2017

STATISTICS

(Major)

Paper : 6.1

(Statistical Inference–2)

Full Marks : 60

Time : 3 hours

The figures in the margin indicate full marks for the questions

1. Choose the correct answer : 1×7=7

   (a) Suppose you want to test $\theta = \theta_0$ against $\theta = \theta_1$ with w as critical region and A as acceptance region, then $P_{\theta_0}(w)$ is

   (i) the probability of type II error

   (ii) the probability of type I error

   (iii) Both type I and type II errors

   (iv) None of the above

   (b) As the same symbols in **1.**(a) $P_{\theta_1}(A)$

   (i) is a correct decision

   (ii) is a wrong decision

(iii) is sometimes correct and sometimes not

(iv) None of the above

(c) In the sign test, we consider

(i) the direction of the differences

(ii) the magnitudes of the differences

(iii) both sign and magnitude of the differences

(iv) None of the above

(d) Let X be a normal variate with mean μ and variance σ^2, then the 95% confidence interval for μ is (with usual notation)

(i) $\left[\bar{x}-2\cdot 58\frac{\sigma}{\sqrt{n}}, \bar{x}+2\cdot 58\frac{\sigma}{\sqrt{n}}\right]$

(ii) $\left[\bar{x}-2\cdot 00\frac{\sigma}{\sqrt{n}}, \bar{x}+2\cdot 00\frac{\sigma}{\sqrt{n}}\right]$

(iii) $\left[\bar{x}-1\cdot 96\frac{\sigma}{\sqrt{n}}, \bar{x}+1\cdot 96\frac{\sigma}{\sqrt{n}}\right]$

(iv) None of the above

(e) The Kolmogorov-Smirnov one sample test is used to test

(i) the independence of attributes

(ii) about the population mean

(iii) the population variance

(iv) the goodness of fit

(f) Uniformly most powerful test

(i) always exists

(ii) never exists

(iii) sometimes exists and sometimes not

(iv) None of the above

(g) With usual notation, if

$$P(c_1 < \theta < c_2) = 1 - \alpha$$

then c_1 and c_2 are called

(i) confidence coefficients

(ii) confidence intervals

(iii) confidence means

(iv) None of the above

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

(a) Define the uniformly most powerful test.

(b) Define a 'run'. Also give an example.

(c) Define the Kendall's tau (τ).

(d) State the Neyman-Pearson lemma.

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

(a) Write a note on confidence interval and confidence limit. 5

(b) Given the frequency function

$$f(x, \theta) = \frac{1}{\theta}, \quad 0 \le x \le \theta$$
$$= 0, \quad \text{elsewhere}$$

and that you are testing the null hypothesis $H_0 : \theta = 1$ against $H_1 : \theta = 2$ by means of a single observed value of x, what would be the size of the type I error if you choose the interval $1 \le x \le 2$? Also find the power of the test. 2+3=5

(c) Describe the sign test. 5

(d) Write a note on likelihood ratio test. 5

(e) Let p be the probability that a coin shows head in single toss. In order to test $H_0 : p = \frac{1}{2}$ against $H_1 : p = \frac{3}{4}$, the coin is tossed four times. H_0 is rejected if more than 3 heads are obtained. Find the probability of type I error and power of the test. 2+3=5

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

(a) *(i)* Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(\mu, \sigma^2)$ where μ and σ^2 are both unknown, find the confidence interval for μ. Also interprete your result. 7

(ii) Define the simple and composite hypotheses. $1\frac{1}{2}+1\frac{1}{2}=3$

(b) Show that for the normal distribution with zero mean and variance σ^2, best critical region for $H_0 : \sigma = \sigma_0$ against the alternative $H_1 : \sigma = \sigma_1$ is of the form

$$\sum_{i=1}^{n} x_i^2 \le a_\alpha \text{ for } \sigma_0 > \sigma_1$$

and $$\sum_{i=1}^{n} x_i^2 \ge b_\alpha \text{ for } \sigma_0 < \sigma_1$$

Also find the size of the critical region when $\sum_{i=1}^{n} x_i^2 \le a_\alpha$.

7+3=10

(c) (i) Write a note on Wilcoxon signed-rank test. 5

(ii) State the advantages of non-parametric test. 5

(d) Describe how you would test for the mean (μ) of the normal population when variance is known using likelihood ratio test. 10

(e) (i) Describe briefly the Mann-Whitney U test. 5

(ii) Define type I and type II errors. Also define the most powerful tests.

$1\frac{1}{2}+1\frac{1}{2}+2=5$

* * *

2017

STATISTICS

(Major)

Paper : 6.2

(Design of Experiments)

Full Marks : 60

Time : 3 hours

The figures in the margin indicate full marks for the questions

1. Answer the following as directed : $1 \times 7 = 7$

(a) The equality of means is tested in analysis of variance.

(State true or false)

(b) Replication in a design of experiment means

(i) block size

(ii) number of treatments

(iii) number of times a treatment is repeated

(iv) number of plots

(Choose the correct option)

(c) Which of the following statements is correct for completely randomised design?

(i) It has two principles—replication and randomization only.

(ii) It has two principles—replication and local control only.

(iii) It has two principles—randomization and local control only.

(iv) It has all the three principles replication, randomization and local control.

(Choose the correct option)

(d) While analysing the data of a $k \times k$ Latin square design, the degree of freedom for error is _____.

(Fill in the blank)

(e) For a 2^2-factorial experiment in an r-randomised block, the sum of squares for the main effect A in the analysis of variance table is _____.

(Fill in the blank)

(f) What is a treatment contrast?

(g) In the linear model of analysis of variance, the error part is assumed to be distributed as

(i) $N(\mu, \sigma^2)$

(ii) $N(0, \sigma^2)$

(iii) $N(\mu, 0)$

(iv) $N(0, 1)$

(Choose the correct option)

2. Answer the following : $2 \times 4 = 8$

(a) In an RBD with 6 treatments and 5 blocks, the following results were obtained :

$$MSB = 20, \; MST = 25, \; TSS = 245$$

Complete the ANOVA table.

(b) Explain why there cannot be a 2×2 Latin square design.

(c) Show that for 2^3-factorial experiment, the main effect A and interaction effect AB are mutually orthogonal contrasts.

(d) In a 2^4-factorial experiment, the key block is given by

$$(1), \; ab, \; cd, \; abcd$$

Identify the confounded effect(s).

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

(a) In ANOVA testing for one-way classification (for fixed effect model), if the null hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ is rejected, how will you proceed to test the significance of the difference between any two treatment means?

(b) Explain the basic principles of good experimental design giving brief explanatory note for each.

(c) Derive the expressions to measure the efficiency of Latin square design over a randomised block design when rows are used as blocks.

(d) Prove that in a 2^4-factorial experiment in 4 blocks of 4 plots, it is impossible to avoid first-order interaction among the three that are confounded. Also give a scheme of balanced design using such blocks.

(e) Explain the concepts of partial, total and balanced confoundings in factorial experiments.

4. *(a)* Derive the expected values of mean squares for two-way classified data with one observation per cell under the fixed effect mathematical model. Also show that they provide unbiased estimate of error variance. 10

Or

Discuss the analysis of covariance technique in a two-way classified data. 10

(b) What do you understand by a 'missing plot' in a design of experiment? How would you estimate a missing value in Latin square design? Give an outline of the analysis of variance of a $p \times p$ Latin square design involving a single missing plot. 10

Or

Discuss the analysis of partial confounding in a 2^4-factorial experiment. 10

(c) Find the standard error of difference between two treatment means, when one of them has a missing observation in a randomised block design. 10

Or

Discuss how analysis of variance technique can be used to test for multiple linear regression model. 10

★★★

2017

STATISTICS

(Major)

Paper : 6.3

(Applied Statistics—2)

Full Marks : 60

Time : 3 hours

The figures in the margin indicate full marks for the questions

1. Answer *all* questions : 1×7=7

   (a) In statistical quality control (SQC), when is *c*-chart used?

   (b) What is expectation of life at birth?

   (c) What is a *p*-chart?

   (d) Define general fertility rate.

   (e) What is meant by a vital event?

   (f) What does the l_x column of a complete life table denote?

   (g) The term Vital Statistics and Population Statistics in wider sense are _____.

   (Fill up the gap)

2. Answer *all* questions : 2×4=8

(a) Distinguish between gross and net reproduction rates.

(b) Distinguish between chance causes and assignable causes in SQC.

(c) What are the important sources of demographic data?

(d) What are the important functions of National Statistical Commission?

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

(a) Discuss different columns of a complete life table together with their importance.

(b) What is infant mortality rate? Discuss the problems in its construction. What are its merits and demerits?

(c) What is meant by sampling inspection plan? Describe the single sampling inspection plan.

(d) What is standardised death rate? What are its advantages and disadvantages over other types of death rates?

(e) Discuss different types of birthrates.

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

(a) What is population projection? Describe completely one method of population projection.

(b) Explain and describe the terms, producer's risk; consumer's risk; AOQL and LTPD.

(c) Write a note on the origin and function of NSSO.

(d) Explain in detail the $\overline{X}$ and R charts. What purposes do they serve? What are their advantages over the p-chart?

(e) Discuss and attempt a comparison between Census 2001 and Census 2011.

(f) Distinguish between attributes and variables in the context of SQC. Discuss different types of attribute control charts.

* * *

2017

STATISTICS

(Major)

Paper : 6.4

(Computer Programming and Multivariate Analysis)

Full Marks : 60

Time : 3 hours

The figures in the margin indicate full marks for the questions

1. Answer the following questions as directed : $1 \times 7 = 7$

(a) Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$. Then its characteristic function is given by

(i) $e^{i\underset{\sim}{t}\underset{\sim}{\mu}+\frac{1}{2}\underset{\sim}{t}'\Sigma\underset{\sim}{t}}$

(ii) $e^{i\underset{\sim}{t}'\underset{\sim}{\mu}-\frac{1}{2}\underset{\sim}{t}'\Sigma\underset{\sim}{t}}$

(iii) $e^{i\underset{\sim}{t}'\underset{\sim}{\mu}+\frac{1}{2}\underset{\sim}{t}'\Sigma\underset{\sim}{t}}$

(iv) None of the above

(Choose the correct option)

(b) Write down the binary equivalent of decimal number 64.

(c) The moment-generating function of bivariate normal distribution with parameters $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ is ____.

(Fill in the blank)

(d) Multinomial pmf can be expressed as generalization of binomial distribution.

(State True or False)

(e) Which language is directly understood by the computer without translation program?

(i) Machine language

(ii) Assembly language

(iii) High-level language

(iv) None of the above

(Choose the correct option)

(f) What does RAM stand for?

(g) Let $(X, Y) \sim \text{BVND}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Then state the probability density function.

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

(a) State any two applications of multivariate distributions.

(b) Write a note on different FORTRAN variables.

(c) State the marginal pdfs of X and Y in case of a BVND $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

(d) Differentiate between Assembler and Compiler.

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

(a) Obtain the moment-generating function of a multinomial distribution.

(b) If X and Y are $p \times 1$ and $q \times 1$ vectors of random variables, and a and b are $p \times 1$ and $q \times 1$ vectors of constants, then prove that

$$\text{cov}(X-a, Y-b) = \text{cov}(X, Y)$$

(c) Write a note on flowchart symbols and their uses.

(d) Write a FORTRAN 77 programme to find correlation coefficient of n pairs of numbers.

(e) If $(X, Y) \sim \text{BVND}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then show that X and Y are independent if and only if $\rho = 0$.

4. Answer any *three* questions : $10\times3=30$

(a) Let $(X, Y) \sim BVND(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Find the conditional distributions of $X / Y = y$ and $Y / X = x$. 10

(b) Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{0}, \Sigma)$, where $\Sigma = (\sigma_{ij})_{p\times p}$. Prove that $\underset{\sim}{X}' \Sigma^{-1} \underset{\sim}{X}$ follows chi-square distribution with p degrees of freedom. 10

(c) Write a program to find mean of n observations $x_1, x_2, \ldots, x_n$. Also write a FORTRAN 77 programme to find the coefficient of variation. $3+7=10$

(d) Write a note on different types of relational operators available in FORTRAN 77. Also write a FORTRAN 77 programme to find the root of $x^4 - x - 10 = 0$, which is nearer to $x = 2$, correct to five places of decimal using Newton-Raphson method. $3+7=10$

(e) Compare Hotelling's T^2 with univariate counterpart. Examine if Hotelling's T^2 is invariant under changes in the units of measurement. $3+7=10$

(f) Derive mean and variance of multinomial distribution. Also compute the variance-covariance matrix. $3+3+4=10$

★ ★ ★