There are totally 37 points in the exam. The students with score higher than or equal to 35 points will receive 35 points. Please write down your name and student ID number below.

NAME: ________________________________
ID: ________________________________
1. (9 points). Fill in the blanks.

(a) (3 points). Let \( X_1, \ldots, X_n \sim iid \ N(\mu, 16) \) be a random sample. If \( n = 100 \) and \( \bar{x} = 3.1 \), then the 95% confidence interval for \( \mu \) is \( \ldots \), the 99% confidence interval for \( \mu \) is \( \ldots \). If we want the length of the 95% confidence interval less than or equal to 0.5, then we need \( n \) to be at least \( \ldots \).

(b) (3 points). Suppose \( x_1, \ldots, x_{100} \sim iid \ N(\mu, \sigma^2) \) be a random sample, with \( \bar{x} = 4 \) and \( s^2 = 9 \). Then, the 95% large sample confidence interval for \( \mu \) is \( \ldots \), the 99% large sample confidence interval for \( \mu \) is \( \ldots \). If we want to test \( H_0 : \mu = 2 \iff H_a : \mu \neq 2 \), then \( H_0 \) is rejected if \( \ldots \) at significance level \( \alpha = 0.05 \).

(c) (3 points). Let \( X \sim Bin(100, p) \). Suppose we observed \( x = 32 \). Then, the 95% confidence interval for \( p \) is \( \ldots \) and the 99% confidence interval for \( p \) is \( \ldots \). To make the 95% confidence interval less than or equal to 0.01, we need the sample size \( n \) at least \( \ldots \).

Answer: (a) [2.16, 3.884]; [2.068, 4.132]; 984. (b) [3.412, 4.588]; [3.226, 4.772]; \( \left| \frac{\bar{x} - 2}{s/\sqrt{100}} \right| > 1.96 \). (c) [0.2286, 0.4114]; [0.19976, 0.4404]; 38416.
2. (6 points) Assume $X_1 \sim N(2, 4), X_2 \sim N(3, 6)$ and $X_3 \sim N(-3, 9)$.

(a) (2 points). Compute $P(2X_1 + X_2 - X_3 < 16)$.

Solution: Let $Y = 2X_1 + X_2 - X_3$. Then $E(Y) = 2(2) + 3 - (-3) = 10$ and $V(Y) = 4(4) + 6 + 9 = 31$. Thus,

$$P(Y < 12) = \Phi\left(\frac{16 - 10}{\sqrt{31}}\right) = 0.8594.$$

(b) (2 points). Compute $P(|2X_1 - X_2 - X_3| < 7)$.

Solution: Let $Y = 2X_1 - X_2 - X_3$. Then, we have $E(Y) = 2(2) - 3 - (-3) = 4$ and $V(Y) = 4(4) + 6 + 9 = 31$. Thus, we have

$$P(|Y| < 7) = P(-7 < Y < 7) = \Phi\left(\frac{7 - 4}{\sqrt{31}}\right) - \Phi\left(\frac{-7 - 4}{\sqrt{31}}\right) = 0.6809.$$

(c) (2 points). Compute the density function (PDF) of $2X_1 + X_3$.

Solution: Let $Y = 2X_1 + X_3$. Then,

$$f(x) = \frac{1}{\sqrt{2\pi}\sqrt{25}} e^{-\frac{(x-1)^2}{2(25)}} = \frac{1}{\sqrt{2\pi}5} e^{-\frac{(x-1)^2}{25}}.$$

3. (6 points). Suppose $X_1, \ldots, X_n$ are iid with common PDF $f(x) = 2x/\theta^2$ for $0 \leq x \leq \theta$, with some $\theta > 0$. Let $X_{(n)} = \max(X_1, \ldots, X_n)$ be the sample maximum. Let $X_{(1)} = \min(X_1, \ldots, X_n)$ be the sample minimum.

(a) (2 points). Compute the CDF and the PDF of $X_{(n)}$ and $X_{(1)}$.

Solution: Let $F_n$ and $F_1$ be the CDF and $f_n$ and $f_1$ be the PDF of $X_{(n)}$ and $X_{(1)}$ respectively. Then, we have

$$F_n(x) = F^n(x) = (\frac{x^2}{\theta^2})^n = \frac{x^{2n}}{\theta^{2n}}$$

and

$$F_1(x) = 1 - [1 - F(x)]^n = 1 - (1 - \frac{x^2}{\theta^2})^n.$$

Moreover, we have

$$f_n(x) = f'_n(x) = \frac{2nx^{2n-1}}{\theta^{2n}}$$

and

$$f_1(x) = f'_1(x) = n(1 - \frac{x^2}{\theta^2})^{n-1} \frac{2x}{\theta^2}.$$

(b) (2 points). Compute $E(X_{(n)})$ and $V(X_{(n)})$.

Solution:

$$E(X_{(n)}) = \int_0^\theta x \frac{2nx^{2n-1}}{\theta^{2n}} dx = \frac{2n}{2n + 1} \theta.$$
and

\[ E(X^2_{(n)}) = \int_0^{\theta} x^2 \frac{2n x^{2n-1}}{\theta^{2n}} dx = \frac{2n}{2n+2} \theta^2. \]

Thus,

\[ V(X^2_{(n)}) = \left( \frac{2n}{2n+2} - \frac{(2n)^2}{(2n+1)^2} \right) \theta^2 = \frac{2n}{(2n+2)(2n+1)^2}. \]

(c) (2 point). Compute \( P(X_{(n)} > \theta - \alpha) \), for some \( \alpha \in (0, \theta) \) and guess the behavior of the probability for large \( n \).

Solution:

\[ P(X_{(n)} > \theta - \alpha) = 1 - F_n(\theta - \alpha) = 1 - \left( \frac{\theta - \alpha}{\theta^{2n}} \right)^{2n}. \]

The limit of the probability goes to 1 when \( n \to \infty \).

4. (6 points). Let \( \bar{X} = \sum_{i=1}^{n} X_i/n \).

(a) (3 points) Suppose \( X_1, \ldots, X_n \) are iid exponential distribution of mean \( \theta \), i.e, the density is \( f(x) = e^{-x/\theta}/\theta \) for \( x > 0 \), where \( \theta > 0 \) is a positive parameter. By CLT, approximately compute \( P(|\bar{X} - \theta| < 0.01 \theta) \) when \( n \) is 100, 10,000 and 1,000,000 respectively.

Solution: When the mean of exponential distribution is \( \theta \), the variance is \( \theta^2 / n \). Then \( E(\bar{X}) = \theta \) and \( V(\bar{X}) = \theta^2 / n \). Thus, by CLT,

\[ P(|\bar{X} - \theta| < 0.01 \theta) = P\left( -0.01 \theta < \bar{X} - \theta < 0.01 \theta \right) = \Phi\left( \frac{0.01 \theta}{\sqrt{\theta^2/n}} \right) - \Phi\left( -\frac{0.01 \theta}{\sqrt{\theta^2/n}} \right) = 2 \Phi(0.01 \sqrt{n}) - 1. \]

When \( n = 100 \), the value is \( 2 \Phi(0.1) - 1 = 0.0797 \); when \( n = 10,000 \), the value is \( 2 \Phi(1) - 1 = 0.6827 \); when \( n = 1,000,000 \), the value is \( 2 \Phi(10) - 1 = 1 \).

(b) (3 points) Suppose \( X_i \) follows \( \text{Poisson}(\mu) \) distribution. Let \( \mu \) and \( \sigma^2 \) be the common mean and variance of \( X_i \). Then, we have \( \mu = \sigma^2 = 2 \). By CLT, approximately compute \( P(\bar{X} \leq 2.2) \) when \( n = 25, 31 \). (Hint: use the continuity correction (0.5 shift)).

Solution: Here, we have \( E(\bar{X}) = 2 \) and \( V(\bar{X}) = 2/n \). Thus, when \( n = 5 \), we have

\[ P(\bar{X} \leq 2.2) = P\left( \sum_{i=1}^{25} X_i \leq 55 \right) = P\left( \sum_{i=1}^{25} X_i \leq 55.5 \right) \approx P(\bar{X} \leq 2.22) \approx \Phi\left( \frac{2.22 - 2}{\sqrt{2/25}} \right) = 0.7816. \]

When \( n = 8 \), we have

\[ P(\bar{X} \leq 2.2) = P\left( \sum_{i=1}^{31} X_i \leq 68.2 \right) = P\left( \sum_{i=1}^{31} X_i \leq 68.5 \right) \approx P(\bar{X} \leq 2.2097) \approx \Phi\left( \frac{2.2097 - 2}{\sqrt{2/31}} \right) = 0.7954. \]
5. (5 points). Suppose $X_1, \cdots, X_n$ are iid $\Gamma(\alpha, 1)$. Let $\mu$ and $\sigma^2$ be the common mean and variance.

(a) (1 point). Give the formulae of $\mu$ and $\sigma^2$ in term of $\alpha$.

Solution: In this case, $\beta = 1$. So we have $\mu = \alpha/\beta = \alpha$ and $\sigma^2 = \alpha/\beta^2 = \alpha$.

(b) (1 points). Show that $\bar{X}$ is an unbiased estimator of $\alpha$.

Solution: Since $E(\bar{X}) = \alpha$, $\bar{X}$ is an unbiased estimator of $\alpha$.

(c) (2 points). Give the MSE of $\bar{X}$ and show that it goes to 0 as $n \to \infty$.

Solution: Since $V(\bar{X}) = \frac{\alpha}{n}$, we have

$$MSE(\bar{X}) = V(\bar{X}) = \frac{\alpha}{n} \to 0$$

as $n \to \infty$.

(d) (1 point). If the data are 1.39, 1.07, 0.72, 0.91, 2.56, 2.54, 0.91, 1.13, 4.72, 2.29, compute the estimate value of $\alpha$.

Solution: In this case, $\bar{X} = 1.824$. Thus, the estimate value of $\alpha$ is 1.824.

6. (5 points). Suppose $X_1, \cdots, X_n$ are iid $\text{Uniform}[0, \theta]$, i.e., the common density is $f(x) = 1/\theta$ if $0 \leq x \leq \theta$. Let $X_{(n)} = \text{max}(X_1, \cdots, X_n)$.

(a) (2 points). Show that $X_{(n)}$ is a biased estimator of $\theta$ and compute the bias and the MSE.

Solution: The density of $X_{(n)}$ is $f_n(x) = nx^{n-1}/\theta^n$. It induces that $E(X_{(n)}) = n\theta/(n+1)$ and $V(X_{(n)}) = n/[(n+2)(n+1)^2]$. The MSE is

$$MSE(X_{(n)}) = \left(\frac{n\theta}{n+1} - \theta\right)^2 + V(X_{(n)}) = \frac{2}{(n+1)(n+2)}$$

(b) (2 points). If one wants to modify a new estimator of $\theta$ based on $X_{(n)}$ by multiplying a constant $c$ so that it is unbiased, what the value of $c$ do you suggest and find the MSE of the new estimator.

Solution: It is clear that $c = (n+1)/n$. In this case,

$$MSE(cX_{(n)}) = c^2V(X_{(n)}) = \frac{1}{n(n+2)}.$$ 

(c) (1 point). Which estimator is better.

Solution: Since the MSE of $cX_{(n)}$ is smaller, $\frac{(n+1)X_{(n)}}{n}$ is better.