2020

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.

Answer ***question no.* 1** and ***any one*** question from the rest.

1. Answer ***any two*** questions : 10×2

(a) Distinguish between

(i) Type-I and Type-II errors of a test

(ii) Size and level of significance of a test.

(b) Define p-value for the one-sided testing problem with the test statistic T. If the null distribution of T is continuous, show that the p-value for the test follows an uniform (0, 1) distribution.

(c) Suppose $X_1, ..., X_m$ be iid $N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, ..., Y_n$ be iid $N(\mu_2, \sigma_2^2)$ and independently. Mentioning necessary assumption, if any, discuss a suitable way of testing $H_0: \frac{\mu_2}{\mu_1} = C$ (known) vs. $H_1: \frac{\mu_2}{\mu_1} \neq C$.

(d) Let $(X_1, Y_1), ..., (X_n, Y_n)$ follow BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ independently. Provide a method for testing $H_0: \frac{\sigma_1^2}{\sigma_2^2} = \delta_0^2$ vs. $H_1: \frac{\sigma_1^2}{\sigma_2^2} > \delta_0^2$ at level α.

(e) Describe the role of combining p-values in tests of significance.

(f) Suppose four pages from the old edition, and three pages from the new edition of a book are randomly selected and the number of misprints in each page are recorded. Under suitable assumption, to be stated clearly, describe an exact test method to test whether printing quality has improved in new edition as compared to old edition of the book.

(g) Describe a test of randomness and mention two real-life situation where you can apply it.

(h) Let the regression of y on the deterministic variable x be $\eta_x = E(y \mid x) = \alpha + \beta x$. Under normality assumption on y, illustrate a method for testing $H_0: \eta_{x_0} \left(= \alpha + \beta x_0\right) =$ 'a' (given). Also provide a $100(1 - \alpha)\%$ confidence interval for η_{x_0}.

2. (a) Define confidence interval for a real-valued parameter θ. What do you mean by an unbiased confidence set? Show that the unbiasedness of a test can be established through this concept.

(b) Let X_1, X_2 follow uniform $(0, \theta)$, $\theta > 0$ independently.

(i) Find a confidence interval for θ in terms of $X_{(2)} = \max\{X_1, X_2\}$ with confidence coefficient $1 - \alpha$.

(ii) Derive another confidence interval for θ, but in terms of $\frac{1}{2}(X_1 + X_2)$ with same confidence coefficient. {You can use the fact that pdf of $X_1 + X_2$ is :

$$f(u) = \begin{cases} u/\theta^2, & 0 < u \leq \theta \\ \dfrac{2\theta - u}{\theta^2}, & \theta < u \leq 2\theta \end{cases} \}$$

(iii) Compare the performances of the two confidence intervals obtained in (i) and in (ii).

12+18

3. (a) What do you mean by linear model? Based on a set of paired data on (y, x) where x is deterministic and y is normal variate, frame a suitable linear model for testing whether y has any true dependence on x. Also discuss the test procedure and express the test statistic in terms of sample correlation ratio (e_{yx}) of y on x.

(b) If a multiple linear regression of y on $x_1, x_2, ..., x_p$ is framed under normality assumption, derive a test procedure to check whether $x_1, x_2, ..., x_p$ are equally effective for predicting the value of y through the said regression.

16+14

4. (a) Let X be a discrete random variable assuming values 0, 1, 2, 3, 4 with positive probabilities. Let $f_0(x)$ and $f_1(x)$ denoting the pmf's under null and alternative hypotheses are respectively, as follows :

x	0	1	2	3	4
$f_0(x)$	0.2	0.4	0.355	0.03	0.015
$f_1(x)$	0.1	$.3-.3\gamma$	γ^2	$.3+.3\gamma$	$.3-\gamma^2$

At level of significance 0.05, find the value of c such that $\{X > c\}$ is the critical region. Also find the maximum possible power of the test with respect to the choice of γ. Is the test with the maximum power unbiased? Verify.

(b) Based on a random sample $X_1, ..., X_n$ from $N(\mu, 2^2)$ distribution, discuss an UMP test procedure for testing $H_0 : \mu = \mu_0$ (given) vs. $H_1 : \mu > \mu_0$ at level α. Verify the unbiasedness of the test.

(c) For the same problem, as in (*b*), find the size of the test $\{X_{(n)} > \mu_0 + b, X_{(1)} > \mu_0 - a\}$, where a and b $(> a)$ are known constants.

10+12+8

5. (a) Suppose the sale volumes are recorded for 5 products of a renowned brand among four developed countries in the world. Based on some suitable assumptions, to be mentioned clearly, discuss a procedure to check whether the sales vary significantly over the countries, or over the products, under each of the following two cases :

Case I : The brand is restricted to those four countries only.

Case II : Those four countries are randomly selected from the countries around the world who are using that brand.

(b) Discuss Wilcoxon signed Rank test procedure for testing population location by stating clearly the necessary assumption for its application. Show that the null distribution of the test statistic is distribution free and symmetric. 18+12