

Statistical Methods

Program: M.Tech, Computer Science (1st and 2nd Year)

Academic Year: 2025-26, First Semester

End-Semester Exam, Date: 14 November, 2025

Duration: 3 hours, Total Points: 110 (including a buffer of 10)

This is a closed-book, closed-notes exam. Calculators are allowed. Please write your answers with concise reasoning. Unless specifically asked to prove, you may use any result that was proved or stated in class. In that case, you must clearly state the result that you are using. Answer as many questions as you can. However, **the maximum possible score is 100.**

1. A box contains 3 red and 2 blue balls. Two balls are drawn without replacement. Find the probability that both are red, given that at least one is red. [8]
2. Assume that X and Y are uncorrelated random variables having means μ_X and μ_Y , and variances σ_X^2 and σ_Y^2 , respectively. Find a condition under which $(2X + Y)$ and $(X - 3Y)$ are uncorrelated. [8]
3. Let X_1, X_2, X_3 be a sample of size $n = 3$ from a distribution with unknown mean μ , $-\infty < \mu < \infty$ where the variance σ^2 is a known positive number. Compare the mean-squared errors (MSEs) of $\hat{\theta}_1 = \bar{X}$ and $\hat{\theta}_2 = (2X_1 + X_2 + 5X_3)/8$. [8]
4. A study was conducted to see the differences between oxygen consumption rates for male runners from a college who had been trained by two different methods, one involving continuous training for a period of time each day and the other involving intermittent training of about the same overall duration. The sample sizes, sample means, and sample standard deviations are shown in the following table.

Continuous training: $n_1 = 15$, $\bar{x}_1 = 46.28$, $s_1 = 6.3$

Intermittent training: $n_2 = 7$, $\bar{x}_2 = 42.34$, $s_2 = 7.8$

If the measurements are assumed to come from normally distributed populations with equal variances, find a 95% confidence interval of the difference between the population means, and interpret. [6+2=8]

5. Assume that the joint probability density function (pdf) of (X, Y) is:

$$f(x, y) = \begin{cases} 2, & 0 < x < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(i) Find the marginal pdf of Y .

(ii) Compute the conditional expectation $E[X | Y = y]$, where $0 < y < 1$. [4+8=12]

6. Let $X_1, X_2, \dots, X_n$ be a random sample from a distribution with pdf

$$f(x; \alpha) = \frac{\alpha x_0^\alpha}{x^{\alpha+1}}, \quad x > x_0 > 0,$$

where x_0 is known. Find a method of moments estimator for α , and investigate if it is a sufficient statistic for α . [8+4=12]

7. Let $X_1, X_2, \dots, X_n$ be a random sample of size $n > 1$ from $N(\mu, \sigma^2)$ where both μ and σ are unknown. Given a specified number $\sigma_0^2 > 0$, we wish to test $H_0 : \sigma^2 = \sigma_0^2$ vs. $H_a : \sigma^2 \neq \sigma_0^2$ at level $\alpha = 0.05$. Find an appropriate likelihood ratio test (LRT) and completely specify the rejection region. [9+3=12]

8. Four different materials, A, B, C, and D, are tested for tensile strength (in kg/cm²). Below are the samples.

Material A: 45, 47, 46, 48

Material B: 40, 42, 41, 39

Material C: 43, 44, 45, 42

Material D: 38, 40, 37, 39

Determine if the materials differ significantly in their mean strength at 1% significance. Which assumptions did you have to make? [10+2=12]

9. Consider the linear model $y_i = \beta_0 + \beta_1(x_i - \bar{x}) + \varepsilon_i$, $i = 1, 2, \dots, n$, where ε_i are independent and identically distributed as $N(0, \sigma^2)$.
 (i) Find the least-square estimates of β_0 and β_1 .
 (ii) Provide an unbiased estimator for σ^2 . Verify that it is indeed unbiased. [6+6=12]
10. Let $X_1, X_2, \dots, X_n$ be a random sample from the uniform distribution on $[0, \theta]$, where $\theta > 0$ is an unknown parameter. Let $\hat{\theta}$ be the maximum likelihood estimator (MLE) of θ .
 (i) Find $\hat{\theta}$.
 (ii) Find c such that $c\hat{\theta}$ is an unbiased estimator of θ .
 (iii) Find k for which the mean squared error of $k\hat{\theta}$ is the smallest. [6+8+4=18]