I.E.S/ISS EXAM-2022

वियोज्य *DETACHABLE*

# STATISTICS
## Paper IV

Time Allowed : Three Hours

Maximum Marks : 200

### QUESTION PAPER SPECIFIC INSTRUCTIONS

**Please read each of the following instructions carefully before attempting questions.**

There are **FOURTEEN** questions divided under **SEVEN** Sections.

Candidate has to choose any **TWO** Sections and attempt the questions therein. All the Sections carry equal marks.

The number of marks carried by a question/part is indicated against it.

Wherever any assumptions are made for answering a question, they must be clearly indicated.

Diagrams/Figures, wherever required, shall be drawn in the space provided for answering the question itself. F-table and Graph paper can be found at the end of both sections in the booklet.

Unless otherwise mentioned, symbols and notations have their usual standard meanings.

Attempts of questions shall be counted in sequential order. Unless struck off, attempt of a question shall be counted even if attempted partly.

Any page or portion of the page left blank in the Question-cum-Answer (QCA) Booklet must be clearly struck off.

Answers must be written in **ENGLISH** only.

## SECTION 'A'

**(Operations Research and Reliability)**

**1.** (a) Given the following information related to a project, draw the arrow diagram and find the critical path and total duration of the project :

| *Activity* $(i-j)$ | *Duration* *(in days)* |
|---|---|
| 0 – 1 | 2 |
| 1 – 2 | 8 |
| 1 – 3 | 10 |
| 2 – 4 | 6 |
| 2 – 5 | 3 |
| 3 – 4 | 3 |
| 3 – 6 | 7 |
| 4 – 7 | 5 |
| 5 – 7 | 2 |
| 6 – 7 | 8 |

10

(b) What do you mean by an assignment problem ? How do you interpret it as a linear programming problem (LPP) ?

If the completion times (in hours) of different jobs by different workers are as given in the following table, determine the optimal assignment schedule of jobs to the workers :

*Jobs*

| | | *A* | *B* | *C* | *D* | *E* | *F* |
|---|---|---|---|---|---|---|---|
| | *I* | 16 | 47 | 14 | 27 | 0 | 26 |
| | *II* | 0 | 7 | 27 | 43 | 59 | 28 |
| *Workers* | *III* | 0 | 12 | 33 | 24 | 5 | 5 |
| | *IV* | 8 | 13 | 11 | 15 | 0 | 6 |
| | *V* | 3 | 14 | 13 | 0 | 4 | 7 |
| | *VI* | 52 | 10 | 10 | 30 | 21 | 0 |

10

(c) The failure law of a component has the following probability density function :

$$f(t) = (r+1)A^{r+1}/(A+t)^{r+2},\ t>0;$$

where $A$ and $r$ are two non-negative constants.

Derive the expressions of the reliability function and the hazard function. Also show that the hazard function is decreasing in $t$. 10

(d) State the Bellman's principle of optimality in dynamic programming. State how an LPP can be formulated as a dynamic programming problem.

Obtain the solution of the following LPP using dynamic programming :

$$\begin{aligned} \text{Maximize } Z &= 2x_1 + 5x_2 \\ \text{subject to } \quad 2x_1 + x_2 &\leq 430 \\ 2x_2 &\leq 460 \\ x_1, x_2 &\geq 0. \end{aligned}$$

10

(e) Let two components operate independently in parallel. Let failure times of both the components have exponential distribution with parameters $\alpha_1$ and $\alpha_2$ respectively. Show that

(i) the pdf of the failure time of the system, say $T$, is not exponentially distributed, and

(ii) the expected value of $T$ is given by

$$E(T) = \frac{1}{\alpha_1} + \frac{1}{\alpha_2} - \frac{1}{\alpha_1 + \alpha_2}.$$

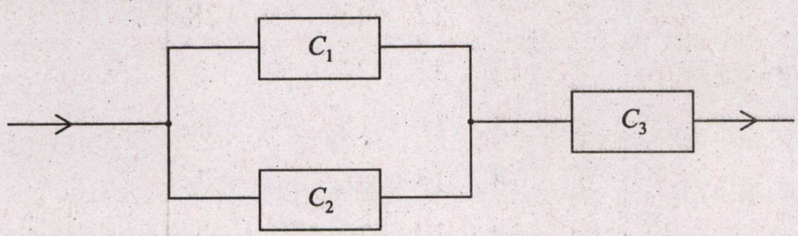
10

**2.** Answer any **two** questions of the following :

(a) (i) Find an expression for the optimum order quantity, Q*, in a single period inventory model without set-up cost when the demand rate is instantaneous and units are continuous.

A company sells a variety of cooked fishes by weight. It makes a profit of ₹35 per kg if the fishes are sold on the day of the catch. If the company fails to sell the fishes on the same day of catch, it disposes them on the next day at the loss of ₹30. If the demand is known to be rectangular between 2000 and 3000 kg, determine the optimum daily amount of inventory of cooked fishes, assuming that the demand is instantaneous. 15

(ii) Three independently functioning components are connected into a single system as shown below :




Suppose that the reliability for each of the components for an operational period of $t$ hours is given by

$R(t) = e^{-0 \cdot 03t}$.

(a) If $T$ is the time to failure of the entire system (in hours), what is the pdf of $T$?

(b) What is the reliability of the system ?

(c) How does the reliability of the system compare with $R(t)$ ? 10

(b) (i) What are the issues which could be answered using sensitivity analysis in linear programming problems ?

For the following LPP

Maximize $Z = 3x_1 + 2x_2$

subject to

$$\begin{aligned} x_1 + 2x_2 &\leq 6 \\ 2x_1 + x_2 &\leq 8 \\ -x_1 + x_2 &\leq 1 \\ x_2 &\leq 2 \\ x_1, x_2 &\geq 0 \end{aligned}$$

The optimal table is obtained as

| *i* | *Basis* | *Cost* | *Solution* | 3 | 2 | 0 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $x_1$ | $x_2$ | $s_1$ | $s_2$ | $s_3$ | $s_4$ |
| 1 | $x_2$ | 2 | 4/3 | 0 | 1 | 2/3 | –1/3 | 0 | 0 |
| 2 | $x_1$ | 3 | 10/3 | 1 | 0 | –1/3 | 2/3 | 0 | 0 |
| 3 | $x_3$ | 0 | 3 | 0 | 0 | –1 | 1 | 1 | 0 |
| 4 | $x_4$ | 0 | 2/3 | 0 | 0 | –2/3 | 1/3 | 0 | 1 |
| | | $Z$ | 38/3 | 0 | 0 | 1/3 | 4/3 | 0 | 0 |

(a) Find the status and worth of each resource.

(b) Find the possible range of the cost coefficient associated with variable $x_1$ for which the optimum solution remains unchanged. 15

(ii) Suppose that each of the three electronic devices has a failure law given by an exponential distribution with parameters $\beta_1$, $\beta_2$ and $\beta_3$. Suppose that these three devices function independently and are connected in parallel to form a single system.

(a) Obtain an expression for $R(t)$, the reliability of the system.

(b) Obtain an expression for the pdf of $T$, the time to failure of the system.

(c) Find the mean time to failure of the system. 10

(c) (i) Develop hazard functions when the life pattern of a system was described by

(a) exponential and

(b) gamma distribution

In each case, find $E(T)$, the mean time to failure. 15

(ii) Distinguish between 'pure' and 'mixed' strategies of players in a two-person zero-sum game. In order to get mixed strategies of the players $A$ and $B$ in the following rectangular game, solve the game graphically :

*Player B*

| | | | | |
|---|---|---|---|---|
| *Player A* | 1 | 3 | −3 | 7 |
| | 2 | 5 | 4 | −6 |

10

(d) (i) What do you mean by sequencing problem ? In this context, explain the terms

(a) total elapsed time,

(b) idle time on a machine and

(c) no passing rule.

Find the sequence that minimizes the total time required in performing the following jobs on three machines in the order $ABC$ :

| *Processing time (in hours) on* | *jobs* | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Machine $A$ | 8 | 3 | 7 | 2 | 5 | 1 |
| Machine $B$ | 3 | 4 | 5 | 2 | 1 | 6 |
| Machine $C$ | 8 | 7 | 6 | 9 | 10 | 9 |

15

(ii) Give the concept and definition of censoring. What are Type-I, Type-II and random censoring ?

Considering that each patient in a hospital has the same exponential density of survival and has the same point of entry into the study and the study is terminated after the survival time of the $d^{th}$ patient ($n \geq d$), find the maximum likelihood estimator of the parameter $\lambda$ of the exponential density. 10

## SECTION 'B'

### (Demography and Vital Statistics)

3. (a) Differentiate between Total Fertility Rate (TFR), Gross Reproduction Rate (GRR) and Net Reproduction Rate (NRR). How does the NRR indicate the growth of the population ? 10

(b) The following part of a life table contains four typographical errors, each of a single digit. Find the errors and correct them.

| $x$ | $l_x$ | $_5d_x$ | $_5q_x$ | $_5L_x$ | $T_x$ | $e_x^0$ |
|---|---|---|---|---|---|---|
| 20 | 95772 | 857 | 0·008953 | 476495 | 4922814 | 51·533 |
| 25 | 94265 | 699 | 0·007368 | 472566 | 4456319 | 47·975 |
| 30 | 94166 | 800 | 0·008496 | 468895 | 3983754 | 42·306 |

10

(c) Explain the meaning of vital statistics. Name some of the important vital statistics and give their usefulness and shortcomings. 10

(d) What do you understand by

(i) True infant mortality rate,

(ii) Neo-natal infant mortality rate,

(iii) Modified infant mortality rate ?

Distinguish between them in terms of their utility in measuring the infant mortality as accurately as possible. 10

(e) Explain the reason for preparing model life tables. In this context, mention what are United Nations Model Life Tables. 10

4. Answer any **two** of the following :

(a) Describe the logistic curve, generally used for the projection of a human population at a time $t$. Describe some of its properties. Illustrate how would you fit a logistic curve to the given set of data using Pearl and Reed method. 25

(b) What is a life table and what are its uses ? Explain various columns of a life table and mention their relationships.

Fill out the several columns of the following abridged life table by employing the usual definitions :

| $x$ | ${}_5q_x$ | $l_x$ | ${}_5d_x$ | ${}_5L_x$ | ${}_5m_x$ | $T_x$ | $e_0^x$ |
|---|---|---|---|---|---|---|---|
| 20 | 0·006338 | 94864 | – | – | – | 5024927 | – |
| 25 | 0·006650 | – | – | – | – | – | – |
| 30 | 0·008087 | – | – | – | – | – | – |
| 35 | – | 92879 | – | – | – | – | – |

25

(c) Why the standardization of death rates is sometimes necessary ? Explain the direct and indirect methods of standardization. Calculate

(i) crude death rate and

(ii) standardization death rate using direct method for the following data :

| *Age group (in years)* | *Population* | *Deaths* | *Standard population* |
|---|---|---|---|
| 0 – 5 | 2000 | 100 | 10000 |
| 5 – 10 | 1600 | 48 | 8000 |
| 10 – 25 | 2400 | 24 | 15000 |
| 25 – 45 | 6000 | 30 | 25000 |
| 45 and above | 8000 | 104 | 40000 |

25

(d) (i) What is stable population theory ? Define a stable population. How does it differ from the stationary population ? 10

(ii) Define crude birth rate, age-specific birth rates, general fertility rate and total fertility rate. State the utility of total fertility rate. 15

## SECTION 'C'

**(Survival Analysis and Clinical Trials)**

**5.** (a) If lifetime distribution of a device follows Weibull distribution with pdf

$$f(t) = \lambda \alpha t^{\alpha-1} e^{-\lambda t^{\alpha}}, \quad t>0; \lambda>0 \text{ and } \alpha>0.$$

Find hazard rate function. 10

(b) What are the principles of ethical clinical trials ? 10

(c) Establish the relationship between Life Table parameters $q_x$ and $\mu_x$. 10

(d) What is Cox proportional hazards model ? Show that ratio of hazard functions for any two subjects is constant. 10

(e) How survival function of Weibull distribution will take form of survival function for Cox proportional hazard model ? 10

**6.** Answer any **two** of the following :

(a) The joint distribution of deaths due to various risks $R_\delta(\delta = 1, 2, ..., k)$ as well as survivors in age group $(t_i, t_{i+1})$ is :

| | *Death due to various Risks* $R_1$ | $R_2$ .... | $R_\delta$ .... | $R_k$ | *Survivors* | *Total* |
|---|---|---|---|---|---|---|
| Frequency | $d_{i1}$ | $d_{i2}$ .... | $d_{i\delta}$ .... | $d_{ik}$ | $l_{i+1}$ | $l_i$ |
| Probability | $Q_{i1}$ | $Q_{i2}$ .... | $Q_{i\delta}$ .... | $Q_{ik}$ | $p_i$ | 1 |

Find the maximum likelihood estimators for $p_i$ and $Q_{i\delta}$. 25

(b) Suppose the data from a clinical trial consist of deaths at 2·1, 2·9, 3·6, 4·5, 5·6 and 6·9 months and censored observations at 3·0, 3·6, 6·9 and 9·1 months. Compute the Kaplan-Meier estimator of the survival function. 25

(c) Researchers would like to test a new therapy. They are planning to conduct a randomized clinical trial in which group $A$ receives the tested therapy and group $B$ receives a therapy that is currently in use. Our purpose is to test :

$H_0 : \mu_A = \mu_B$ against $H_1 : \mu_A > \mu_B$ with $\mu_A - \mu_B = \delta$ is minimum detectable difference. If $\alpha$ and $\beta$ are size of type-I and type-II errors respectively, show that sample size $(n)$ required for testing $H_0$ satisfies equations :

$$K = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \text{ and}$$

$$\Phi\left(K + \frac{\delta}{2\sqrt{\frac{\sigma}{n}}}\right) - \Phi\left(-K + \frac{\delta}{2\sqrt{\frac{\sigma}{n}}}\right) = \beta,$$

where $\sigma$ is population standard deviation and

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy.$$ 25

(d) Discuss various phases involved in a clinical trial. Also discuss the pros and cons of each phase. 25

## SECTION 'D'

**(Quality Control)**

7. (a) (i) What is the need for quality control ?

(ii) What are the advantages of Statistical Quality Control ? 10

(b) Derive the control limits for $\overline{X}$-chart. 10

(c) Determine the control limits for number of defectives when

(i) the standards are given

(ii) the standards are not given 10

(d) What do you mean by process capability ? Explain the different measures of process capability and write their interpretations. 10

(e) What is a CUSUM (cumulative sum) control chart ? State its advantages and disadvantages. 10

8. Answer any **two** of the following :

(a) (i) Explain how would you interpret Shewhart control charts to decide whether the process is in control or not. 10

(ii) When should the control charts for fraction defectives be prepared ?

How will you prepare the control charts for fraction defectives when

(a) the standards are given

(b) the standards are not given 15

(b) (i) Write a note on operating characteristic (OC) curve for control charts. 10

(ii) State the situation for use of Exponentially Weighted Moving Average (EWMA) control charts and also determine its control limits. 15

(c) (i) Mention the situations where Acceptance Sampling is most likely to be useful.

(ii) State the advantages and disadvantages of Acceptance Sampling.

(iii) Explain the operating procedure of single sampling plan. Also determine the effect of $n$ and $C$ on the OC curve. 25

(d) Explain the operating procedure of a Double Sampling Inspection plan. Derive its OC curve, Average Sample Number (ASN) and Average Total Inspection (ATI). 25

## SECTION 'E'
### (Multivariate Analysis)

**9.** (a) Let $(X_1, X_2)$ have a bivariate normal density with mean vector $\underline{\mu} = (\mu_1, \mu_2)'$, variance-covariance matrix $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$ and correlation coefficient $\rho_{12} = \text{corr}(X_1, X_2)$. Show that $X_1$, $X_2$ are stochastically independent if and only if $\rho_{12} = 0$. 10

(b) Let $\underline{X} \sim N_3(\underline{\mu}, \Sigma)$ with $\underline{\mu}' = (2, -3, 1)$ and $\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}$

(i) Find the distribution of $3X_1 - 2X_2 + X_3$.

(ii) Relabel the variables if necessary, and find a $2\times1$ vector $\underline{a}$ such that $X_2$ and $X_2 - \underline{a}'\begin{bmatrix} X_1 \\ X_3 \end{bmatrix}$ are stochastically independent. 10

(c) Find the maximum likelihood estimates of the $2\times1$ mean vector $\underline{\mu}$ and the $2\times2$ variance-covariance matrix $\Sigma$ based on the random sample

$$\underline{X} = \begin{pmatrix} 3 & 6 \\ 4 & 4 \\ 5 & 7 \\ 4 & 7 \end{pmatrix}$$

from a bivariate normal population. 10

(d) Let $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$ be a random sample of size $n$ from a $p$-variate normal distribution with mean vector $\underline{\mu}$ and variance-covariance matrix $\Sigma$. Then show that

(i) $\overline{\underline{X}} \sim N_p\left(\underline{\mu}, \left(\frac{1}{n}\right)\Sigma\right)$.

(ii) $(n-1)S$ is distributed as a Wishart distribution with $(n-1)$ degrees of freedom where $S = \frac{1}{(n-1)} \sum_{j=1}^{n} (\underline{X}_j - \overline{\underline{X}})(\underline{X}_j - \overline{\underline{X}})'$. 10

(e) Give formal mathematical definition of the $i^{th}$ pair of canonical variables and corresponding canonical correlation, for $i = 1, 2, ..., k$, in terms of the standard notations involving variance-covariance matrix $\Sigma$. Provide the necessary details, without proofs, for obtaining the canonical variables and their correlations. 10

**10.** Answer any **two** of the following :

(a) Let $\underline{X}_1$, $\underline{X}_2$, $\underline{X}_3$, and $\underline{X}_4$ be independent $Np(\underline{\mu}, \Sigma)$ random vectors.

(i) Find the marginal distributions for each of the random vectors

$$\underline{V}_1 = \frac{1}{4}\underline{X}_1 - \frac{1}{4}\underline{X}_2 + \frac{1}{4}\underline{X}_3 - \frac{1}{4}\underline{X}_4$$

and $\underline{V}_2 = \frac{1}{4}\underline{X}_1 + \frac{1}{4}\underline{X}_2 - \frac{1}{4}\underline{X}_3 - \frac{1}{4}\underline{X}_4$. 10

(ii) Find the joint distribution of the random vectors $\underline{V}_1$ and $\underline{V}_2$ defined in (*i*). 15

(b) (i) Let $\underline{X}_1, \underline{X}_2, ..., \underline{X}_n$ be a random sample from an $N_p(\underline{\mu}, \Sigma)$ population. Show that the Hotelling's $T^2$ statistic for testing $H_0 : \underline{\mu} = \underline{\mu}_0$ vs $H_1 : \underline{\mu} \neq \underline{\mu}_0$ is a generalization of the square of the *t*-statistic for testing $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ in the univariate case. 10

(ii) Evaluate $T^2$, for testing $H_0 : \underline{\mu}' = (7, 11)$, using the data

$$X = \begin{pmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{pmatrix}$$

Specify the distribution of $T^2$.

Test $H_0$ at the $\alpha = 0{\cdot}05$ level. What conclusion do you reach ? 15

(c) (i) Let $Y_1 = \underline{e}_1' \underline{X}, Y_2 = \underline{e}_2' \underline{X}, ..., Y_p = \underline{e}_p' \underline{X}$ be the principal components obtained from the variance-covariance matrix $\Sigma$. Obtain the correlation coefficient between $Y_i$ and $X_k$ where $i, k = 1, 2, ..., p$. Here $(\lambda_1, \underline{e}_1), (\lambda_2, \underline{e}_2), ..., (\lambda_p, \underline{e}_p)$ are the eigenvalue-eigenvector pairs for $\Sigma$. 15

(ii) Let $\Sigma = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

Determine the principal components $Y_1$, $Y_2$ and $Y_3$. What can you say about the eigenvectors (and principal components) associated with eigenvalues that are not distinct. 10

(d) (i) Consider the linear function $Y = \underline{a}'\underline{X}$. Let $E(\underline{X}) = \underline{\mu}_1$ and $\text{cov}(\underline{X}) = \Sigma$ if $\underline{X}$ belongs to population $\pi_1$. Let $E(\underline{X}) = \underline{\mu}_2$ and $\text{cov}(\underline{X}) = \Sigma$ if $\underline{X}$ belongs to population $\pi_2$. Let $m = \frac{1}{2}(\underline{a}'\underline{\mu}_1 + \underline{a}'\underline{\mu}_2)$. Given that $\underline{a}' = (\underline{\mu}_1 - \underline{\mu}_2)'\,\Sigma^{-1}$, show that

(a) $E(\underline{a}'\underline{X} | \pi_1) - m = \underline{a}'\underline{\mu}_1 - m > 0$

(b) $E(\underline{a}'\underline{X} | \pi_2) - m = \underline{a}'\underline{\mu}_2 - m < 0$ 15

(ii) Show that $-\frac{1}{2}(\underline{x} - \underline{\mu}_1)'\,\Sigma^{-1}(\underline{x} - \underline{\mu}_1) + \frac{1}{2}(\underline{x} - \underline{\mu}_2)'\,\Sigma^{-1}(\underline{x} - \underline{\mu}_2)$ is equal to $(\underline{\mu}_1 - \underline{\mu}_2)'\,\Sigma^{-1}\underline{x} - \frac{1}{2}(\underline{\mu}_1 - \underline{\mu}_2)'\,\Sigma^{-1}(\underline{\mu}_1 + \underline{\mu}_2)$. 10

## SECTION 'F'

**(Design and Analysis of Experiment)**

**11.** (a) Characterize a Completely Randomized Design (CRD). State its merits and demerits. 10

(b) Discuss the Statistical Analysis of $m \times m$ Latin Square Design for one observation per experimental unit. 10

(c) Give the Yates' method of computing effect totals in $2^3$ factorial experiment. 10

(d) State the advantages and disadvantages of confounding. 10

(e) Explain the Factorial Experiment with an example. 10

**12.** Answer any **two** of the following :

(a) For the two-way classification (with one observation per cell) model

$$X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \qquad \begin{pmatrix} i = 1, 2, \ldots. k \\ j = 1, 2, \ldots. h \end{pmatrix}$$

where $\varepsilon_{ij}$ is the error effect due to chance and $\sum_{i=1}^{k} \alpha_i = 0 = \sum_{j=1}^{h} \beta_j$

(i) Obtain the least square estimates of $\mu$, $\alpha_i$ and $\beta_j$.

(ii) Obtain the expectation of various sum of squares.

(iii) Also give the ANOVA Table. 25

(b) (i) The Table below gives the yield of wheat (kgs/plot) as observed in an experiment carried out in a 4×4 Latin Square. The Four manurial treatments are indicated by $A$, $B$, $C$ and $D$. Analyse the data and write a brief report on your findings. (Given that $F_{0\cdot05}(3, 5) = 5\cdot41$)

| | | | |
|---|---|---|---|
| *A* 12 | *C* 19 | *B* 10 | *D* 8 |
| *C* 18 | *B* 12 | *D* 6 | – |
| *B* 22 | *D* 10 | *A* 5 | *C* 21 |
| *D* 12 | *A* 7 | *C* 27 | *B* 17 |

15

(ii) Write down the special features of a Split-Plot Design. Give Statistical model and appropriate analysis of variance table for Split-Plot Design. 10

(c) (i) Define main effects and interactions with reference to a $2^3$ factorial experiment. 15

(ii) Explain a confounding scheme to run a $3^3$ experiment using blocks of 9 plots each by considering to confound the $AB^2C^2$ component and write the ANOVA for completely confounding $AB^2C$ in replicates. 10

(d) (i) Explain Strip-plot Design with layout for 3 replications and state its significance.

(ii) Illustrate the Strip-Plot Design as a randomization and layout with

(a) 3 varieties of wheat (Horizontal Treatments)

(b) 3 levels of nitrogen (vertical treatment)

(c) 6 Replicates

Also, outline the ANOVA Table. 25

## SECTION 'G'
### (Computing with C and R)

**13.** (a) The Fibonacci numbers are defined recursively as follows :

$$x_1 = x_2 = 1,$$
$$x_n = x_{n-1} + x_{n-2},\ n > 2.$$

Write a function that will generate and print the first $n$ Fibonacci numbers. 10

(b) Given bivariate data $(x_i, y_i)$, $i = 1, 2, ..., n$. Write a C-program to compute variance-covariance matrix $(\Sigma)$ of the data. 10

(c) Write a C-program that returns 1 if its argument is a prime number and returns 0 otherwise. 10

(d) Write R-code for fitting binomial distribution with parameters $n, p$ of data $(i, fi)$; $i = 0, 1, 2, ..., k$. 10

(e) Write R-code to find highest common factor of two positive numbers. 10

**14.** Answer any **two** of the following :

(a) Write a C-program for counting characters, words and lines in the input text. 25

(b) Write a C-program to construct linear linked list. Add functions to carry out the following operations on linear linked list.

(i) Count the number of nodes.

(ii) Write out contents.

(iii) Adding node at the end. 25

(c) Suppose $X$ follows Normal distribution with mean $a$, standard deviation $b$ and pdf $f_x(x)$, where $-\infty < a < \infty$, $b > 0$.

Write R-code to compute :

(i) $P(X > t) = \int_t^{\infty} f_x(x)dx$,

(ii) $MD = \int_{-\infty}^{\infty} |x - \text{mean}| f_x(x)dx$ 25

(d) In usual notations, consider the following linear model :

$$y_1 = \beta_1 - 2\beta_2 + 3\beta_3 + \epsilon_1$$
$$y_2 = 2\beta_1 + 3\beta_2 - \beta_3 + \epsilon_2$$
$$y_3 = -\beta_1 + 7\beta_2 + 4\beta_3 + \epsilon_3$$

Write an R-program to find the estimate of parameters and their variances. 25