STATISTICS

Paper I : Non-Parametric Methods and Regression Analysis

Note : Attepmt **all** questions from **Section-A** (Objective type questions), **five** questions from **Section-B** (Short answer type questions) and **three** questions from **Section-C** (Long/Essay type questions).

Section—A $1 \times 10 = 10/0{\cdot}5 \times 10 = 5$

1. If X_i ($i = 1 \ldots n$) are iid as $Np(\mu, \Sigma)$ then the distribution of $\overline{X} \frac{1}{n} \sum_{i=1}^{n} X_i$ is :

(a) $N\left(n\mu, \frac{\Sigma}{n}\right)$ (b) $N\left(\mu, \frac{\Sigma}{n}\right)$ (c) $N\left(\frac{\mu}{n}, \frac{\Sigma}{n}\right)$ (d) $N(n\mu, \Sigma)$.

2. The multivariate normal distribution is the extension of :

(a) Uniform distribution (b) Normal distribution

(c) Exponential distribution (d) Chi-square distribution.

3. The off-diagonal elements in the variance-covariance matrix contain :

(a) Regression parameter (b) Mean

(c) Covariance estimates (d) Variance estimates.

4. The CDF of $F_1(x)$ is given by :

(a) $[F(x)]^n$ (b) $1 - \{1 - F(x)^n$ (c) $[1 - F(x)]^n$ (d) $1 - [F(x)]^n$.

5. Relative efficiency in non-parametric distributions in the ratio of :

(a) Power of the tests (b) Size of two tests

(c) Size of the samples (d) All of the above.

6. In Mood's test one can test the hypothesis :

(a) $H_0 : \sigma_1 = \sigma_2$ Vs $H_1 : \sigma_1 \neq \sigma_2$ (b) $H_0 : \sigma_1 < \sigma_2$ Vs $H_1 : \sigma_1 > \sigma_2$

(c) $H_0 : \sigma_1 \geq \sigma_2$ Vs $H_1 : \sigma_1 < \sigma_2$ (d) All the above.

7. Kolmogorov-Smirnov test is a :

(a) One left sided test (b) One right sided test

(c) Two sided test (d) All of the above.

8. Coefficient of determination is given by :

(a) $1 - \frac{SSE}{TSS}$ (b) $1 - \frac{SSR}{TSS}$ (c) $1 - \frac{SSE}{SSR}$ (d) $1 - \frac{SSR}{SSE}$.

9. For $Y_{n \times 1} \sim \underset{n \times k,\, k \times p,\, n \times 1}{(X\,\beta, \varepsilon)}$, where $\varepsilon \sim N(0, \sigma^2 I)$, The MLE of σ^2 is :

(a) $\hat{\sigma}^2 = \frac{\Sigma e_1{}^2}{n-k}$ (b) $\hat{\sigma}^2 = \frac{\Sigma e_i{}^2}{n}$ (c) $\hat{\sigma}^2 = \frac{\Sigma e_1{}^2}{k}$ (d) $\hat{\sigma}^2 = \frac{\Sigma e_1{}^2}{n-1}$.

10. The test statistic for significance of all the explanatory variable is :

(a) $t = \frac{\hat{\beta} - \beta_i}{SE(\beta_i)}$ (b) $F = \frac{SSR/k-1}{TSS/n-1}$

(c) $F = \frac{SSR/k-1}{SSE/n-k}$ (d) $\psi^2 = \frac{(n-k)SSE}{\sigma_0{}^2}$.

Section—B $2 \times 5 = 10/1 \times 5 = 5$

1. If $X_1, X_2, \ldots, X_n$ are iid as $Np(\mu, \Sigma)$, then find the distribution of mean.

2. Let $X \sim Np(\mu, \Sigma)$, obtain its characteristics function.

3. Derive CDF for largest $X_{(n)}$ and lowest $X_{(1)}$ order statistic.

4. Let X_i ($i = 1, 2, \ldots n$) be a random sample from a population with continuous density. Show that $Y_1 = \text{Min}(X_1, X_2, \ldots, X_n)$ is exponential with parameter $n\lambda$ iff each X_i is exponential with parameter λ.

5. Describe sign test for paired samples.

6. Following are the yields of maize in quintal per hectare recorded from an experiment and arrnaged in ascending order with median $M = 20$.

Test $H_0 : M = 20$ Vs $H_1 : M \neq 20$ at $\alpha = 0{\cdot}05$ Given, $P(X \leq 4) = 0{\cdot}5$.

7. What are the assumptions for Non-parametric test ?

8. Explain Spearman's rank correlation test.

9. For a simple linear regression model $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $i = 1, 2, \ldots n$ where $\varepsilon_i \sim N(0, \sigma^2)$. Find an unbiased estimator of error variance (σ^2).

10. For a simple linear regression model, prove that :

(a) $\Sigma Y_i = \Sigma \hat{Y}_1$, (b) $\text{Cov}(\hat{\beta}_0, \hat{B}_1) = -\dfrac{\sigma^2}{\Sigma(X_i - \overline{X})^2}$.

Section—C $10 \times 3 = 30/5 \times 3 = 15$

1. If $X \sim Np(\mu, \Sigma)$, then show that :

$$Q = (X - \mu)\,\Sigma^{-1}(X - \mu) \sim \psi^2_p$$

2. $X_1, X_2, \ldots, X_n$ be n independent observations from $Np(\mu, \Sigma)$ and $A = \sum_{r=1}^{n} (X_r - \overline{X})(X_r - \overline{X})'$. Find the MLE of μ and Σ.

3. Define order statistics and obtain the distribution of Range $W = X_{(n)} - X_{(1)}$, where $X_{(n)}$ and $X_{(1)}$ are largest and lowest observations.

4. Explain Mann-Whitney-Wilcoxon test and also prove that :

$$U = T - \frac{n(n+1)}{2}$$

where U and T are drived under Mann-Whitney and Wilcoxon rule.

5. Stating the assumptions, carry out the test of significance for α and β separately in case of two variable linear model $Y_i = \alpha + \beta X_i + \varepsilon_i$ ($i = 1, 2, \ldots, n$),and construct the confidence interval for the same.