CSM – 68/21
Statistics
Paper – I

Time : 3 hours

Full Marks : 300

The figures in the right-hand margin indicate marks.

*Candidates should attempt Q. No. 1 from Section – A and Q. No. 5 from Section – B which are compulsory and any **three** of the remaining questions, selecting at least **one** from each Section.*

SECTION – A

1. Attempt any **three** of the following sub-parts : 20×3 = 60

(a) Define convergence in probability and almost sure convergence. In usual notations, show that :

(i) $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} Y \Rightarrow \frac{X_n}{Y_n} \xrightarrow{p} \frac{X}{Y}$.

(ii) $X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{p} X$.

(b) Define probability generating function and find the generating function of $P(X > n)$. By giving a counter example, show that the random variable X may have no moments although its moment generating function exists.

(c) Show that the correlation coefficient between the residual $X_{1.23}$ and $X_{2.13}$ is equal and opposite to that between $X_{1.3}$ and $X_{2.3}$.

(d) Let $X_1, X_2, \ldots\ldots, X_n$ represent a sample from the multivariate normal population with mean vector μ and covariance matrix Σ. Obtain the maximum likelihood estimates of μ and Σ.

2. (a) State and prove Borel Cantelli Lemma. Discuss briefly the strong law of large numbers and Kolmogorov's theorem.

(b) Two discrete random variables X and Y have the joint probability density function :

$$pxy(x, y) = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!}, \ y = 0, 1, 2, \cdots, x;$$

$x = 0, 1, 2, \cdots$, where λ, p are constants with $\lambda > 0$ and $0 < p < 1$. Obtain the conditional distribution of Y for a given X and of X for a given Y.

(c) For any continuous distribution, show that the mean deviation is least when measured from the median. $20 \times 3 = 60$

3. (a) Define Cauchy distribution. Obtain its characteristic function. If X is a random variable having a (standard) Cauchy distribution, find the p.d.f for X^2 and identify its distribution.

(b) Discuss beta variate of first kind. Obtain its mean and variance. Also define beta variate of second kind and state its relation with gamma variate.

(c) If $C_1'y$ and $C_2'y$ are two best linear unbiased estimators of $l_1'\beta$ and $l_2'\beta$, respectively, show that $C_1'y + C_2'y$ is best linear unbiased estimator of the same function of parametric vector. $20 \times 3 = 60$

4. (a) Let (X, Y) have the joint p.d.f. given by

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

Show that the regression of Y on X is linear but regression of X on Y is not linear.

(b) Discuss principal component analysis. Show that they are uncorrelated and have variances equal to the eigen values of the covariance matrix.

(c) If X and Y are such that $X + Y \sim N_p(\mu_1 + \mu_2, \Sigma_1 + \Sigma_2)$.

then show that $X \sim N_p(\mu_1, \Sigma_1)$ and $Y \sim N_p(\mu_2, \Sigma_2)$. Also, obtain conditional distribution of $X \mid Y$. 20×3 = 60

SECTION – B

5. Answer any three of the following sub-parts : 20×3 = 60

(a) Explain minimal sufficient statistic. Also, derive the sufficient statistic for the exponential family of distribution.

(b) Use Neyman-Pearson Lemma to obtain the region for testing $\theta = \theta_0$ against $\theta = \theta_1 > \theta_0$ and $\theta = \theta_1 < \theta_0$, in the case of a normal population $N(\theta, \sigma^2)$, where σ^2 is known. Hence find the power of the test.

(c) Discuss a comparison among stratified random sampling with proportional allocation stratified random sampling with Neyman allocation and simple random sampling.

(d) Distinguish between the fixed effect model and random effect model. Also, estimate the parameters in two-way classified data (one observation per cell) with fixed effect model.

6. (a) Define Bayes estimator. Obtain the Bayes estimator under squared error loss function (SELF).

(b) State and prove Lehmann-Scheffe theorem.

(c) Explain likelihood ratio test and mention some of its important properties. Let X_1, X_2,

......, X_n be i.i.d random variables from $N(\mu, \sigma^2)$, where μ and σ^2 are unknown. Use likelihood ratio test to test $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$. 20×3 = 60

7. (a) Discuss Kolmogorov-Smirnov two-sample test.

   (b) Let X have the distribution.

   $f(x, \theta) = \theta^x(1-\theta)^{1-x}; x = 0, 1; 0 < \theta < 1.$

   For testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$, construct sequential probability ratio test and obtain its A.S.N. and O.C. functions.

   (c) What do you understand by unordered estimator ? Obtain Horvitz-Thompson unbiased estimator of the population total. Also find the expression for variance of the above estimator. 20×3 = 60

8. (a) Explain cluster sampling and discuss its merits and demerits. Also obtain an estimate of the population variance based on above method.

(b) Give the method of calculating sum of squares for factorial experiments in general. Obtain the ANOVA table for a 2^3 design.

(c) Discuss RBD with one missing observation and estimate its value. $20 \times 3 = 60$