(b) Write an explanatory note on sign test.

(c) Write a note on likelihood ratio test.

(d) Write notes on non-parametric and distribution-free tests.

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

(a) Describe any non-parametric test consisting of two samples.

(b) (i) Write an explanatory note on confidence interval. 6

(ii) Describe briefly the run test. 4

(c) (i) State the advantages of Kolmogorov-Smirnov one-sample statistic over chi-square test. 3

(ii) Suppose you are testing $H_0 : \lambda = 2$ against $H_1 : \lambda = 1$, where λ is the parameter of the Poisson distribution. Obtain the best critical region of the test. 7

(d) Describe how you will test the equality means of two univariate normal distributions using likelihood ratio test.

★★★

2019

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) In a testing of hypothesis, power of a test is

(i) the probability of accepting a hypothesis when it is true

(ii) the probability of rejecting a hypothesis when it is false

(iii) the probability of accepting a hypothesis when it is false

(iv) None of the above

(b) In a sign test, we use

(i) only the sign of the differences

(ii) only the magnitude of the differences

(iii) both the magnitude and sign of the differences

(iv) None of the above

(c) In a testing of hypothesis problem, we try to fix
  (i) the type-II error and minimize type-I error
  (ii) the type-I error and maximize type-II error
  (iii) the type-I error and maximize the power of the test
  (iv) None of the above

(d) In a testing of hypothesis problem for a normal distribution, if we specify the mean but not the variance, then it will be the case of
  (i) simple hypothesis
  (ii) composite hypothesis
  (iii) alternative hypothesis
  (iv) None of the above

(e) The Neyman-Pearson lemma
  (i) always gives us the uniformly most powerful (UMP) test
  (ii) sometimes gives us the UMP test
  (iii) never gives us the UMP test
  (iv) None of the above

(f) The Kolmogorov-Smirnov test statistic
  (i) uses the concept of empirical distribution
  (ii) never uses the concept of empirical distribution

  (iii) sometimes uses the concept of empirical distribution
  (iv) None of the above

(g) Suppose we put forward an interval which we expect would include the true parametric value, then the process is called
  (i) testing of hypothesis
  (ii) non-parametric inference
  (iii) interval estimation
  (iv) None of the above

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

(a) Define type-I and type-II errors.

(b) Define Kendall's τ.

(c) State the Neyman-Pearson lemma.

(d) Define the most powerful test.

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

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

Or

(b) (i) What is a factorial experiment? In what respect is it different from a number of single-factor experiments, the number being equal to the number of factors in the factorial experiment? 5

(ii) Explain three basic principles of experimental design. 5

★★★

2019

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 : 1×7=7

(a) Replication means the execution of an experiment more than _____.

(Fill in the blank)

(b) _____ is the simplest design making use of all the three basic principles of design.

(Fill in the blank)

(c) What is a contrast?

(d) The error degrees of freedom in a $p \times p$ LSD is _____.

(Fill in the blank)

(e) In an RBD, the number of plate in a block is equal to the number of _____.

(Fill in the blank)

(f) In the linear model of analysis of variance, the error term 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)

(g) What will be the error d.f. in an RBD with 3 blocks comparing 4 treatments having one missing observation?

2. Answer the following : 2×4=8

(a) Explain what you understand by 'analysis of variance'. State the basic assumptions in an analysis of variance.

(b) In a 4×4 LSD, the following results were obtained :

RMS = 87, CMS = 52, TMS = 457, Total SS = 1943

Complete the ANOVA table.

(c) Explain the use of local control in Latin square design.

(d) In a 2^3-factorial experiment in blocks of 4 plots with three fertilizers N, P and K, the control blocks are given below :

Replicate I	*np*	*npk*	(1)	*k*
Replicate II	(1)	*npk*	*nk*	*p*
Replicate III	*pk*	*nk*	(1)	*np*

Identify the confounded effects.

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

(a) Describe how ANOVA technique can be used to test for multiple linear regression model.

(b) What is meant by confounding in a factorial experiment? Why is confounding used even at the cost of loss of information on the confounded effect? Explain the terms 'complete confounding', 'partial confounding' with example.

(c) What do you understand by missing plot technique in a design of experiment? Obtain the estimate of missing observation in an LSD.

(d) Describe the layout of a 2^3-experiment where all the interactions are partially confounded. Give the structure of the AOV table in this case.

(e) What is the principle of allocating treatments in the blocks of an RBD? Give the layout of an RBD with 5 blocks and 4 treatments *A*, *B*, *C* and *D*.

4. *(a)* Discuss the analysis of covariance technique in a two-way classified data. 10

Or

(b) Explain the assumptions underlying the results of a Latin square design and the types of hypotheses that can be tested. Draw up the analysis of variance table and give the expected values of mean squares. 2+2+3+3=10

5. *(a)* What is a split plot design? Give the analysis of this design. Why is it said that this design confounds main effects? 2+6+2=10

Or

(b) Give an outline of the analysis of variance of a randomized block design. Find the standard error of the difference between two treatment means, when one of them has a missing observation in a randomized block design. 5+5=10

6. *(a)* Suppose we have a 2^3-experiment with three factors each at two levels. There are three replicates each divided into two blocks. Show how will you confound *ABC* in the first replication, *AC* in the second replication and *BC* in the third replication. Give the analysis of the design. 3+7=10

2019

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 the following as directed : 1×7=7

(a) Define total fertility rate.

(b) What does the L_x column of a complete life table denote?

(c) In SQC, when C chart is used?

(d) What are the control limits of the $\overline{X}$ chart?

(e) What was the literacy rate of Assam as per Census 2011?

(f) What is infant mortality rate?

(g) LTPD means _____.

(Fill in the blank)

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

(a) Distinguish between product control and process control in SQC.

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

(c) How can population projection be used to determine the future population of a country?

(d) Distinguish between stationary and stable pouplations.

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

(a) Distinguish clearly between control charts for variable and control charts for attributes.

(b) What is epidemiology? Write a note on the importance of its study.

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

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

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

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

(a) Explain Rhodes method of fitting a logistic curve. Also explain why fitting of logistic curve is not satisfactory for Indian population.

(b) Define GRR and NRR in detail with their merits and demerits. Also derive the relationship between them.

(c) What are different types of control chart for variables? Explain in detail.

(d) Explain and describe the terms producer's risk, consumer's risk and AOQL.

(e) Discuss and compare the important highlight of Census 2001 and Census 2011.

(f) Distinguish between tolerance limit and specification limit. Also give the statistical basis of 3-sigma limits.

★ ★ ★

4. Answer the following questions : 10×3=30

(a) State the pdf of BVND $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Derive bivariate normal density as a particular case of multivariate normal distribution. 1+9=10

Or

(b) If $\underset{\sim}{X} \sim N_p(\mu, \Sigma)$, then prove that the quadratic form in the multivariate normal density function

$$Q = (\underset{\sim}{X} - \mu)' \Sigma^{-1} (\underset{\sim}{X} - \mu)$$

follows χ^2 distribution with pdf.

(c) Write a FORTRAN 77 program to find the regression coefficient of Y on X.

Or

(d) *(i)* Write an explanatory note on 'Arithmetic IF' statement used in FORTRAN 77.

(ii) Explain briefly about WHILE–DO statement. 5+5=10

(e) Derive mean and variance of multinomial distribution. Also compute the variance, covariance matrix Σ. 4+4+2=10

Or

(f) *(i)* Let $\underset{\sim}{X} \sim N_5(\mu, \Sigma)$. Then find the distribution of $(X_2\ X_4)'$.

(ii) Write a FORTRAN 77 program to calculate harmonic mean of n observations $x_1, x_2, \cdots, x_n$. 4+6=10

★ ★ ★

2019

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 as directed : 1×7=7

(a) Let $\underset{\sim}{X} \sim N_3(\mu, \Sigma)$ with

$$\Sigma = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Are X_1 and X_3 independent?

(b) State the built-in mathematical function in Fortran 77 to find the square root of x.

(c) Marginal distribution of any x_j of a multinomial distribution with parameters $(n, p_1, p_2, \ldots, p_k)$ follows binomial distribution. (State True or False)

(d) Let (X, Y) ~BVND $(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$. Then the conditional variance of $X/Y = y$ is ____. (Fill in the blank)

(e) Write the decimal equivalent of octal number 2534_8.

(f) What is the value of M in the following Fortran 77 expression?

$$M = 2 * 7 / 5$$

(i) $M = 2 \cdot 8$

(ii) $M = 3$

(iii) $M = 2$

(iv) None of the above

(Choose the correct option)

(g) Define Hotelling's T^2 statistic.

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

(a) State any two properties of multivariate normal distribution.

(b) Write equivalent FORTRAN 77 statements for each of the following expressions :

(i) $e^{-x}x^{kx}$

(ii) $e^{-\left(\frac{x-a}{b}\right)^2}$

(c) Let $\underset{\sim}{X} \sim N_p(\mu, \Sigma)$. Then find var $(C\underset{\sim}{X})$ where C is a $p \times p$ non-singular matrix.

(d) Write an algorithm to find the arithmetic mean of three numbers A, B, C.

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

(a) Let $(X, Y) \sim$ BVND $(0, 0, 1, 1, \rho)$. Then show that

$$Q = \frac{X^2 - 2\rho XY + Y^2}{(1-\rho^2)}$$

is distributed as chi-square with $n = 2$ d.f.

(b) What is the final value of a in the following sequence of statements in FORTRAN 77?

$$a = 2 \cdot 45$$
$$a = (a + 0 \cdot 06) * 10$$
$$k = a$$
$$a = k$$
$$a = a / 10 \cdot 0$$

If $a = 2 \cdot 45$ is replaced by $a = 2 \cdot 43$ above, what is the final value of a ? 3+2=5

(c) Obtain the probability-generating function of multinomial distribution with parameters $(n, p_1, p_2, \cdots p_k)$.

(d) Draw a flowchart to find the largest among three numbers M, N, P.

(e) Examine if Hotelling's T^2 is invariant under changes in the unit of measurement.