☐ **Write your name and PID on the top of EVERY PAGE.**

☐ **Write the solutions to each problem on separate pages. CLEARLY INDICATE on the top of each page the number of the corresponding problem. Different parts of the same problem can be written on the same page (for example, part (a) and part (b))**

☐ **Remember this exam is graded by a human being. Write your solutions NEATLY AND COHERENTLY, or they risk not receiving full credit.**

☐ **From the moment you access the midterm problems on Gradescope you have 70 MINUTES to COMPLETE AND UPLOAD your exam to Gradescope. Plan your time accordingly.**

☐ **You are allowed to use the textbook, lecture notes and your personal notes. You are not allowed to use the electronic devices (except for accessing the online version of the textbook) or outside assistance. Outside assistance includes but is not limited to other people, the internet and unauthorized notes.**

1. (25 points) A research company conducted a phone poll in the city of San Diego to study the music tastes of the city residents. The interviewees were asked to choose which music genre they prefer: rock or rap. Out of 2500 interviewed residents, 2000 replied that they prefer rap.

   Using the above data, construct a 92% confidence interval for the unknown fraction of the population that prefers rap. Clearly state (in words) your conclusion about the fraction of rap fans in the population.

   [You may leave your answer in terms of $\Phi(x)$]

   Solution. Let p denote the fraction of the population that prefers rap music. Then we use the normal approximation of the binomial distribution to construct the confidence interval
$$P(|p - \widehat{p}| < \varepsilon) \geq 0.92, \tag{1}$$
   where $\widehat{p} = 2000/2500 = 0.8$. It has been shown in Lecture 14 that
$$P(|p - \widehat{p}| < \varepsilon) \geq 2\Phi(2\varepsilon\sqrt{n}) - 1, \tag{2}$$
   where n is the number of experiments (number of interviewed San Diegans). In order to find the smallest interval (the smallest $\varepsilon > 0$) that guarantees (1), we take ε such that
$$2\Phi(2\varepsilon\sqrt{n}) - 1 = 0.92. \tag{3}$$
   Solving (3) with respect to ε gives
$$\varepsilon = \frac{\Phi^{-1}(0.96)}{100}. \tag{4}$$
   We conclude that with probability at least 0.92, the actual fraction of San Diegans who prefer rap music is in the interval
$$\left(0.8 - \frac{\Phi^{-1}(0.96)}{100}, 0.8 + \frac{\Phi^{-1}(0.96)}{100}\right). \tag{5}$$

2. (25 points) Felix has one dollar and a biased coin. Tossing this biased coin results in heads with probability $p > 1/2$.

   Felix plays a game with the following rules: (i) he tosses the coin until he sees heads, and (ii) after each toss his fortune (money he possesses before tossing the coin) doubles.

   (a) (15 points) Compute the expected fortune of Felix at the end of the game.

   (b) (10 points) Compute the probability that Felix wins more than 20 dollars.

   [For full credit, present your answer in the closed form (not as an infinite series)]

   Solution.

(a) Let X denote the number of the tosses it takes Felix to see heads for the first time. Then X has geometric distribution with parameter p. Since at the beginning of the game Felix has one dollar, and his fortune doubles after each toss, at the end of the game Felix has 2^X dollars. Using the formula for the expectation of a function of a discrete random variable from Lecture 10, we compute Felix's expected fortune after the game

$$E(2^X) = \sum_{k=1}^{\infty} 2^k (1-p)^{k-1} p = 2p \sum_{\ell=0}^{\infty} (2(1-p))^{\ell} = \frac{2p}{1-2(1-p)} = \frac{2p}{2p-1}. \qquad (6)$$

(b) Felix wins more than 20 dollars if and only if he tosses the coin at least 5 times, so

$$P(\text{wins more than 20 dollars}) = P(X \geq 5) = \sum_{k=5}^{\infty} p(1-p)^{k-1} = (1-p)^4. \qquad (7)$$

3. (25 points) A study showed that 2% of San Diego residents own a boat.

(a) (15 points) Estimate the probability that among 100 randomly interviewed San Diego residents there are at least 3 boat owners.

(b) (10 points) Explain why the approximation that you used in part (a) is better compared to other approximations that you know.

[For full credit, present your answer in the closed form (not as an infinite series); you may leave your answer in terms of e^x or $\Phi(x)$]

Solution.

(a) If X is the number is interviewed residents of San Diego that own boat. Then $X \sim \text{Bin}(100, 0.02)$. To compute the probability $P(X \geq 3)$ we first rewrite it as

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - P(X=0) - P(X=1) - P(X=2), \qquad (8)$$

and then use the Poisson approximation of the binomial distribution with $\lambda = 100 \cdot 0.02 = 2$

$$P(X \geq 2) \approx 1 - e^{-2} - 2e^{-2} - \frac{2^2}{2} e^{-2} = 1 - 5e^{-2}. \qquad (9)$$

(b) Using the criterion from Lecture 14, compute the bounds for the approximation error for normal and Poisson approximations of $\text{Bin}(100, 0.02)$. The error in the Poisson approximation is bounded by

$$np^2 = 100 \cdot (0.02)^2 = 0.04, \qquad (10)$$

while the error in the normal approximation is bounded by

$$\frac{1}{\sqrt{np(1-p)}} = \frac{1}{100 \cdot 0.02 \cdot 0.98} \approx 0.5. \tag{11}$$

By comparing the above numbers it is clear that the Poisson distribution ensures better approximation.

4. (25 points) Let X be a continuous random variable with density

$$f_X(x) = \begin{cases} xe^{-x}, & x > 0. \\ 0, & x \le 0. \end{cases}$$

   (a) (15 points) Compute the moment generating function $M_X(t)$. Indicate for which values $t \in \mathbb{R}$ function $M_X(t)$ is finite.

   (b) (10 points) Compute the moments of X using the power series expansion of $M_X(t)$. [Hint: use

$$\sum_{k=0}^{\infty} a_k t^k \cdot \sum_{\ell=0}^{\infty} b_\ell t^\ell = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} a_k\, b_\ell\, t^{k+\ell}$$

   and the geometric series formula]

Solution.

(a) We compute the moment generating function using the definition $M_X(t) = E(e^{tX})$

$$M_X(t) = \int_0^\infty e^{tx} x e^{-x} dx = \int_0^\infty x e^{(t-1)x} dx. \tag{12}$$

This integral diverges if $t \ge 1$, and for $t < 1$ using the integration by parts we get

$$M_X(t) = \frac{1}{t-1}\Big(xe^{(t-1)x}\Big|_0^\infty - \int_0^\infty e^{(t-1)x} dx \Big) = \frac{1}{(1-t)^2}. \tag{13}$$

So

$$M_X(t) = \begin{cases} \frac{1}{(1-t)^2}, & t < 1, \\ \infty, & \text{otherwise.} \end{cases} \tag{14}$$

(b) To compute the moments of X, write the Taylor series of $M_X(t)$ at $t = 0$ using the geometric series expansion $(1-t)^{-1} = \sum_{k=0}^\infty t^k$

$$M_X(t) = \frac{1}{1-t} \cdot \frac{1}{1-t} = \sum_{k=0}^{\infty} t^k \sum_{\ell=0}^{\infty} t^\ell = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} t^{k+\ell}. \tag{15}$$

By introducing the index $m = k + \ell$, (15) can be rewritten as

$$M_X(t) = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} t^{k+\ell} = \sum_{m=0}^{\infty} (m+1)t^m = \sum_{m=0}^{\infty} (m+1)! \frac{t^m}{m!}, \tag{16}$$

and we conclude that

$$E(X^m) = (m+1)!\,. \tag{17}$$