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, student ID number below.

NAME: ________________________________
ID: ________________________________

This is the distribution of the midterm
7.5 13.0 13.0 16.5 17.0 18.0 18.0 20.0 21.0 21.0 21.5 22.0
22.5 23.0 23.0 23.0 23.5 24.5 26.0 26.5 27.0 27.5 28.0 28.0
28.5 28.5 28.5 29.0 29.5 29.5 30.0 30.5 30.5 31.0 31.0 31.0
31.0 31.5 31.5 31.5 32.0 32.0 32.5 32.5 32.5 33.0 33.0 33.5
34.0 34.0 34.5 35.0 35.0 35.0 35.0 35.0 35.0 35.0 35.0 35.0

This is the summary of the midterm
Min. 1st Qu. Median Mean 3rd Qu. Max.
7.50 23.00 29.50 27.83 32.62 35.00
1. (17 points, 1 point each). Fill in the blanks.

(a) (4 points). Suppose \( E(X_1) = 1.0, E(X_2) = 2.0, V(X_1) = 1.44, V(X_2) = 2.56, \) and \( \text{Cov}(X_1, X_2) = 0.768 \). Let \( Y = 1.5X_1 + 1.2X_2 \). Then, \( E(Y) = ________, V(Y) = ________, \) \( \text{Corr}(X_1, X_2) = ________, \) and \( \text{Cov}(X_1, Y) = ________. \)

(b) (4 points). Let \((X, Y)\) be a bivariate random variable with joint PDF \( f(x, y) = 3x^2 + 2xy, \) \( 0 < x, y < 1 \). Then, \( P(X \leq 0.8, Y \leq 0.8) = ________, \) \( E(X) = ________, \) \( V(X) = ________, \) and \( \text{Cov}(X, Y) = ________. \)

(c) (3 points). Let \( x_1, \cdots, x_{20} \) are independently observed from \( N(\mu, 16) \). Assume \( \bar{x} = 4.32 \). Then, the 95% confidence interval for \( \mu \) is ________ and the 99% confidence interval for \( \mu \) is ________. If we want the 99% confidence interval to be shorter than 0.5, we need the sample size \( n \) to be at least ________.

(d) (3 points). Let \( X \sim \text{Bin}(378, p) \). Suppose we observed \( x = 249 \). Then, the 95% confidence interval for \( p \) is ________ and the 99% confidence interval for \( p \) is ________. If we want to test \( H_0 : p = 0.5 \) against \( H_1 : p \neq 0.5 \) at 0.05 significance level, then we ________ the null hypothesis.

(e) (3 points). Let \( X_1, \cdots, X_n \) be identically and independently distributed with a common mean \( \mu = 1 \) and a common variance \( \sigma^2 = 0.36 \). Let \( a = P(\bar{X} \leq 1.02) \). Using the Central Limiting Theorem, if \( n = 10^3 \), then \( a \approx ________ \); if \( n = 10^4 \), then \( a \approx ________ \); and if \( n = 10^5 \), then \( a \approx ________ \).

Answer: (a) \( -0.9; 4.1616; 0.4; 1.2384 \). (b) \( 0.36864; 2/3; 1/18; 0 \). (c) \([2.567, 6.073]; [2.012, 6.628]; 1075 \). (d) \([0.6109, 0.7065]; [0.596, 0.722]; \) Reject. (e) 0.8541; 0.9996; and 1.0000.
2. (8 points). Derive the maximum likelihood estimator (MLE) in the following problems. You need to display your entire approach (i.e., the maximum likelihood estimation).

(a) (2 points). Let \(X_1, \ldots, X_n\) be identically and independently distributed of \(\text{Poisson}(\theta^2), \theta > 0\). Derive the MLE of \(\theta\).

\[ \text{Solution:} \] The loglikelihood function is

\[ \ell(\theta^2) = \log \prod_{i=1}^{n} \frac{\theta^{2X_i}}{X_i!} e^{-\theta^2} = -\sum_{i=1}^{n} \log X_i! + 2 \sum_{i=1}^{n} X_i \log \theta - n\theta^2. \]

Its derivative is

\[ \ell'(\theta) = \frac{2}{\theta} \sum_{i=1}^{n} X_i - 2n\theta \Rightarrow \hat{\theta} = \bar{X}^{1/2}. \]

Thus, the MLE is \(\hat{\theta} = \bar{X}^{1/2}\).

(b) (2 points). Let \(X_1, \ldots, X_n\) be identically and independently distributed with \(\text{N}(0, \theta), \theta > 0\). Derive the MLE of \(\theta\).

\[ \text{Solution:} \] The loglikelihood function is

\[ \ell(\theta) = \sum_{i=1}^{n} \log \left( \frac{1}{\sqrt{2\pi \theta}} \right) e^{-\frac{X_i^2}{2\theta}} = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \theta - \frac{1}{2\theta} \sum_{i=1}^{n} X_i^2. \]

Its derivative is

\[ \ell'(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^{n} X_i^2 \Rightarrow \hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} X_i^2. \]

Thus, the MLE is \(\hat{\theta} = \sum_{i=1}^{n} X_i^2 / n\).

(c) (2 points). Let \(X_1, \ldots, X_n\) be identically and independently distributed with a common PDF \(f_\theta(x) = \theta^2 x e^{-\theta x}, x, \theta > 0\). Derive the MLE of \(\theta\).

\[ \text{Solution:} \] The loglikelihood function is

\[ \ell(\theta) = \sum_{i=1}^{n} \log(\theta^2 X_i e^{-\theta X_i}) = \sum_{i=1}^{n} \log X_i + 2n \log \theta - \theta \sum_{i=1}^{n} X_i. \]

Its derivative is

\[ \ell'(\theta) = \frac{2n}{\theta} - \sum_{i=1}^{n} X_i \Rightarrow \hat{\theta} = \frac{2}{\bar{X}}. \]

Thus, the MLE is \(\hat{\theta} = 2 / \bar{X}\).

(d) (2 points). Let \(X_1, \ldots, X_n\) be identically and independently distributed with a common PDF \(f_\theta(x) = (\theta + 1)x^\theta, x \in (0, 1)\) and \(\theta > -1\). Derive the MLE of \(\theta\).

\[ \text{Solution:} \] The loglikelihood function is

\[ \ell(\theta) = \sum_{i=1}^{n} \log[(\theta + 1)^n X_i^\theta] = n \log(\theta + 1) + \theta \sum_{i=1}^{n} \log X_i. \]

Its derivative is

\[ \ell'(\theta) = \frac{n}{\theta + 1} + \sum_{i=1}^{n} \log X_i \Rightarrow \hat{\theta} = 1 - \frac{n}{\sum_{i=1}^{n} \log X_i}. \]

Thus, the MLE is \(\hat{\theta} = 1 - n / \sum_{i=1}^{n} \log X_i\).
3. (6 points). The following problems are related to the concepts of bias and mean square error (MSE).

(a) (2 points). Let $X_1, \cdots, X_n$ be identically and independently distributed of $N(\theta, 10)$. Compute the bias and the MSE of $\hat{\theta} = \bar{X}$.

**Solution:** Since $\bar{X} \sim N(\theta, 10/n)$, $E(\hat{\theta}) = \theta$ and $V(\hat{\theta}) = 10/n$. Therefore,

$$\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta = 0$$

and

$$\text{MSE}(\hat{\theta}) = V(\hat{\theta}) + \text{Bias}^2(\hat{\theta}) = V(\hat{\theta}) = 10/n.$$ (b) (2 points). Let $X \sim \text{Bin}(n, \theta)$. Compute the bias and the MSE of $\hat{\theta} = X/n$.

**Solution:** Since $E(X) = n\theta$ and $V(X) = n\theta(1 - \theta)$, $E(\hat{\theta}) = \theta$ and $V(\hat{\theta}) = \theta(1 - \theta)/n$. Thus,

$$\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta = 0$$

and

$$\text{MSE}(\hat{\theta}) = V(\hat{\theta}) + \text{Bias}^2(\hat{\theta}) = V(\hat{\theta}) = \theta(1 - \theta)/n.$$ (c) (2 points). Let $X_1, \cdots, X_n$ be identically and independently distributed with a common PDF $3x^2/\theta^3$, $0 \leq x \leq \theta$ for some $\theta > 0$. Compute the bias and the MSE of $\hat{\theta} = \max_{i \leq n}(X_i)$.

**Solution:** The CDF of $X_I$ is $x^3/\theta^3$ for $0 \leq x \leq \theta$. For any $x \in (0, \theta)$, there is

$$F(x) = P(X \leq x) = \prod_{i=1}^n P(X_i \leq x) = \left( \frac{x^3}{\theta^3} \right)^n = \frac{x^{3n}}{\theta^{3n}}.$$ The PDF of $\hat{\theta}$ is

$$f(x) = F'(x) = \frac{3nx^{3n-1}}{\theta^{3n}}.$$ Then,

$$E(\hat{\theta}) = \int_0^\theta xf(x)dx = \int_0^\theta \frac{3nx^{3n}}{\theta^{3n}} dx = \frac{3n\theta}{3n+1}$$

and

$$E(\hat{\theta}^2) = \int_0^\theta x^2f(x)dx = \int_0^\theta \frac{3n^2x^{3n+1}}{\theta^{3n}} dx = \frac{3n^2\theta^2}{3n+2} = \frac{3n\theta^2}{3n+2}.$$ Then,

$$V(\hat{\theta}) = E(\hat{\theta}^2) - E^2(\hat{\theta}) = \frac{3n\theta^2}{3n+2} - \left( \frac{3n\theta}{3n+1} \right)^2 = \frac{3n\theta^2}{(3n+2)(3n+1)^2}.$$ Thus,

$$\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta = \frac{\theta}{3n+1}$$

and

$$\text{MSE}(\hat{\theta}) = V(\hat{\theta}) + \text{Bias}^2(\hat{\theta}) = \frac{3n\theta^2}{(3n+2)(3n+1)^2} + \frac{\theta^2}{(3n+1)^2} = \frac{2\theta^2}{(3n+2)(3n+1)}.$$
4. (6 points). Let $X_1, \ldots, X_{25} \sim N(\mu, 16)$. Consider a hypothesis testing problem for

$$ H_0 : \mu \geq 5 \leftrightarrow H_1 : \mu < 5. $$

Assume the rejection region is $C = \{ \bar{X} \leq 3.5 \}$.

(a) (2 points). Compute the type I error probability at $\mu = 6$.

**Solution:** Note that $\bar{X} \sim N(\mu, 16/25) = N(\mu, 0.64)$. The type I error probability at $\mu = 6$ is

$$ P(\bar{X} \leq 3.5|\mu = 6) = \Phi\left(\frac{3.5 - 6}{\sqrt{0.64}}\right) = \Phi(-3.125) = 0.0006. $$

(b) (2 points). Compute the type II error probability at $\mu = 3$.

**Solution:** The type II error probability at $\mu = 3$ is

$$ P(\bar{X} > 3.5|\mu = 3) = 1 - \Phi\left(\frac{3.5 - 3}{\sqrt{0.64}}\right) = 1 - \Phi(0.625) = 0.2660. $$

(c) (2 points). Compute the significance level of the test.

**Solution:** The significance level of the test is

$$ P(\bar{X} \geq 3.5|\mu = 5) = \Phi\left(\frac{3.5 - 5}{\sqrt{0.64}}\right) = \Phi(-1.875) = 0.0304. $$