.

$P(A \cap B) \overset{?}{=} P(A)P(B)$

$P(A)P(B) = 0.5 \cdot 0.6 = 0.3$

$P(A \cap B) = 0.5 + 0.6 - 0.8 = 0.3$

3. **True or False:** If X and Y are independent random variables, each taking values in $\{+1, -1\}$, then

$$\mathbb{E}\left[(X-Y)^2\right] = 2 - 2\,\mathbb{E}[X] \cdot \mathbb{E}[Y].$$

$X^2, Y^2 = 1$

$E(X^2) = E(Y^2) = 1$

$E[(X-Y)^2] = E(X^2 - 2XY + Y^2) = 2 - 2E(XY)$

4. **True or False:** For any random variable X, $E(5^X) = 5^{E(X)}$.

$E(g(X)) \neq g(E(X))$

5. **True or False:** Let X be a normal random variable with parameters μ (mean) and σ^2 (variance). Then,

$$\mathbb{P}(|X - \mu| \geqslant k\sigma) = 1 - \Phi(k) + \Phi(-k).$$

$k > 0$

$P(|X - \mu| \geq k\sigma)$

$= P\left(\frac{|X-\mu|}{\sigma} \geq k\right) = P(|Z| \geq k) = P(Z \geq k) + P(Z \leq -k) =$

6. Multiple Choice: Suppose that $X_1, \ldots, X_n$ are i.i.d. samples from a normal distribution with unknown mean μ and variance 9. How big does n need to be so that μ is in $[\overline{X} - 0.03, \overline{X} + 0.03]$ with probability at least 0.97. (As usual $\overline{X} = (\sum_{i=1}^{n} X_i)/n$.) You should use the fact that $\Phi^{-1}(0.985) = 2.17$. **Circle the correct answer.**

(a) $n \geqslant \frac{9}{0.03} \cdot 2.17$.

(b) $n \geqslant \left(\frac{3}{0.03} \cdot 2.17\right)^2$.

(c) $n \geqslant \left(\frac{0.03}{3} \cdot 2.17\right)^2$.

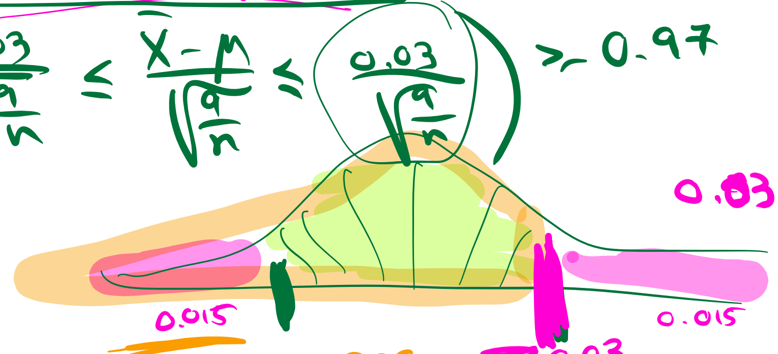
(d) $n \geqslant \left(\frac{0.03}{3} \cdot \frac{1}{2.17}\right)^2$.

$P(\mu \in \overline{X} \pm 0.03) \geq 0.97$

$P(|\overline{X} - \mu| \leq 0.03)$

$P(\mu - 0.03 \leq \overline{X} \leq \mu + 0.03) \geq 0.97$

$P\left(-\frac{0.03}{\sqrt{\frac{9}{n}}} \leq \frac{\overline{X} - \mu}{\sqrt{\frac{9}{n}}} \leq \frac{0.03}{\sqrt{\frac{9}{n}}}\right) \geq 0.97$



$\Phi\left(\frac{0.03}{\sqrt{\frac{9}{n}}}\right) \geq 0.985$

$E(\overline{X}) = \mu$

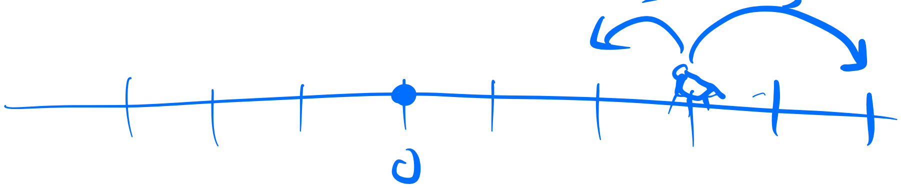
$Var(\overline{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$

$= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{9}{n}$

$n\,9$

$\sqrt{n} \geq \frac{2.17}{0.03} \cdot 3$

$\frac{1}{3}$ $\frac{2}{3}$



$$\frac{0.03}{\sqrt{\frac{9}{n}}} \geq 2.17$$

Task 5 – Random Frog

A frog takes a random walk on an infinite (discrete) line, where coordinates correspond to all integers (positive and negative), i.e., the elements of $\mathbb{Z}$, with the numbers ordered as usual increasing to the right and decreasing to the left.

- The frog starts at the coordinate $x = 0$.
- At every step, the frog jumps 2 to the right with probability $2/3$, and jumps 1 to the left with probability $1/3$. That is, the frog jumps from x to x' where $x' - x$ is 2 with probability $2/3$ and -1 with probability $1/3$.

Let X be the position of the frog after 200 steps.(Note that each jump is independent of all other jumps.)

(NOTE: One could model the frog's position as a Markov chain, but you should NOT use the language of Markov chains here! Remember to briefly justify your answers.)

1. Use linearity of expectation to compute $\mathbb{E}[X]$.

X_i: 2 w.p. $\frac{2}{3}$, -1 w.p. $\frac{1}{3}$

$$E(X_i) = 2\cdot\frac{2}{3} + (-1)\frac{1}{3} = 1$$

$$P_{X_i}(x) = \begin{cases} \frac{2}{3} & x = 2 \\ \frac{1}{3} & x = -1 \\ 0 & \text{o.w.} \end{cases}$$

$$X = \sum_{i=1}^{200} X_i$$

$$\boxed{\mathbb{E}[X] = \sum_{i=1}^{200} E(X_i) = 200}$$

2. What is $\mathrm{Var}(X)$?

$$\mathrm{Var}(X) = \sum_{i=1}^{200} \mathrm{Var}(X_i)$$

$$\mathrm{Var}(X_i) = 3 - E(X_i)^2 = 2$$

$$E(X_i^2) = (2)^2\frac{2}{3} + (-1)^2\frac{1}{3} = 4\cdot\frac{2}{3} + 1\cdot\frac{1}{3} = 3$$

$$\boxed{\mathrm{Var}(X) = 200 \cdot 2 = 400}$$

3. Use the CLT, including the continuity correction, to estimate the probability that after 200 steps the frog is at a position that is 250 or larger, i.e., to estimate $\mathbb{P}(X \geqslant 250)$. Write your result as a function of Φ. (You do NOT need to simplify numerical values.)

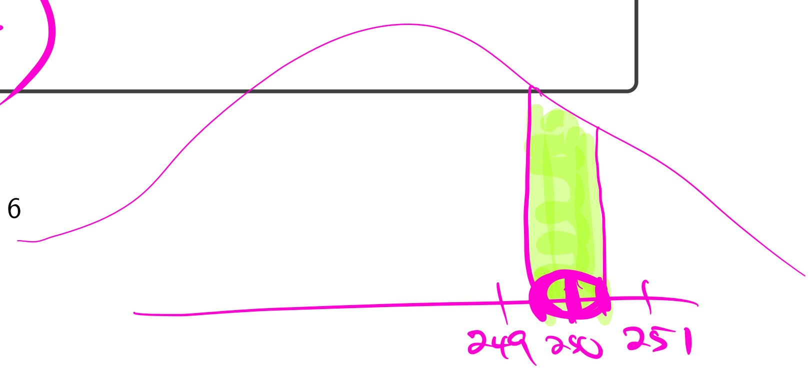
$$X \approx N(200, 400)$$

$$Y \sim N(200, 400)$$

$$P(X \geq 250) \approx P(Y \geq 249.5)$$

$$= P\left(\frac{Y-200}{\sqrt{400}} \geq \frac{249.5-200}{20}\right) = 1 - \Phi\left(\frac{49.5}{20}\right)$$

$$\boxed{\mathbb{P}(X \geqslant 250) = 1 - \Phi\left(\frac{49.5}{20}\right)}$$



Task 4 – MLE

Let $x_1, \dots, x_n$ be i.i.d. samples from a random variable that follow a so-called Borel distribution with unknown parameter θ, i.e., a distribution from the family

$$P(X=k;\theta) = \mathbb{P}(k;\theta) = \frac{e^{-\theta k}(\theta k)^{k-1}}{k!},$$

where $0 < \theta \leqslant 1$ is a real number, and $k \geqslant 1$ is an integer. (Note that the values Borel random variables take are positive integers.) What is the maximum likelihood estimator for θ?

$$L(x_1,\dots,x_n;\theta) = \prod_{i=1}^{n} P(x_i;\theta) = \prod_{i=1}^{n} \frac{e^{-\theta x_i}(\theta x_i)^{x_i-1}}{x_i!}$$

$$\ln L = \sum_{i=1}^{n} \ln \big(\quad \big)$$

$$= \sum_{i=1}^{n} -\theta x_i + \sum_{i=1}^{n} (x_i - 1)\ln\theta + \sum_{i=1}^{n} (x_i-1)\ln x_i - \sum_{i=1}^{n} \ln(x_i!)$$

$$\frac{d \ln L}{\partial \theta} = -\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} (x_i - 1)\frac{1}{\theta} = 0$$

$\hat{\theta} =$

This quarter - we made it!

- You learned foundations of probability & some stats

I recommend 421, 422, 427 (CB), 446, 447 (NLP), 426 (crypto), 431

493Q quantum computation

- You learned about some applications
  - Naive Bayes spam filtering
  - Bloom filters
  - Distinct elts & Minhash
  - Markov chains & Pagerank
  - auctions

- You learned some Python

  great headstart for 446

- Many of you learned LaTex

- You asked a lot of questions

Special shout-out to our wonderful TAs!

Thanks for a great quarter!