2021

STATISTICS — HONOURS

Fifth Paper

(Group - B)

Full Marks : 50

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

1. Answer ***any four*** questions : 5×4

(a) Compare between

(i) simple hypothesis and composite hypothesis

(ii) Type-I and Type-II error

(iii) size and level of significance of a test.

(b) Assuming suitabe probability distributions, discuss a test method for testing, whether the weekly average number of road accidents in two different cities are same or not, based on sample road-accidents data for n_1 and n_2 weeks (n_1, n_2 small) respectively from those two cities.

(c) Based on a random sample of size n from $\mathcal{N}(\mu, \sigma^2)$ population, suggest a test for testing $H_0 : \sigma = \sigma_0$ (given) versus $H_1 : \sigma > \sigma_0$. Also write down the power function of the test.

(d) What do you mean by an UMAU confidence set? Justify its duality with UMPU test.

(e) Develop a suitable testing rule to test whether CIBIL (Credit Information Bureau India Limited) score (y) has any dependence on the income (x), based on paired data on (y, x) from a number of debters.

(f) What do you mean by Mann-Whitney U-test? Show that the null distribution of the test statistic is exactly distribution free.

(g) Let X be a random variable with pmf f_0 if H_0 is true and with pmf f_1 if H_1 is true. The pmf values are as follows :

x	1	2	3	4	5	6
$f_0(x)$	0.01	0.02	0.015	0.45	0.35	0.155
$f_1(x)$	$0.03-0.05\gamma$	0.005	$\gamma(0.2-\gamma)$	$0.365-0.15\gamma$	$0.4+\gamma^2$	0.2

If the critical region is of the form $\{ X < C \}$, find a suitable value of C where level of significance of the test is 0.05. Also find the maximum value of the power depending on the choice of γ. Is the test unbiased at its maximum power? Give reason.

(h) If $\{x_1, ..., x_m\}$ and $\{y_1, ..., y_n\}$ be two independent samples drawn from $\mathcal{N}(\mu_1, \sigma^2)$ and $\mathcal{N}(\mu_2, \sigma^2)$ populations respectively, derive a 100 $(1 - \alpha)$% confidence interval for μ_2/μ_1.

Answer ***any two*** questions from question numbers **2-5**.

2. (a) Based on a random sample of size n (>1) drawn from an exponential population with density

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & , x > 0 \\ 0 & , \textit{elsewhere} \end{cases}$$

$\theta > 0$, find a 100 $(1 - \alpha)$% confidence interval of θ.

Also find its expected length.

(b) When is a test called biased? Show that an MP or an UMP test is necessarily unbiased.

(c) Prove that the test given by the critical region $\left\{\overline{X} > \mu_0 + \frac{\sigma}{\sqrt{n}} \tau_\alpha\right\}$ for testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$ is a biased one, where $X_1, ..., X_n \sim$ IIDN (μ, σ^2), σ : known and τ_α : upper 100α% point of N (0,1) distribution, α being level of significance of the test. 7+5+3

3. (a) Suppose a random sample of size n be drawn from a bivariate normal population N_2 $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$. Discuss how you can get an exact (and not an asymptotic) test for testing $H_0 = \frac{\sigma_2}{\sigma_1} = \delta_0$ against $H_1 = \frac{\sigma_2}{\sigma_1} \neq \delta_0$

(b) Describe the MP size-α test based on single observation X for testing $H_0 : X \sim N$ (0, 1) against $H_1 : X \sim$ Laplace (0, 1). 8+7

4. (a) Discuss ANOVA technique for two-way classified data with m (≥ 2) observations per cell under random effects model. In this context highlight the role of valid error in ANOVA technique.

(b) We know that usually expenditure of a person depends on his/her income. As a result there should be a regression of expenditure on income for a group of individuals. Suppose income (X) (in ₹) and expenditure (Y) (in ₹) are jointly normally (bivariate normal) distributed. Based on n paired readings on (X,Y), obtained from n randomly chosen individuals, develop a test procedure to test whether a person with no income, can have on an average expenditure of ₹ α_0 (given) or not. Also discuss a 100 $(1 - \gamma)$% confidence interval of the average expenditure of a person with no income. 8+7

5. (a) On the basis of a random sample of size n_1 from $N(\mu_1, \sigma_1^2)$, and of a random sample of size n_2 from $N(\mu_2, \sigma_2^2)$, independently, determine a test procedure through LRT for testing $H_0: \sigma_1^2 = \sigma_2^2$ versus $H_1: \sigma_1^2 \neq \sigma_2^2$ at level α. Also check unbiasedness of the test.

(b) Define p-value of a two-tailed test. Describe combination of tests as an application of p-values.

8+(2+5)

———————