INDIAN STATISTICAL INSTITUTE

B Stat I, 1st Semester, 2025-26

Statistical Methods I

Semestral Examination

Date: ~~27~~ 25.11.2025 Time: 3 hours

Answer all questions. Maximum you can score is 100.

1. Given numbers $x_1, x_2, \ldots, x_n$, the mean and the variance of the data set are defined as

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

respectively. Consider the following sample of five numbers: 1.0, 2.3, 3.0, 4.2, 1.5. Let their mean and variance be denoted as $x^\star$ and $s^{\star 2}$.

(a) Evaluate $x^\star$ and $s^{\star 2}$.

(b) Find five numbers such they have the same variance as the original set but have mean three units higher (i.e., mean is $x^\star + 3$).

(c) Find five numbers such that they have the same mean as the original set, but have a variance which is four times higher than the variance of the original set.

(d) Find five numbers such they have a mean which is 3 units higher than the mean of the original set and variance which is four times higher than the variance of the original set.

[3+5+5+5=18]

2. Suppose that a system of three components fails when any of the components fail. Assume that the lifetime of the components are independently distributed, each with the common probability density function

$$f(x) = \frac{1}{200}e^{-x/200}, \quad x > 0.$$

(a) Find the probability that the system lasts for at least 150 hours.

(b) Find the probability density function of the time till the system fails.

[8+10=18]

3. A random variable is said to have the gamma distribution with parameters α and θ if its density function is given by

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-x/\theta}, \quad x > 0.$$

Suppose that X and Y have independent gamma distributions with parameters (α, θ) and (β, θ), respectively. Show that

$$W = \frac{X}{X+Y}$$

has a beta distribution with parameters (α, β). [16]

4. Suppose $X_i \sim N(i\mu, 1)$, where the X_is are independent, $i = 1, 2, \ldots, n$. Derive a 90% confidence interval for μ based on $X_1, \ldots, X_n$.

   (a) Identify an appropriate pivotal element for finding a confidence interval for μ.

   (b) Find a 90% confidence interval for μ based on the above pivotial element.

[8+8=16]

5. Consider the usual linear regression model

$$Y_i = \alpha + \beta x_i + \epsilon_i, \quad i = 1, 2, \ldots, n,$$

where ϵ_i are independent, each having a $N(0, \sigma^2)$ distribution. Let the terms S_{xx}, S_{xY} and S_{YY} have their usual meanings. We know that the least square estimators of α and β given by

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}, \quad \hat{\beta} = \frac{S_{xY}}{S_{xx}}$$

are unbiased estimators of α and β respectively.

   (a) Derive $\mathrm{Var}(\hat{\beta})$, $\mathrm{Var}(\hat{\alpha})$ and $\mathrm{Cov}(\hat{\alpha}, \hat{\beta})$.

   (b) Derive the 95% confidence interval for the mean of the response variable Y for a new observation for which the value of the explanatory variable is x^*.

   (c) Compute the confidence interval in part (b) when $n = 10$, $\hat{\alpha} = -0.07926$, $\hat{\beta} = 0.40875$, $S_{xx} = 840.4$, $S_{YY} = 141.433$ and $S_{xY} = 343.514$ and $x^* = 10$.

[(2+2+3)+6+6=19]

6. We follow the notation used in class: $f(x, y)$ represents the joint pmf, $f_X(x)$, $f_Y(y)$ represent the indicated marginals, and $f_{Y|X}(y|x)$ represents the indicated conditional pmf. Let $f_X(x) = 1/10, x = 0, 1, \ldots, 9$ and $f_{Y|X}(y|x) = 1/(10-x), y = x, x+1, \ldots, 9$. Find

   (a) $f(x, y)$

   (b) $f_Y(y)$

   (c) $E(Y|X = x)$

[6+6+6=18]