3 (Sem–6) STS M1

2020

STATISTICS

(Major)

Paper : 6·1

(Statistical Inference-2)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Choose the correct answer from the given options : 1×7=7

   (a) A confidence interval of confidence coefficient $(1-\alpha)$ is considered best when it has

   (i) smallest width

   (ii) largest width

   (iii) average width

   (iv) upper and lower limits equidistant from the parameter

(b) For a certain test, you are given that $\alpha = 0{\cdot}05$ and $\beta = 0{\cdot}10$. The power of the test is

(i) 0·95

(ii) 0·90

(iii) 0·05

(iv) 0·10

(c) Area of the critical region depends on

(i) number of observations

(ii) value of the statistic

(iii) size of the Type I error

(iv) size of Type II error

(d) If there are 10 symbols of two types, equal in number, the minimum possible number of runs is

(i) 1

(ii) 2

(iii) 3

(iv) 5

(e) For an unbiased critical region, in testing a simple null hypothesis $H_0 : \theta = \theta_0$ (say) against a simple alternative $H_1 : \theta = \theta_1$ (say), we have,

(i) $1 - \beta < \alpha$

(ii) $1 - \beta \geq \alpha$

(iii) $\beta \geq \alpha$

(iv) $\beta + \theta_1 < \theta_0 + \alpha$

(f) Which of the following tests is equivalent / analogous to the χ^2-test of goodness of fit ?

(i) Mann-Whitney U test

(ii) Wilcoxon signed rank test

(iii) Kolmogorov-Smirnov test

(iv) Median test

(g) The ratio of the likelihood function function under H_0 and under the entire parametric space is called

(i) probability ratio

(ii) sequential probability ratio

(iii) likelihood probability ratio

(iv) likelihood ratio.

2. Answer the following questions : 2×4=8

(a) Explain simple and composite hypotheses with examples.

(b) Define Uniformly Most Powerful (UMP) Critical Region and UMP test.

(c) What is the difference between sign test and Wilcoxon signed rank test ?

(d) State the asymptotic properties of Likelihood-ratio test.

3. Answer ***any three*** of the following questions : 5×3=15

(a) Obtain 100 $(1-\alpha)\%$ confidence intervals for the parameters θ and σ^2 of the normal distribution.

(b) Discuss the Kolmogorov-Smirnov two-sample test.

(c) What do you understand by a Statistical Hypothesis and a Null hypothesis ? Describe the errors involved in testing of hypothesis.

(d) State and prove the Neyman Pearson lemma.

(e) Define Spearman's rank correlation coefficient and Kendall's Tau. Discuss the similarities and differences between them.

4. Answer ***any three*** of the following questions : 10×3=30

(a) (i) Explain level of significance, confidence probability, power of a test and critical region. 4

(ii) Given the frequency function

$$f(x,\theta)=\begin{cases}\frac{1}{\theta}, & 0\le x\le\theta\\ 0, & \text{otherwise}\end{cases}$$

For testing the null hypothesis $H_0:\theta=1$ against $H_1:\theta=2$, by means of a single observed value x, what would be the size of Type I and Type II errors and the Power functions of the tests for the following critical regions : 6

(a) $0{\cdot}5 \le x$ (b) $1 \le x \le 1{\cdot}5$

(b) (i) Write a note on Association and Contingency. 4

(ii) Explain the Mann Whitney U-test. To which parametric test of significance is it analogous ?

5+1=6

(c) What do you mean by Best Critical Region (BCR) ?

Using Neyman Pearson lemma, obtain the BCR for testing $H_0:\theta=\theta_0$ against $H_1:\theta=\theta_1>\theta_0$ in case of a normal population $N(\theta,\sigma^2)$, where σ^2 is known.

Also find the power of the test.

2+6+2=10

(d) What are the advantages of Likelihood ratio test over Neyman Pearson test ? Derive the Likelihood ratio test procedure for testing the equality of variance of two univariate normal populations. 2+8=10

(e) Define a Run. When do we use the One-sample Run test ? Describe the Wald-Wolfowitz Runs test stating the underlying assumptions, if any.

1+2+7=10

---

3 (Sem–6) STS M 2

2020

STATISTICS

(Major)

Paper : 6·2

(Design of Experiments)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer the following as directed : 1×7=7

(a) For one-way classified data with equal number of observation $(n_1 = n_2 = ... n_k = n)$ per class to compare the class means two at a time we are to calculate —

(i) $t_{\alpha/2,\, k(n-1)} \times \sqrt{\frac{MSE}{k}}$ *(ii)* $t_{\alpha/2,\, (n-1)} \times \sqrt{\frac{2MSE}{n}}$

(iii) $t_{\alpha/2,\, k(n-1)} \times \sqrt{\frac{2MSE}{n}}$ *(iv)* None of the above.

(Choose the correct option)

Contd.

(b) If r is the number of replications for each treatment then variance of difference between two treatment means is given by —

(i) σ^2/r^2

(ii) σ/r^2

(iii) $2\sigma^2/r^2$

(iv) None of the above.

(Choose the correct option)

(c) What will be the error d.f. in a RBD to compare 5 treatments in 4 blocks, having one missing observation ?

(d) In a 4×4 LSD, the following results were obtained —

Row MS = 87 Column MS = 52

Treatment MS = 457 Total SS = 1943

Compute error MS.

(e) The Concept of Confounding —

(i) is not deliberately introduced in factorial experiment

(ii) is deliberately introduced in factorial experiment

(iii) is sometimes deliberately introduced and sometimes not

(iv) None of the above.

(Choose the correct option)

(f) What will be the total number of factorial effects in 2^n factorial experiment ?

(g) In a split-plot design —

(i) main effect is confounded

(ii) interaction effect is confounded

(iii) sometimes main effect is confounded and sometimes interaction effect is confounded

(iv) None of the above.

2. Answer the following : 2×4=8

(a) With 3 factors A, B, C each at 2 levels, write the factorial effects. Also give the layout of a replication if interaction BC is confounded.

(b) Write the linear model of an one-way classified data when there is a concomitant variable.

(c) Explain why there cannot be a 2×2 LSD.

(d) In a factorial experiment there are two factors V and M. The factor V is to be applied at three levels v_1, v_2, v_3 and the factor M is to be applied at 4 levels m_1, m_2, m_3, m_4. Construct the layout of the experiment in split-plot design, using 3 replications.

Or

State *any two* limitations of A.O.V.

3. Answer ***any three*** of the following : 5×3=15

(a) Discuss the types of model and their underlying assumptions that are associated with the Analysis of Variance (AOV) technique.

(b) Show that in RBD treatment and error effects are mutually orthogonal.

(c) The part of the AOV table of a RBD is given below :

Source of variation	d.f.	MS
Blocks	3	60
Treatments	4	36
error	12	30

Test the significance of treatment variation and examine how far your conclusion would change when Block classification is not taken into consideration.

Given —

$F_{0.05;\nu_1,\nu_2} = 3.26$ for $\nu_1 = 4, \nu_2 = 12$

$F_{0.01;\nu_1,\nu_2} = 5.41$ for $\nu_1 = 4, \nu_2 = 12$

$F_{0.05;\nu_1,\nu_2} = 3.06$ for $\nu_1 = 4, \nu_2 = 15$

$F_{0.01;\nu_1,\nu_2} = 4.89$ for $\nu_1 = 4, \nu_2 = 15$

(d) An agricultural field was laid out in a LSD for comparing 5 varieties of paddy. Due to the negligence of the caretaker, the crop in one of the border-lying plots was damaged by cattle long before harvesting. Describe the procedure you would adopt to analyse the resulting yield-data, briefly indicating the basic theory of involved therein.

(e) What is balanced factorial design ? Give the confounding subgroups for a 2^4 balanced design in 6 replications in 4 blocks of size 4 each.

4. *(a)* What do you mean by efficiency of a design? Discuss how efficiency of a design can be increased. Obtain an expression for efficiency of LSD over RBD. 2+2+6=10

Or

For a RBD with one missing observation, let V_1 and V_2 denote respectively the correct treatment SS and the treatment SS for the completed data. Which one is greater? Justify your answer. 10

(b) Discuss the necessity of confounding in factorial design. How does partial confounding differ from complete confounding? Give your answer with suitable illustration. 3+4+3=10

Or

Obtain the expression for unrestricted residual sum of square in case of analysis of covariance for two-way classified data. 10

(c) The content of one of the block in a 2^5 factorial experiment involving the factors *A*, *B*, *C*, *D* and *E* tested in block of size 8 are —

(*abc* *abcde* *c* *ad* *be* *cde* *ae* *bd*)

give the contents of the principal block and hence identify the confounded effects. Discuss the analysis of the experiment assuming 5 such replications. Further indicate what you would do under the following two different situations —

(i) All 2nd order interactions are of equal importance

(ii) Experimental material does not permit to have more than one replication. 2+2+4+1+1=10

Or

Find the standard error of difference between two treatment means when one of them has a missing observation in RBD. 10

———

3 (Sem–6) STS M 3

2020

STATISTICS

(Major)

Paper : 6·3

(Applied Statistics – 2)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer the following questions as directed : 1×7=7

   (a) The first Indian census took place in the year _____. *(Fill in the blank)*

   (b) In the construction of which chart Poisson distribution is used ?

   (i) $\overline{X}$ – chart *(ii)* R–chart

   (iii) P–chart *(iv)* C–chart

   (Choose correct answer)

(c) The control chart for fraction defective is called *np*-chart.

(Write true or false)

(d) An assumed number of newly born babies at the same time denoted by l_0 is called _____. *(Fill in the blank)*

(e) The probability of accepting a lot with fraction defective p_t is termed as _____.

(Fill in the blank)

(f) Most serious drawback of crude death rate is that it does not take into account of the age and sex distribution.

(Write true ***or*** *false)*

(g) What is meant by the statement 'NRR of a country is 0·508'?

2. Answer the following questions : 2×4=8

(a) Distinguish clearly between defect and defective.

(b) Explain the usefulness of *R*-chart.

(c) Name different measures of population growth.

(d) What do you understand by abridged life table?

3. Answer ***any three*** of the following questions : 3×5=15

(a) Explain the main control charts for attributes and obtain their control limits.

(b) Is the crude death rate an accurate measure of mortality of a population of a country? If not, how will you modify it to give reliable results?

(c) What is Average Sample Number (ASN) and Average Total Inspection (ATI)? Explain the method of their calculation for single sampling plan.

(d) State the assumptions regarding the population under which a life table is constructed. Also mention the uses of a life table.

(e) What is clinical trial? Write a note on the importance of its study.

4. Answer ***any three*** of the following questions : 10×3=30

(a) What do you mean by fertility of a population? Explain the various measures of fertility in common use and discuss their merits and demerits.

(b) Describe single sampling plan. Obtain OC and AOQ curve for this plan.

(c) Starting from a suitable assumption regarding the relative growth rate of population, derive the logistic equation. Also mention the properties of this curve.

(d) What do you mean by statistical quality control? Discuss its utility and limitations. Also state the role of SQC in the industrial world.

(e) Write brief notes on the origin and functions of the

(i) Central Statistical Organisation and

(ii) National Sample Survey Organisation.

(f) Define (i) Central mortality rate (m_x)

(ii) Force of mortality rate (μ_x)

and (iii) Probability of death (q_x) in a life table.

Hence show that

(i) $q_x = \dfrac{2m_x}{2 + m_x}$

(ii) $m_x = \mu_x + \dfrac{1}{2}$.