1st year question (topic wise)

Descriptive statistics [Upto Moment, Skewness, Kurtosis]

1. How do ordinal data differ from nominal data? Explain with one example.
2. Describe the situation where you would recommend the use of (i) a multiple bar diagram (ii) a pie diagram. Illustrate by suitable examples
3. What do you mean by symmetric frequency distribution? Show that for a frequency distribution symmetric about a value "m", the arithmetic mean equals "m".
4. The AM and median of the radii of 50 circular discs are 3.1 cm and 3.2 cm respectively. How much can be said about the AM and median of the surface area of these disc? justify your answer.
5. If, for a collection of spheres, the frequency distribution of radicus is symmetric, show that the frequency distribution of volume is positively skew.
6. If a 3-point frequency distribution has the values x_0-1_2, $x_{0,}$ x_0+1_2 with respective frequencies (1-p)/, p, (1-p)/2. Calculation pearson's b_2 coefficient and find the limiting value of b_2 as p=0 and p=1. Comment on your findings.
7.
   (a) Indicate how central tendency, dispersion and skewness of a set of data on a continuous variable can be examined with the help of its cumulative frequency diagram.
   (b) Is the median of the logarithm of a set of positive real numbers equal to the logarithm of the median? What will be your answer for the A.M. ? Justify your answers.
8. Distinguish(with one example in each case) between (i) cross. Sectional data and time series data (ii) interval scale and ratio scale.
9. Indicate how central tendency, dispersion and skewness of a given data set can be examined with the help of a single diagram. Can you use the same diagram to detect possible outliers in the data set?
10. An assessee depreciated the machinery of his factory by 10% each in the first two years and by 40% in the third year and there by claimed 21% average depreciation relief from income tax department, where as the income tax officer objected and allowed only 20% which assessment would you like to support and why ?
11. The mean and the sd of life-time of bulbs produced by a factory are 'm' hours and 's' hours, respectively. Provide a loan bound for the percentage of bulbs having life-time between m-25 hours and m-25hours. Clearly state the relevant result.
12. Prove, by a geometrical argument, that for a J-shaped frequency distribution with its longer tail towards the higher value of the variable, the median is nearer to the first quantity than to the third what is the significance of this result.
13. Show that the mean deviation about mean cannot exceed the s.d. When are the two equal?
14. (a)Given the estimates of percentage of national income going to five different income groups (bottom 20%, next 20% middle 20%, next 20% and top 20%) in a country in the years 2000, 2005 and 2010, suggest a suitable diagrammatic method to represent the data and provide a rough sketch.

(b) Which measure of the average would you consider to be the most suitable in each of the following two cases :

(i) Size of ready made shoes

(ii)Percentage growth rates of an economy over consecutive years.

Justify your answers.

(c) Show that the mean deviation about A (MDA) based on n values may be obtained by the formula

$$n\ MD_A=(S_2-S_1)+A(n_1-n_2)$$

where S_1 is the sum of the values that are less than A and n_1 is the number of such values, while S_2 is the sum of the values that are greater than A and n_2 is the number of such values. Hence, show that MD_A is a minimum when A is the median.

15. (a) Explain two different real-life cases wher you would like to recommend coefficient of variation rather than s.d as the appropriate measure of dispersion.

(b) Show that the s.d s of a set of n values $x_1,x_2,\ldots\ldots\ldots\ldots,x_n$ is given by $ns^2= \sum_{j=1}^{i} x_j/i$ for

i=2(1)n

(c) Prove that $b_2>=1$ and $b_2-b_1-1>=0$, where b_1 And b_2 are the usual Pearson's measure of skewness and kurtosis respectively .

16. Distinguish (with an example in each case) between :

(a) Frequency data and non frequency data

(b) A discrete data and a continuous data

17. Agricultural production in a country comprises paddy , wheat , pulses and other items . suggest a diagrammatic method for studding how the total agricultural production of the country, as also the above four components , change over given period , indicating the method of drawing .

LINEAR ALGEBRA

1. Let D_n be a determinant of a matrix of order n in which the diagonal elements are all unity and in each row, the elements just above and just below the diagonal elements are x. The remaining elements of the determinant are zero. Show that $D_n - D_{n-1} + x^2 D_{n-2} = 0$ (5)
2. Define symmetric and skew-symmetric matrices. Show that a square matrix can be uniquely expressed as the sum of a symmetric matrix and a skew-symmetric matrix. (5)
3. Show that a real quadratic form $X'AX$ in n variables may be reduced to the form $\sum_{i=1(1)r} \lambda_i y_i^2$,where r=rank (A). (5)
4. Consider the two lines in the Euclidian plane given by the system of non-homogeneous equations involving two variables as $a_{11}x_1 + a_{12}x_2 = b_1$ and $a_{21}x_1 + a_{22}x_2 = b_2$.Writing the equations as : Ax=b ; derive the conditions involving the ranks of the matrix A and the augmented matrix (A|b) for the line (i)intersecting at a point, (ii)being parallel and (iii)being coincident. (5)
5. Define rank and basis of a vector space. Show that rank is unique but basis is not unique. State and prove the inequality involving ranks of the matrices A, B and AB. (4+6+5)
6. Define orthogonality and linear independence. Show that a set containing a null vector cannot be linearly independent. Explain the Gram-Schmidt orthogonalization process. Indicate a use of the method. (4+5+5+1)
7. Show that the Vandermonde determinant defined through the following square matrix

$$\begin{pmatrix} 1 & 1 & \cdots\cdots & 1 \\ d_1 & d_2 & \cdots\cdots & d_n \\ . & . & \cdots\cdots & . \\ . & . & \cdots\cdots & . \\ . & . & \cdots\cdots & . \\ d_1^{n-1} & d_2^{n-2} & \cdots\cdots & d_n^{n-1} \end{pmatrix}$$

has the value $\prod_{i<j} (d_i - d_j)$. (5)

8. Let a nonsingular matrix and U and V be two column vectors. Then $[A+UV^{/}]^{-1} = A^{-1} - \{(A^{-1}U)(V^{/}A^{-1}) / (1+V^{/}A^{-1}U)\}$ (5)

9. Define eigenvalues and eigenvector of a matrix. Show that eigenvectors corresponding to distinct eigenvalues of a matrix are orthogonal and if the matrix is positive definite then the eigenvalues are strictly positive. (5)

10. Define systems of non-homogeneous and homogeneous linear equations. Show that the solution for the non-homogeneous equation is unique when the corresponding homogeneous equation has the null vector as the only solution. (5)

11.(a) Suppose A and B are two square matrices of size n. Then they are similar if there exists a nonsingular matrix S of size n, such that $A=S^{-1}BS$. Show that these two similar matrices A and B have same eigenvalues. Show that the converse of the result is not true.

(b)Reduce any quadratic form $X^{/}AX$ to its canonical form. Show that if the matrix A is of Rank r, then only r such components will arrive in that form.

(8+7)

12.(a)Describe the Pivotal Condensation method for evaluation of a determinant of a square matrix $A=(a_{ij})$ of order m. Justify why this method is computationally more tractable than the usual formula given by $|A|=\Sigma\pm a_{1i}a_{2j}.......a_{mp}$,where the summation is taken over all permutations (i,j,.....p) of (1,2,……,m) with a plus sign if (I,j,……,p) is an even permutation and a minus sign if it is an odd permutation.

(b)Let V be a finite dimensional subspace and α , a vector not belongs to V. Show that there exists two vectors ϒ and β (which are unique) such that α=β+ϒ, β not belongs to V and ϒ≠0 is orthogonal to V. (8+7)

13. Prove that all bases of a vector space consist of the same number of vectors. (5)

14. Let A and B be two matrices such that the product AB is defined. Then prove that $r(AB) \leq \min\{r(A), r(B)\}$, where $r(A)$ denotes the rank of the matrix A. (5)

15. If A and B are two square matrices of the same order then prove that AB and BA have the same characteristic roots. (5)

16. If A and B are positive definite matrices of order n, prove that

$$|A+B|^{1/n} \geq |A|^{1/n} + |B|^{1/n} \qquad (5)$$

17.(a) Let $M = \begin{pmatrix} A & B \\ B' & C \end{pmatrix}$ be a symmetric matrix where A and C are symmetric square matrices. Show that M is positive definite if

(i) A is positive definite and
(ii) $C - B'A^{-1}B$ is positive definite.

(b) Prove that every square matrix A satisfies its non-characteristic equation.

(8+7)

18.(a) Prove that every real quadratic form $X'AX$, where X is an nx1 vector,

can be reduced by an orthogonal transformation to the diagonal form

$$\lambda_1 y_1^2 + \dots\dots\dots\dots + \lambda_n y_n^2$$

where λ_is (i=1,2,………..,n) are the characteristic roots of A.

(b) If A is a symmetric matrix and if P and Q are non-singular matrices such that $P'AP$ and $Q'AQ$ are diagonal matrices, then show that the number of positive elements in the two diagonal matrices is the same. (8+7)

19. Show that every vector space V containing n-vectors (not all null vectors) has a basis. (5)

20.(a) If E is a pxp symmetric and non-singular matrix, show that there exists a non-singular matrix F, such that $\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} = F'EF$

where the order of I is the number of positive characteristic roots of E and that of –I is the number of negative characteristic roots of E.

(b)Show that if A and B are real, symmetric and positive definite matrices of the same order, then so are (A-B) and $A(A-B)^{-1}B$. (8+7)

21.(a)State and prove a necessary and sufficient condition for solvability of a set of non homogeneous equations.

(b)Let $A=\begin{pmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ . & . & \cdots\cdots & . \\ . & . & \cdots\cdots & . \\ a_{n1} & a_{n2} & \cdots\cdots & a_{nn} \end{pmatrix}$ be a positive definite matrix.Then show that for any r(>n)

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{vmatrix} \begin{vmatrix} a_{r-1r-1} & \cdots & a_{r-1n} \\ a_{r-2r-1} & \cdots & a_{r-2n} \\ \cdots & \cdots & \cdots \\ a_{nr-1} & \cdots & a_{nn} \end{vmatrix} \geq |A|$$

(8+7)

22. Prove that if a vector space V has a basis, all its bases include the same number of vectors. (5)

23. Let A be an mxn matrix of rank m and B be an rxm matrix of rank r. Then show that for $r \leq \min(m,n)$, rank (BA)=r . (5)

24.(a)Prove that eigenvalues of a real symmetric matrix are real.

(b)Show that every system of linear homogeneous equations in n unknowns generates two vector spaces R and S with Rank (R) + Rank (S) =n where R is generated by the rows of coefficients and S consists of the solutions. Also show that a vector is orthogonal to R if and only if it belongs to S. (7+8)

25. Define skew-symmetric matrix. Show that the determinant of a skew-symmetric matrix is zero when its order is odd. (2+3)

26. When is a set of vectors said to be linearly dependent? Can a set of linearly dependent vectors be orthogonal? Show that zero-vector is a dependent vector. (2+2+1)

27. Suppose A is a positive-definite matrix. Show that (a) the diagonal elements of A are all positive. (b) the determinant of A is positive. (c) A^{-1} is a positive=definite matrix. (3+4+3)

28.(a)Discuss the Gram-Schmidt orthogonalization process to find an orthogonal set of vectors from a set of linearly independent vectors.

(b)Let A be a square matrix of order p and A' = Transpose of A. Suppose AA' = pI_p ,whereI_p is the identity matrix of order p. What can you say about A'A? (5+5)

29.(a)Show that for every matrix A, there are non-singular matrices P and Q such that PAQ= $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ where r is the rank of matrix A.

(b)Discuss when the system of non-homogeneous linear equations Ax=1 has (i) a unique solution , (ii) no solution . (6+4)

30. Let A be a square matrix with |A|=0 .Then find a relation among the columns of A. Justify whether there is any non-zero vector **x** for which A**x**=**0**. (2+3)

31. Let $(a_i-a_j)^2$ be the (i,j)-th element of a square matrix A of order p. Show that |A|=0 when p>3. (5)

32.(a)What is non-singular matrix? Show that the rank of a matrix remains unchanged by non-singular transformation.

(b)If $\boldsymbol{\alpha}$ is a non-null p-component column vector , then find the characteristic roots of $\boldsymbol{\alpha\alpha'}$.

(c)If r (0<r<infinitive) is the radius of convergence of the power series $\Sigma a_n x^n$, then discuss the behaviour of the power series over various subsets of

(-infinitive, +infinitive). [(1+2)+3+4]

33.(a)Show that the representation of a vector of a vector space in terms of its basis is unique.

(b)Discuss how the basis of a vector space cannot be unique.

(c)Discuss the nature of the characteristic roots of a positive –semi-definite quadratic form. (3+4+3)

34.(a)Define an improper integral over (i) an unbounded interval, and (ii) a bounded interval having a point of discontinuity at an interior point of the interval. Give an example of each case.

(b)Show that the inverse of a non-singular matrix is unique. [(2+2+3)+3]

35. Show that for every matrix A of rank r, there is a non-singular matrix P such that PA has r non-zero independent rows. (6)

36. Let S be any set of vectors and $(O)_s$ be the set of all vectors orthogonal to S . Show that $(O)_{(O)s}$ contains S. Discuss the situation where $(O)_{(O)s}$ is identical to S.

(6)

37.(a)Define a matrix. Show that commutative law does not hold good for matrix multiplication.

(b)When a quadratic form is said to be positive semi-definite? Explain the method of reduction of a positive semi-definite quadratic form to a sum of squares by a non-singular transformation. (4+6)

38. Find the value of c for which the following equations admit a solution

$4x_1 + 6x_3 = 1$

$4x_3 - 2x_2 = 7 + c$

$2x_1 - x_2 + 5x_3 = 4$ (4)

39.(a)Find the determinant of the following matrix:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2a & 3b & 4c & 5d \\ 4a^2+1 & 9b^2+1 & 16c^2+1 & 25d^2+1 \\ 8a^3-5 & 27b^3-5 & 64c^3-5 & 125d^3-5 \end{pmatrix}$$

(b)Let A be a square matrix of order p such that for some non null vector $\boldsymbol{\alpha}$, $A\boldsymbol{\alpha}=\mathbf{0}$. Show that for any square matrix B of order p, $|AB| = 0$. (5+5)

40. Define a vector space and a vector subspace with examples. (5)

41.(a)Define a negative definite matrix. For a negative definite matrix what can you say about the principle diagonal elements and the principal submatrices? Justify your answer.

(b)How can you examine the consistency or inconsistency of the system of equations Ax=b ? Cite examples of both. (6+4)

42. Define linearly dependent set of vectors. Show that any (m+1) vectors from the m-dimensional Euclidean space are linearly dependent. (5)

43. Show that for any square matrix, the determinant of the transposed matrix is the same as the determinant of the matrix itself. (5)

44. State a necessary and sufficient condition for the existence of a solution to $A^{nxn}\, x^{nx1} = b^{nx1}$. Show that if a system of linear equations has two distinct solutions then there exists an infifnite number of solutions. (5)

45.(a)Assume $A=\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}$ and $P=\begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix}$

(i) Show that P is an orthogonal matrix for any Θ.

(ii) Find Θ for which P'AP is a diagonal matrix and hence find A^t. (8)

(b)Show that for a matrix A^{nxn}, AA=A if and only if $R(A) + R(I_n-A) = n$, where I_n is the (nxn) identity matrix and R(A) is the rank of A. (7)

46.(a)If i-th and j-th rows of a matrix are alike ,prove that the i-th and j-th row of any product AB must be alike. (3)

(b)Consider a vector space V of the Euclidean space. What do you mean by the dimension of V? What role is being played by the basis of V? Given a basis of this vector space V, describe the method of finding an orthogonal basis. (6)

47.(a)Define rank of a matrix. Show that ,in a matrix, the number of linearly independent rows is same as the number of linearly independent columns.

(b)Find the angle between two non-zero vectors. Hence derive the Cauchy-Schwartz Inequality. (10+5)

48. Let A and B two matrices such that B is obtained from A by interchanging its first and second rows. Find a non-singular matrix P such that PA=B. (5)

49. Define a positive definite (p.d.) matrix. If Υ_{ij}=1 or 0 as i=or≠j , show that the matrix ($\Upsilon_{ij}+x_i x_j$), i,j=1,2,........, p is p.d. Stating the necessary results, show that the characteristic roots of a p.d. matrix are all positive. Hence establish that the sum of the diagonal elements of a p.d. matrix is positive. (15)

50. Prove that for any two square matrix A^{nxn} and B^{nxn} . $|AB|=|A||B|$ (5)

51. What is the echelon form of a matrix? Prove that the rank of a matrix is equal to the number of non-zero rows in its echelon form. (8)

52. Verify whether the following set of vectors are independent or not. Are they orthogonal? (1,-2,3), (5,6,-1), (2,2,1) (8)

53. Describe the methods of finding the rank of a matrix. (5)

54. Show that all vectors (x_1, x_2, x_3) in a vector space V_3 which obey $x_1-x_3=x_2$ form a subspace V. Then show that V is spanned by a_1=(1,0,1) and a_2=(0,1,-1). Find an orthogonal basis of V. Also find a vector which is orthogonal to all vectors in V. (8)

55. If $A^{pxp}=(a_{ij})$ is a real symmetric positive definite matrix show that $|A|\leq\Pi_{i=1(1)p}a_{ij}$ When does the equality hold? (8)

56. Suppose an mxn matrix A has rank r. Show that there exist non-singular matrices P and Q of appropriate orders such that $PAQ=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$

Hence show that for any two nxn matrices A and B ,

$\operatorname{rank}(AB) \geq \operatorname{rank}(A)+\operatorname{rank}(B)-n$ (8)

57. Consider the vectors $\mathbf{x_1}=(1,3,2)$ and $\mathbf{x_2}=(-2,4,3)$ in R^3. Show that the set spanned by $\mathbf{x_1}$ and $\mathbf{x_2}$ is given by $\{(\varepsilon_1,\varepsilon_2,\varepsilon_3):\varepsilon_1-7\varepsilon_2+10\varepsilon_3=0\}$ (8)

58. Let $A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ and $|A|\neq 0$, where A is of order pxp and A_{11} is of order qxq . Show that $|A|=|A_{11}|\,|A_{11}-A_{12}A_{22}^{-1}A_{21}|$ (8)

STATISTICS-HONOURS-PRACTICAL

SECOND PAPER

GROUP-B

01. Length of ear-head (in cm.) has observed for two varieties of wheat. The data are given in the following table:

Variety-1	10.7	9.8	9.3	8.6	10.4	9.8	11.3	8.4	11.2	8.2
Variety-2	9.3	10.0	9.7	10.1	9.8	10.7	11.3	10.1	11.0	9.5

(a) Compute a relative measure of dispersion for each of the two data sets and make a comment.

(b) Obtain a stem-and-leaf display for the combined set of data. (4+1)+3

02. A frequency distribution of some successful students in their university examination is given below:

Income	English Medium Schooling	Other Schooling
High Income Family	38	97
Low Income Family	13	61

Compute a measure of association between income of family and medium of schooling on the basis of the data and comment. (4)

03. Fifteen students of a honours class have assessed with respect to both a class-test (full marks = 25) and a seminar presentation (full marks =10) and the following data have been obtained:

Student Serial No.	Marks obtained in the class-test	Marks obtained in the seminar presentation
1	7	5

2	18	7
3	24	9
4	12	6
5	12	7
6	18	9
7	18	6
8	7	6
9	24	9
10	24	9
11	24	8
12	20	8
13	21	6
14	15	7
15	22	8

Use Spearman's rank correlation coefficient to find the degree of association between the performance in the class-test and that in the seminar presentation.

(6)

04. Let the rank of a matrix be r, where $A = \begin{pmatrix} 0.5 & 1.0 & 1.5 \\ 1.5 & 2.5 & 3.5 \\ 4.5 & 8.0 & 11.5 \end{pmatrix}$

(a) Express A as a sum of r matrices of unit rank each.

(b) Find an orthonormal basis of the row-space of A. (6)

05. Age-groups (in years) of mothers, female population (in '000) and number of live-births (in '00) to the mothers are given for a small country with total population 2286×10^3 .

Age-group (in years)	Female population (in '000)	No. of live-births (in '00)
15-19	85	23.4
20-24	70	145.4

25-29	73	167.4
30-34	76	102.2
35-39	75	51.3
40-44	72	14.2
45-49	67	1.0

Obtain (a) the crude birth rate, (b) the general fertility rate, and (c) the total fertility rate for the country. If the crude rate of natural increase is 10 (per 1000), what is the crude death rate for the country? (6)

06. Using computer and EXCEL package, obtain a scatter plot for the data on marks obtained in two class-tests on the same topic as given in the following table:

[Marks converted out of 15]

Roll No.	1	2	3	4	5	6	7	8	9	10
Class-Test 1	10.7	9.8	9.3	8.6	10.4	9.8	11.3	8.4	11.2	8.2
Class-Test 2	9.3	10.0	9.7	10.1	9.8	10.7	11.3	10.1	11.0	9.5

Obtain the linear regression equation of marks in class-test 2 on that in class-test 1 and interpret the regression coefficient thus obtained. (10)

07. The following are the weights of peas in hundredths of a gram in four pea pods each containing eight peas.

Pea pod 1: 43 46 48 42 60 45 45 49

Pea pod 2: 33 34 37 39 32 35 37 41

Pea pod 3: 56 52 50 51 54 52 49 52

Pea pod 4: 36 37 38 40 40 41 41 41

Calculate the coefficient of intra-class correlation. (6)

08. Compute the following life-table:

Age(x)	l_x	d_x	$1000q_x$	L_x	T_x	e_x^0
25	78046					39.6

26	77614	440				
27						
28	76723		6.06			
29						
30	75523				2750943	

Hence, determine the probability that a person of age 25 l.b.d. will die before reaching age 30 l.b.d.

(4+2)

09.Consider the quadratic form Q(x) = x'Ax , where

$$A = \begin{pmatrix} 1.64 & 1.20 & 4.95 \\ 1.20 & 2.01 & 1.28 \\ 4.95 & 1.28 & 5.38 \end{pmatrix}$$

(a) Determine the type of this quadratic form.

(b) Find the transformation that will diagonalize the quadratic form.

(2+4)

10. Obtain a solution of the system A**x**=b ,

$$\text{where } A = \begin{pmatrix} 0.1228 & -0.0508 & 0.1434 & 0.2013 \\ 0.1281 & 0.1342 & 0.0196 & 0.4703 \\ 0.0434 & 0.2172 & 0.1034 & 0.1054 \\ 0.2023 & 0.1523 & 0.2045 & 0.8562 \end{pmatrix}$$

with $\mathbf{x} = (\mathbf{x_1},\mathbf{x_2},\mathbf{x_3},\mathbf{x_4})'$ and b = (0.843,0.6713,1.31412,0.83235)' (6)

11.Using computer and EXCEL package (i) find summary statistics (i.e. one measure each of central tendency, dispersion, skewness and kurtosis) and (ii) draw histogram for the given data set of marks scored by 30 students.

31 55 48 47 53 48 33 32 42 55 44 38 60 65 71

80 41 53 47 48 55 30 31 34 42 51 35 35 26 25 (5+5)

12.A car dealer offers a number of second-hand models for sale. To estimate the price of a car of a given model, he uses a linear regression of price of the car on its mileage, based on the following data on 15 cars of the same model. Determine the fitted regression line by the method of least squares:

(4)

Mileage (x 1000 km)	0	11	14	24	24	31	55	50	5	60
Price (x 1000 rupees)	40	32	25	20	16	20	13	12	20	7

13.A chemical compound containing 12.5% of iron was given to two technicians A and B for chemical analysis. A made 15 determinations and B made 10 determinations of the percentage of iron. Their results are given in the following table:

Determinants of A			Determinants of B	
12.46	12.43	12.77	12.21	12.33
11.89	12.24	12.33	11.97	12.45
12.76	11.85	12.56	12.45	12.39
11.95	12.12	12.65	12.22	12.37
12.77	14.43	12.12	12.05	12.65

(a) Find separately for A and B the mean, sd and mean deviation about the median. Also find their respective coefficients of variation.

(b) Based o the above measures, prepare a report comparing the accuracy and consistency of the two technicians.

(4+2)

14. A drug, supposed to have some effect in curing diabetes was treated on 100 patients in a certain hospital and their records were compared with 100 other patients not treated with the drug. Study the efficacy of the drugs in curing diabetes: (3)

	Cured	Not Cured
Treated	54	46
Untreated	24	76

15. A part of a life table is given here with most of the entries missing. On the basis of the available figures, supply the missing ones:

Age(x)	l_x	d_x	$1000q_x$	L_x	T_x	e_x^0
10	93102		0.62			
11			0.66			
12			0.72			
13			0.80			
14			0.90			
15			1.00			
16			1.12			
17			1.23			
18			1.33			
19			1.40		4842446	

Hence determine the probability (a) that a child of age 10 will live at least 5 years more, (b) that two children aged 10 and 11 will each live at least 5 years more, and (c) that of two children aged 10 and 11 at least one will die within 9 years.

(6+3)

16. Find an orthogonal matrix that diagonalizes the following matrix:

$$A = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{pmatrix}$$

Hence or otherwise , find the determinant of A, its characteristic roots and vectors and comment on the definiteness of A. (8)

17. Using the computer and EXCEL package, draw box plot and histogram for the given data set of the monthly wages of 30workers in a factory and comment: (5)

830,835,890,810,835,836,869,845,898,890,820,860,832,833,855,

845,804,808,812,840,885,835,836,878,840,868,890,806,840,890

18. Using EXCEL package, draw the two regression lines for the following data: (5)

X:	14	19	24	21	26	22	15	20	19
Y:	31	36	48	37	50	45	33	41	39

19. The following table shows the height(x, in cm) and weights(y, in kg.) of 15 students.

Height(x)	152.5	157.5	160.0	170.0	175.0
Weight(y)	56.0	56.5	55.0	55.0	58.0
	61.5	59.0	62.5	65.0	63.0
	57.5		50.0		61.0
	58.5		56.5		

(a) Estimate the weight of a student of height 172 cm.

(b) Calculate the correlation coefficient of weight and height.

(c) Calculate the correlation index of order 2 as well as the regression equation of weight on height.

(d) Calculate the correlation ratio of weight on height. (1.5x4)

20. An office clerk is regularly given 10 individual routine tasks to carry out the time taken for each of these tasks is observed in two occasions, once in the morning and once in the afternoon. We wish to know whether the time taken is longer in the afternoon than in the morning.

Task	A	B	C	D	E	F	G	H	I	J
Time taken in morning (X)	1.5	2.5	3.0	3.8	5.0	5.6	6.0	7.0	8.5	9.0
Time taken in	3.7	3.6	5.0	6.5	8.3	8.3	7.6	10.8	10.0	12.0

afternoon (Y)										

(a) Draw a scatter plot of the time taken in the afternoon against the time taken in the morning and comment briefly on any relation.

(b) Determine the least squares estimates of the parameters in the simple linear regression model $Y_i = a+bX_i+e_i$.

(c) What proportion of the variation in the data is explained by the model?

(7)

21. A study was conducted to find out what people of different educational level feels about the role of religion or caste in social life.

Educational Level	Role of Religion or Caste			
	No. Role	Little Role	Large Role	Complete Role
Group I (upto class 8)	2	8	25	36
Group II (upto HS)	10	15	14	12
Group III (Graduate)	29	24	2	1

Compute a suitable measure to find if there is association between educational level and the individual's perception towards the role of religion/caste in social life. Interpret your finding.

(3)

22. Let X be a symmetric random variable with 3 as the point of symmetry and coefficient of variation 50%. Find the minimum probability for which X lies in the interval [2.445 , 3.555]

(3)

23. An examination has 3 multiple-choice questions (Q1-Q3) of increasing difficulty. Each question has 4 alternatives, A, B, C and D, only one of which is correct. The probability of a student knowing the answer to Q1 is 0.8, to Q2 is 0.75 and to Q3 is 0.72. If a student does not know the answer he chooses one of the alternatives randomly. Find the probabilities that a student

(a) answers all three questions correctly.

(b) knows all the answers given that he has answered all of them correctly.

(3)

24. Find the transformation that diagonalizes the quadratic form

$Q=4x_1^2-2x_2^2+x_3^2+5x_4^2-2x_1x_2+4x_1x_3+3x_1x_4-5x_2x_3-2x_2x_4+6x_3x_4$ (4)

25. Find the solution to the following system of equations:

$$9.8x + 2.14y + 1.67z + 5.82 = 0$$

$$2.14x + 10.62y + 7.64z + 4.21 = 0$$

$$1.67x + 7.64y + 33.08z - 2.67 = 0$$ (3)

26. Use EXCEL package to display the following data through two diagrams, one highlighting the absolute feature and the other the relative feature of the data. (Give print-outs)

REGION	NO. OF MALARIA CASES
North	27
South	84
Central	12
East	73
West	36

27. Consider the following data on marks scored by 30 students in an examination:

41 55 48 47 53 48 33 32 42 55

44 38 60 65 71 80 41 53 47 48

55 20 31 34 42 51 35 35 26 25

Using EXCEL package,

(i) Draw the histogram and comment. (Give print-out)

(ii) Find the mean and standard deviation.

(iii) Obtain three different measures of skewness, one based on the moments, one on the quantiles and one on the central tendencies respectively and compare. Also find the kurtosis and comment on the nature of the distribution.

(7)

28. The following table gives the scores of 16 students on a psychological test X and an achievement test Y.

X	65	68	67	42	58	72	70	75	47	52	93	71	69	86	63	79
Y	78	61	62	48	50	72	42	76	70	55	123	65	73	110	77	86

(a) Calculate the product moment correlation coefficient and Spearman's rank correlation coefficient and comment.

(b) Calculate the medians for X and Y.

From a two-by-two contingency table of classifying the paired data points with respect to their components being greater than or less than the corresponding medians.

From this table find a suitable measure of association and comment.

29. The following table shows the quantities consumed and the values (price x quantity) of 5 commodities for three successive years.

Commodities	2005		2006		2007	
	quantity	value	quantity	value	quantity	value
I	50	350	60	420	70	490
II	120	600	140	700	160	800
III	30	330	20	200	15	225
IV	20	360	15	300	10	220
V	5	40	5	50	5	60

Calculate the price index number for 2007 taking 2005 as the base period adopting chain base formula and Paasches' formula at each stage. Also verify whether the circular test is satisfied by the Paasches' formula or not on the basis of the above data.

30. The population of a certain country, as recorded in each of the ten decennial censuses, is shown below:

Year	Population (in million)	Year	Population (in million)
1911	238.3	1961	361.0
1921	252.0	1971	439.1
1931	251.2	1981	547.0

1941	278.9	1991	683.8
1951	318.5	2001	823.8

Fit a logistic curve to the data and comment on the nature of fit.

31. Consider the following matrix:

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 2 & 3 & 5 & -1 \\ 8 & 13 & 21 & -3 \\ 13 & 8 & 5 & -2 \end{pmatrix}$$

Find: (a) Rank (A), (b) An orthogonal basis of the row space of A.

32. Data from a case-control study of esophageal cancer in Ile-et-Vilaine, France is given below. The data records 45 combinations of age groups and alcohol consumption (alcgp) in gms/day. Fit an appropriate regression line to the data. Also compute the value of Pearson's product moment correlation coefficient and the appropriate correlation ratio. Comment on your findings.

Sl No.	Age-group	alcgp
1	25-34	0-39
2	25-34	0-39
3	25-34	0-39
4	25-34	0-39
5	25-34	40-79
6	25-34	40-79
7	25-34	40-79
8	25-34	40-79
9	25-34	80-119
10	25-34	80-119
11	25-34	80-119
12	25-34	120-159
13	25-34	120-159
14	25-34	120-159
15	25-34	120-159
16	35-44	0-39
17	35-44	0-39
18	35-44	0-39

19	35-44	0-39
20	35-44	40-79
21	35-44	40-79
22	35-44	40-79
23	35-44	40-79
24	35-44	80-119
25	35-44	80-119
26	35-44	80-119
27	35-44	80-119
28	35-44	120-159
29	35-44	120-159
30	35-44	120-159
31	45-54	0-39
32	45-54	0-39
33	45-54	0-39
34	45-54	0-39
35	45-54	40-79
36	45-54	40-79
37	45-54	40-79
38	45-54	40-79
39	45-54	80-119
40	45-54	80-119
41	45-54	80-119
42	45-54	80-119
43	45-54	120-159
44	45-54	120-159
45	45-54	120-159

33. A study was conducted to find out what people of different educational level feel about the role of religion or caste in social life.

Edu. Level/Role	No Role	Little Role	Large Role	Complete Role
Group I (upto class 8)	2	8	25	36
Group II (upto HS)	10	15	14	12
Group III (Graduate)	29	24	2	1

(a) Compute a suitable measure to find if there is any association between educational level and the individual's perception towards the role of religion/caste in social life. Interpret your finding.

(b) Merge the first two columns "No Role" and "Little Role" and rename it as "Minor Role". Similarity merge the last two columns "Large Role" and "Complete Role" and rename it as "Major Role". Find the Odds Ratio of (i) Group II with Group I and (ii) Group III with Group I, with respect to Minor and Major Roles and interpret your results.

34. The following table shows the group indicates and the corresponding weights for the year 1995, with 1981 as the base (=100), of a given community.

Group	Group Index	Weight
Food	212.45	65.3
Clothing	328.06	4.8
Fuel and light	345.89	8.5
House rent	173.41	7.6
Miscellaneous	201.35	13.8

(a) Find the cost of living index for the year 1995.

(b) What is the purchasing power in 1995 as compared to 1981?

(c) If Mr. Ghosh's salary increased from Rs 7500.00 in 1981 to Rs 15000.00 in 1995, how had his economic status changed?

35. Consider the following data set for two countries.

Country-1

Age group	Population size (in '000)	Number of males (in '000)	Number of deaths	Number of deaths in males
<1	100	50	750	400
1-4	400	200	3000	1600
5-20	1500	750	12000	6500
21-100	4000	1500	31000	12000

Country-2

Age group	Population size (in '000)	Number of males (in '000)	Number of deaths	Number of deaths in males
<1	15	7.5	12	6
1-4	60	30	45	24
5-20	200	100	160	85
21-100	440	170	350	200

(i) Compute the crude death rates for both the countries.

(ii) Compute the age specific death rates for both the countries separately for Male, Female and Total population.

(iii) In order to compute the mortality situation of the two countries propose a good measure and calculate its value for both the countries.

36. (a) Let $\mathbf{a_1} = [1 \;\; 5 \;\; 7]_{3x1}$; $\mathbf{a_2} = [4 \;\; 0 \;\; \alpha]_{3x1}$; $\mathbf{a_3} = [1 \;\; 0 \;\; 0]_{3x1}$. For what values of α will the above vectors form a basis for E^3?

(b) Let **W** = {(2,0,3,1,1), (1,0,2,1,1), (2,0,3,1,3)}

V = {(2,1,1,0,1), (3,2,3,2,3), (1,1,1,1,1)}

Obtain a basis and the dimension of (i) [**W**] (ii) [**V**] (iii) [**W**]+[**V**] ,where [**X**] denotes the span of **X**.

37. The following table gives data on income in thousand dollars (x), the number of families (N) at income x and the number of families owning a house (n).

x: 10 13 15 20 25 30 35 40

N: 60 80 100 70 65 50 40 25

n: 18 28 45 36 39 33 30 20

Suggest an appropriate regression equation to explain the effect of income on owning a house. Also estimate the parameters of this equation using the above data and predict therefrom the proportion of families at income 32 thousand dollars who own a house.

38. The following data relate to the urban middle-class people of a particular region in the years 2003 and 2005.

Group	**Percentage of total expenditure**	**Group Index (base : 2000)**	
		2003	**2005**
Food	35	118	122
Clothing	15	112	118
Housing	20	113	115
Transport	10	112	117
Durable Goods	8	105	110
Others	12	120	125

(a) Compute the cost of living indices for 2003 and 2005 with base year2000.

(b) If a family saved 15% of its monthly income in 2003, find the relative change in its average savings over the years 2003-2005, assuming that it maintained the same standard of living as in 2003.

39. From the following data relating to a particular community compute the gross Reproduction Rate and the Net Reproduction Rate. Interpret your results.

Age of Mother	**No. of women**	**No. of births**	**Survival factor**
15-19	9000	140	0.920
20-24	9200	1312	0.914
25-29	8900	1067	0.908
30-34	8600	771	0.891
35-39	8400	468	0.878
40-44	8500	160	0.869

Assume that 48.7% of the total are female births. Survival factor gives the rate of survival from birth to the mid-point of the corresponding age-group.

40. Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 5 & 7 \\ 9 & 16 & 23 \end{pmatrix}$

Suppose rank (A) =r

(i) Find r.

(ii) Write A as the sum of r matrices each of rank unity.

(iii) Find an orthogonal basis of the row space of A.

41. An experiment was carried out by an agricultural farm to determine the most economic dose of a certain fertilizer for cultivation of wheat. The table below shows the net profit in Rs./plot (y) when dose (x) in some suitable units are used.

Dose:	5	10	15	20	25	30	35	40
Profit:	126	180	230	252	263	253	230	194

(a) Draw a scatter diagram for the above data and fit a suitable regression equation of y on x. Also plot the fitted values on the same diagram.

(b) Obtain an appropriate measure of correlation between the two variables x and y.

(c) Derive the most economic fertilizer dose.

42. For a stationary population with cohort l_o = 100000 , out of the children born in 1990, the number of children decreased in 1990 was 20000 and the number of children decreased in 1991 was 5000. Given below are the age-specific mortality rates of this population for the following ages x (last birthday):

x:	0	1	2	3	4
m_x:	0.4075	0.0123	0.0024	0.0018	0.0013

(a) Compute the complete expectation of life at age 4,given the same at birth is 65.90 years.

(b) What is the chance that two newborn babies would survive 4 years after their birth?

43. Check whether the following vectors are linearly independent or not:

$\boldsymbol{\alpha_1}$ = (0,1,2,3), $\boldsymbol{\alpha_2}$ = (2,-1,5,4), $\boldsymbol{\alpha_2}$= (4,0,6,1), $\boldsymbol{\alpha_3}$ = (0,-2,4,7)

Also find a vector in the space orthogonal to the vector space spanned by $\boldsymbol{\alpha_1}$,$\boldsymbol{\alpha_2}$, $\boldsymbol{\alpha_3}$, $\boldsymbol{\alpha_4}$.

44. Practical Note Book. (5)

45. Viva-Voce. (5)

PROBLEMS ON PROBABILITY THEORY PART-1

1. What is sigma field?Give an example of sigma field and an example which is not a sigma field ? Show that every sigma field is a field. [5]

2. Give the Kolmogorov's axiomatic definition of probability . Use it to show that P (ϕ)=0 and P(A)+P(B)>=P(AUB),where A and B are any two events associated with corresponding probability space and ϕ is null event. [5]

3. Distinguish between mutual independence and pairwise independence of a set of events citing examples. [5]

4. (a) Consider an experiment with sample space $\Omega=\{\omega 1,\omega 2,\ldots\ldots,\omega \mathbf{n}\}$,where it is known that outcome $\omega\mathbf{j}$ is twice as likely as outcomes $\omega\mathbf{j-1}$,j=2,.......,n.If A_k $=\{\omega 1,\omega 2,\ldots\ldots,\omega \mathbf{k}\}$, show that the probability of occurrence of A_k is $(2^{k}-1)/(2^{n}-1)$.

(b) State and prove the Theorem of Total probability.

(c)Let $B_{1,\ldots\ldots}B_n$ be mutually disjoint and B= $\bigcup_{i=1}^{n} B_i$ Suppose $P(B_i)>0$ and $P(A/B_i)=P$,i=1,2,.........,n. Show that P(A/B)=p

(d) A single card is removed from a deck of 52 cards.From the remainder two cards are drawn at random and is found that both are spades.What is the probability that the first card removed was also a spade. [4+4+3+4=15]

5. Give the classical definition of probability.What are its limitations?Suggest any further development. [2+2+1=5]

6. Show that A_1,$A_2$and A_3 are mutually independent events if and only if A_1^c,A_2^c, and A_3^c are mutually independent events.

7. (a) Use Kolmogorov's axiomatic definition of probability to establish Boole's in-equation.

(b) Two fair dice are thrown.Let A be an event that the first throw shows an odd number, B be the event that the second throw shows an even number, and C be the event that either both are odd or both are even.Show that A, B, C are pairwise independent but not mutually independent. [5+5=10]

8. Define conditional probability.Show that it satisfies Kolmogorov's axioms on the definition of probability.

9. Consider a collection of (N+1) urns each containing a total of N red and white balls;the urn number K contains K red and (N-K) white balls (K=0,1,2,.....,N). An urn is chosen at random and 11 random drawings are made from it,the ball drawn being replaced each time.Suppose that all n balls turn out to be red.then find the (conditional) probability that the next drawing will also yield a red ball. [5]

10. Construct four events A,B,C,D in a random experiment such that 0<P(A)<P(A/B)<1 and 0<P(C/D)<P(C)<1 and comment.

11. (a) For any two events A and B ,show that max{0,P(A)+P(B)-1<=P(A B)<=m;n{P(A),P(B)}<=max{P(A)P(B)}<=P(AUB)<=m;n{P(A)+P(B),1}

(b) Suppose n balls are distributed randomly into N cells.What is the probability that exactly K cells will be occupied and the remaining (N-K) cells will remain empty? [5+5=10]

12. (a) Describe the intuitive idea of probability in terms of relative frequency.What do you mean by 'events' and 'elementary event'?Describe with examples. [5]

(b) A bar of unit length is broken into three parts.Find the probability that a triangle can be formed from the resulting parts. [5]

13. There are two batches of articles of the same kind.The first batch contains of N articles of which n are defective and the second batch consists of M articles of which m are defective,α and β articles are selected at random from the two batches respectively(α<N,β<M).Three articles are selected at random from the mixed batch of ($\alpha+\beta$)

articles.Find the probability that at least one of the three articles is defective. [5]

14. Define mutually exclusive ,exhaustive and mutually independent events.let the two events be mutually exclusive.Are they mutually independent? [5]

15. R indistinguishable balls are distributed randomely in n cells.Show that the probability that exactly m cells remain empty is $P_m=({}^{n}{}_{m})\text{x}({}^{r-1}{}_{n-m-1})/({}^{n+r-1}{}_{r})$.

16. In a cinema hall that can accommodate n+kpeoples,n peoples are seated.What is the probability that r<=n given seats are all occupied? [4]

17. On a very rainy day, n members of a society came tp what eventually turned out to be a very stormy meeting.At the end of the meeting,each member took one of n umbrellas that they had left in the anterrrom.assuming that eachj member in his tired state of body and mind,chose an umbrella at random.What is the probability that none carried away his own umbrella? [5]

18. Three urns of the same appearance have the following proportions of white and black balls.

Urn 1:1 white,2 black

Urn 2:2 white,1 black

Urn 3:2 white,2 black

One of the urns is chosen at random and a ball is drawn out of it also at random.It turns out to be white.What is the probability that urn 'i' was chosen ,for i=1,2 and 3. [5]

19. A card is chosen at random from a deck of playing cards.Show that the appearance of an ace and the appearance of a spade are independent events. [4]

20. Let P(B)>0,P(C)>0.If B and C are independent ,show that P(A/B)=P(C).P(A/B C)+P(C^c).P(A/B C^c). Conversely if this relation holds P(A/B C)=P(A/B) and P(A)>0, then show that B and C are independent. [5]

Probability I

[2015]

1) Define Conditional Probability. Show that it satisfies Kolmogrov's axioms on the definition of Probability. [2+3]

2) In a sample space of eight equally likely points find:
   a. Three events that are pair wise independent but not mutually independent.
   b. Three events those are mutually independent.

[5]

3) A biased coin with probability $p (0<p<1)$ of success is tossed until for the first time the same result occurs three times in succession . Find the probability that the game will end at the 7^{th} throw.

[5]

4) Derive an expression for the probability of realisation of exactly k out of n events which are not necessarily mutually exclusive.

[5]

5) Discuss Bayes' theorem on Probability.

[5]

6) Missiles are fired at a target. The probability of each missile hitting the target is p. The hits are independent of one another. Each missile which reaches the target brings it down with probability p_1. The missiles are fired until the target is brought down or the missile reserve is exhausted. The reserve consists of n (>2) missiles. Find the probability that no less than two missiles remain in the reserve after the target is brought down. [5]

7) There are two batches of articles of the same kind. The first batch consists of N articles of which n are defective and the second batch consists of M articles of which m are defective, α and β articles are selected at random from the two batches respectively ($\alpha<N$, $\beta<M$). Three articles are selected at random from the mixed batch of $(\alpha+\beta)$ articles. Find the probability that at least one of the three articles is defective.

[5]

8) State and prove Boole's inequality.

[5]

[2014]

9) State the axiomatic definition of probability. Let for two events A and B. $A \Rightarrow B$. Find the relationship between P(A) and P(B).

[5]

10) Suppose each of three persons tosses a coin. If the outcome of one of the tosses differs from the other outcomes, then the game ends. If not, then the person starts over and re-tosses their coins. Assuming fair coins. Find the probability that the game will end with the first round of tosses. If all three coins are biased and have the probability ¼ of landing heads. What is the probability that the game will end at the first round. [5]

11) Distinguish between pair wise independence and mutual independence. Show that if A and B are independent then so are A and B^c, A^c and B and A^c and B^c.

[5]

12) You are given that at least one of the events A_r , r= 1,2,.....,n is

certain to occur. If $P(A_r)$=p and $P(A_r \cap A_s)$ = q ,r≠s show that p≥ 1/n and q≤2/n.

[5]

13) Suppose all n persons at a conference deposit their mobile at a counter. Afterwards, each person is forced to select a mobile at random. Find the probability that none of the n persons select the right mobile. [5]

14) State and prove the multiplicative law.

[5]

15) With the help of Bayes' theorem, show that for any event 'S'.

[5]

$$P(S|T) = \frac{\sum_{i=1}^{n} P(Ai)P(T|Ai)P(S|Ai \cap T)}{\sum_{t=1}^{n} P(At)P(T|At)}$$

Where $A_1, \dots, A_n$ aren mutually exclusive and exhaustive events and event T is such that $P(A_i \cap T)>0$ for all i≠1,2,....,n

16) Consider n independent trials in which each trials results in one of the outcomes 1,2,....., k(<n) with respective probabilities $p_i \geq 0$, i=1,2,...,k.

[5]

[2013]

17)) Give Kolmagrov 's axiomatic definition of probability. Use it to show that P ($\boldsymbol{\Phi}$) = 0 and P (A)+P (B) ≥ P (A$\cup B$). Where A and B are two events associated with the corresponding probability space and $\boldsymbol{\Phi}$ is a null set. [5]

18) A hospital administrator codes incoming patients suffering gunshot wounds according to whether they have insurance (coding 1 if they do and 0 if they do not) and according to their condition, which is rated as good(g) , fair(f) or serious(s). Consider an experiment that consists of coding of such a patient.

i) Give sample space of the experiment.
ii) Let A be the event that the patient is in serious condition. Specify the outcomes in A.
iii) Let B be the event that the patient is uninsured. Specify the outcomes in B.
iv) Give all outcomes in the event $B^c \cup A$.

[5]

19) Two football teams A and B play a series of independent games until one of wins the games. The probability of each team winning in each game equals ½. Find the probability that the series will end. (i) In at most 6 games. (ii) In 6 games given that team A won the first two games.

[5]

20) Show that a pair of events A and B cannot be simultaneously mutually exclusive and independent. [5]

21) A six faced die is thrown twice

a) Write down the probability space of the experiment.
b) Let B be the event that the first number thrown is no larger than 3 and let C be the event that the sum of the two numbers thrown equals 6. Find the probabilities B and C and the conditional probabilities C given B and of B given C.

[3]

22) Let $\Omega = \{a, b, c, d\}$ and let P assign probabilities 0.2, 0.3, 0.1, 0.3 and 0.1 respectively . Let A={a, b,c} and ={b,c,d}. Find P (A), P(B), P(A|B) P(B|A) andP(A^c|B^c).

[7]

23) Show that

a) P(A^c|B)=1-P(A|B)

b) PA$\cup B|C$)=P(A|C)+P(B|C)-P(A∩B|C)

Also show by means of counter examples that the following equations need not be true.

c) P(A|B^c)=1-P(A|B)

d) P(C|A$\cup B$)=P(C|A) + P(C|B)

[2+3+2+3]

24) A ball is drawn at random from a box containing balls numbered 1,2,.....,n.

a) What is the probability that its label is divisible by 3 or 4?

b) Examine the case in (i) as n→∞.

[3+2]

25) Establish the monotonicity property of the probability measure using Kolmogrov's axioms. [3]

26) Show that P(A Δ C)≤ P(A Δ B)+ P(B Δ C)

[4]

27) In terms P(A), P(B), P(C), P(A∩B), P(B ∩ C) and P(A ∩ B ∩ C) express for k= 0,1,2,3, the probabilities that :

a) Exactly k of events a, B and C occur.

b) At least k of the events A,B and C occur.

[5]

28) Two fair dice are thrown. Consider three events E, F and G. The event 'E' indicates that the first throw shows an odd number, 'F' indicates that the second throw shows an even number, and 'G' be the event that either both are odd or both are even. Show that E , F and G are pair-wise independent but not mutually independent.

[5]

29) A laboratory blood test is 95% effective in detecting a certain disease when it is in fact, present. However , the test also yields a 'false positive' result for1% of the healthy persons tested. (That is, if a healthy person is tested, then, with probability 0.01, the test result will imply that the person has the disease). If 0.5% of the population actually has the disease, what is the probability a person has the disease given that test result is positive? [5]

30) One of the sequence of letters AAAA, BBBB , CCCC is transmitted over a communication channel with the probability p_1p_2,p_3 ($p_1+p_2+p_3$ =1). The probability that each transmitted letter would be correctly understood is α, and the probabilities that the letter would be confused with other two letters are equally probable that AAAA was transmitted when ABCA was received.

[5]

31) State and prove total theorem of probability.

[5]

32) Let B= $\bigcup_{i=1}^{n} Bi$, where B_i, i=1,2,.........,n. Are mutually disjoint. If P(A|B_i)=p and P(B_i) > 0 for all i, then show that P(A|B)=p.

[4]

33) State and prove Boole's inequality. Hence or otherwise prove Bonferroni's inequality. [6]

34) Write a note on Bayes's theorem.
[4]

35) Candidate A wins an election by securing n votes, defeating her nearest rival candidate B who secures m votes (n>m). Assuming that all oderings are equally likely, show that the probability that A is always ahead in count of votes is (n-m)/(n+m).
[5]

[2012]

36) Let F and G be sigma field of subsets of Ω. Show that $(F \cup G)$. The collection of subsets of Ω lying in either F or G is not necessarily a sigma-field.
[5]

37) Find two fields such that their union is not a field.
[5]

38) Describe the underlying probability space when two balls are drawn without replacement from an urn containing two white and two black balls.
[5]

39) The events A is said to be repelled by the event B if PA|B) <P(A), and to be attracted by B if P(A|B)>P(A). Show that if B attracts A, then A attracts B and B^c repels A, if A attracts B and B attracts C. Does A attract C? [5]

40) Waiting in a line for a morning movie shows are 2n children. Tickets are priced at a quarter for each. Find the probability that nobody will have to wait for a change if before a ticket is sold to the first customer. The cashier has 2k (k<n) quarters. Assume that it is equally likely that each ticket is paid for with a quarter or a half dollar coin. [5]

41) Let A and B be two events such that P (A) = p_2>0 and P (B) = p_2>0, and $p_1 + p_2 > 1$. Show that P(A|B)≥ 1-[(1-p_2)/p_1]
[3]

42) Each of n urns contains four white and six black balls, while another urn contains five white and five black balls. An urn is chosen at random from the (n+1) urns and two balls are drawn from it, both being black. The probability that five white and three black balls remain in the urn is 17. Find n.
[6]

43) A fair coin is tossed three times. Let A= {at least one of the first two tosses is a head}, B={ same result on tosses 1 and 3}, C={no heads}, D={ same result on tosses 1 and 2}. Among these four events there are six pairs. Which of these pairs are independent? Which are mutually exclusive?
[6]

44) If P (A_n) = 1, 2, 3.... Then P ($\bigcap_{n=1}^{x}$ An) = 1.
[3]

45) Let Ω = {1, 2, 3, 4} with uniform probability and A = {1, 2}. List all B ⊂Ω such that A and B are independent. [[3]

46) Prove that P (AB) = P (BA)P (A)P (B) whenever P (A)P (B)≠0.

Show that, if P (AB)>P (A), then P (BA) > P (B). [3]

[2011]

47) What is sigma field? Give an example of sigma field and an example which is not a sigma field. Show that every sigma field is a field. [5]

48) Give the Kolmogrov's axiomatic definition of probability. Use it to show that P($\boldsymbol{\Phi}$)=0 and P(A)+P(B)≥P($A \cup B$) where A and B are any two events associated with corresponding probability space and $\boldsymbol{\Phi}$ is a null event. [5]

49) Distinguish between mutual independence and pair wise independence of a set of events citing examples. [5]

50) Define discrete random variable and its probability distribution with examples. [5]

51) A single card is removed from a deck of 52 cards. From the remainder, two cards are drawn at random and is found that both are spades. What is the probability that the first card removed was also spade. [4]

52) A coin is tossed repeatedly, and head appears at each toss with probability p, 0<p<1. Find the expected length of the initial run (this is a run of heads if the first toss gives head and of tails otherwise). [4]

53) Let B_1 , ... , $B_n$will be mutually disjoint, and B=$\bigcup_{i=1}^{n} Bi$. suppose P(B)>0 and P(A|B)=p_i, i=1,2,.....,n. show that P(A|B)=p. [3]

STATISTICS-HONOURS (2014)

1. Discuss about the different sources of error in Census and registration data.
(5)

2. Discuss the limitations of **Crude Birth Rate** as a measure of fertility in order to compare fertility situations of two different places. Also state a good measure which can be used for comparison.
(5)

3. Distinguish among **Infant Fertility Rate**, **Neonatal Mortality Rate** and **Perinatal Mortality Rate.**
(5)

4. Discuss the situation under which the Logistic curve can be used for **population forecasting**. What are the important properties of Logistic curve?
(5)

5. a) What is an **abridged life table**?

b) Derive the relations between different functions of a complete life table.

c) Given a complete life table with

$l_x = (100-x)/190,\ 5 \leq x \leq 100$

Find the probability that a man of age 30 will live up to age 80.
(4+6+5)

6. a) What is population growth?

b) Define reproduction rates and explain how far these rates may be looked upon as indices of population growth.

c) What is meant by saying **NRR for a country is 1.5**? Show that for any community **NRR < GRR**.

d) Explain the nature of growth when NRR=1.
(2+7+4+2)

STATISTICS-HONOURS-PRACTICAL (2014)

Age-groups (in years) of mothers. Female population (‘000) and number of live-births (’00) to the mothers are given for a small country with total population 2286×10^3.

Age-group(in year)	**Female population(in ‘000)**	**Number of live births(in ’00)**
15-19	85	23.4
20-24	70	145.4
25-29	73	167.4
30-34	76	102.2
35-39	75	51.3
40-44	72	14.2
45-49	67	1.0

Obtain:

(a) The crude death rate,

(b) The general fertility rate and

(c) The total fertility rate of the country.

If the crude rate of natural increase is 10 (per thousand), what is the Crude Death Rate for the country?

(6)

STATISTICS-HONOURS (2013)

1. Distinguish between rates and ratios. Define **Infant Mortality Rate** and justify why it is not a proper rate.

(5)

2. Distinguish between cohort and current life tables. Establish the relation among the different functions of a complete life table.

(5)

3. Distinguish among population estimation, projection and forecasting. Describe **AP** and **GP** methods of population estimation.

(5)

4. Examine the merits and demerits of **Net Reproduction Rate** (**NRR**) as a measure of population growth. Explain the situation where NRR<1 and NRR>1.

(5)

5. (a) Starting from a suitable assumption regarding the relative growth rate of population, derive the logistic equation.

(b) What are the important properties of the logistic curve?

(c) Describe the Rhode's method of fitting a logistic curve to the population data. (4+4+7)

6. (a) What is a complete life table?

(b) How does an abridged life table differ from a complete life table?

(c) Describe the uses of a life table.

(d) Why is the q_x column of a complete life table called the "**pivotal column**" of the table?

Show that under suitable assumption (to be started) q_x may be estimated from the observed age -specific death rates m_x (without the multiplier 1000) by the approximate formula:

$$q_x = 2m_x \div (2 + m_x)$$

Is the assumption underlying the above formula suitable for q_x at x= 0? If not suggest an alternative formula for q_0. (2+4+3+6)

STATISTICS-HONOURS-PRACTICAL (2013)

Complete the following life table:

Age (x)	l_x	d_x	$1000q_x$	L_x	T_x	e_x^0
25	78046					39.6
26	77614	440				
27						
28	76723		6.06			
29						
30	75523				2750943	

Hence, determine the probability that a person of age 25 lbd will die before reaching age 30 lbd.

(4+2)

STATISTICS-HONOURS (2012)

1. Describe few sources of raw data in demography. Distinguish between errors of coverage and errors of response in connection with errors in census and registration data.

(5)

2. Define crude rate of natural increase and vital index as measures of population growth .Explain why they are unsuitable to serve the purpose. What other measures are available to serve the purpose?

(5)

3. Define CBR, GFR and ASFR, and indicate why each is considered an improvement on the preceding measures of fertility.

(5)

4. What is meant by life table stationary population? Show that for a life table stationary population, the crude death rate(CDR) IS $1/e^0$(except for the multiplier 1000)

(5)

5. (a) Distinguish between population estimates and population projections. Describe the inter-censal and post-censal estimates by mathematical method (assume both linear and exponential growths of the population).

(b) With appropriate assumptions, state the logistic curve equation for population growth and examine the important properties of this curve.

(c) Briefly describe the component method of population projection.

STATISTICS-HONOURS-PRACTICAL (2012)

1. A part of a life table is given here with most of the entries missing. On the basis of the available figures, supply the missing ones:-

Age(x)	l_x	d_x	$1000q_x$	L_X	T_X	e_x^0
10	93102		0.62			
11			0.66			
12			0.72			
13			0.80			
14			0.90			
15			1.00			
16			1.12			
17			1.23			
18			1.33			
19			1.40		48,42,446	

Hence determine the probability:

(a) That a child of age 10 will live at least 5 years more,

(b) That two children aged 10 and 11 will each live at least 5 years more, and

(c) That of two children aged 10 and 11 at least one will die within nine years.

(6+3)

STATISTICS-HONOURS (2011)

1. Distinguish between rate and ratio with suitable examples. How does crude rate differ from specific rate? Explain with suitable examples.

(5)

2. Define GRR and NRR. Interpret the statement "NRR of the country *X* is 1.211". Show that for any community NRR<GRR.

(5)

3. What do you mean by standardization of death rate? Explain both the direct and indirect methods of standardization in this context.

(5)

4. Explain the role of TFR and IMR in the context of development.

(5)

STATISTICS-HONOURS-PRACTICAL (2011)

1 .The following information is obtained from a rural sample survey:

Age-group	No. of women	Proportion married	No. of births	Proportion surviving from birth to midpoint of age-group among married women.
15-19	16952	0.82	2642	0.902
20-24	19139	0.83	4270	0.891
25-29	10858	0.84	2181	0.878
30-34	4993	0.85	796	0.865
35-39	2460	0.74	205	0.849
40-44	925	0.64	48	0.430

(a) Calculate the GFR, TFR, GRR and NRR assuming (i) there is no legitimated birth and (ii) ratio of female to total birth is 0.485. Comment on your results.

(b) If the number of deaths among the 0-year age group is 400 and the total number of births is 0.05, find the Infant Mortality Rate.

(4)

STATISTICS-HONOURS (2010)

GROUP A

1. Name two statistical offices of government of India and mention two major publications of each.

2. What are price relatives? Examine the statement:

"If n the average of price relatives we ignore their weights, we shall not get an unweight or simple price index but an inappropriate weighted price index."

3. (a) Discuss the different functions of Central Statistical Organization (CSO) in brief.

(b) Define national income and explain two methods for its estimation.

(c) Indicate two uses of a cost of living index.

4. (a) If L_p, P_p, $L_q$denote, respectively, Laspeyres' price index, Paasche's price index and Laspeyres' quantity index, show that L_q (P_p - L_p) may be looked up on as a weighted covariance between price relatives and quantitative relatives. How will you interpret the result?

(b)Examine one index number in each of the following three cases:

The index satisfies neither the time reversal test nor factor reversal test.

The index satisfies the first but not the second.

The index satisfies both.

(a) Distinguish between a fixed base index and chain index using Laspeyres' price index number.

GROUP B

1. Define Infant Mortality Rate (IMR). Do you consider IMR as a true probability rate in the true sense of the term?

(5)

2. What do mean by a life-table stationary population? Show that the crude death rate (CDR) for a life-table stationary population, except for the multiplier 1000 equals to $\mathbf{1/e_0^0}$.

(5)

3. Distinguish between **Total Fertility Rate** and **Gross Reproduction Rate**. For a country the sex-ratio at birth maybe taken as 105 males to 100 females. If for a given year the TFR is 3500 per 1000, estimate the GRR.

(5)

4. A population of size 5×10^6 close to migration has fewer births than deaths. If it decreases at an annual rate of **0.75%**, how many years will it take to reach half of its present size?

(5)

5. (a) Explain why Crude Death Rate (CDR) is not always a suitable index to compare the mortality situations in two countries. In this regard describe

the direct and indirect methods of **standardization of death rates**. (6)

(b) Distinguish between the Morbidity Incidence Rate (**MIR**) and the Morbidity Prevalence Rate (**MPR**) in a community. (4)

6. (a) Why is the q_x column of a complete life-table called the **pivotal column** of the table? Show that under suitable assumption (to be stated) q_x maybe estimated for the

observed age-specific death rates m_x (without the multiplier 1000) by the approximate formula:

$$q_x = 2m_x \div (2+m_x)$$

Is the assumption underlying the above formula suitable for q_x at x=0? If not, suggest an alternative formula for q_0.

(6)

(b) Distinguish between a stationary population and a stable population. (4)

7. (a) Distinguish between population estimates and population projection.

(4)

(b) Briefly describe the component method of population projection.

(6)

8. (a) Define the Net Reproduction Rate (NRR) for a country and interpret the case: NRR= 1.5. Also comment on the use of this figure as an index of population growth.

(6)

(b) State he equation of logistic curve of population growth and examine the important properties of the curve.

(4)

STATISTICS HONOURS PRACTICAL (2010)

1. The following table shows the quantities consumed and the values (price × quantity) of 5 commodities for 5 successive years.

Commodities	**2005**		**2006**		**2007**	
	quantity	**value**	**quantity**	**value**	**quantity**	**value**
I	50	350	60	420	70	490
II	120	600	140	700	160	800
III	30	330	20	200	15	225
IV	20	360	15	300	10	220
V	5	40	5	420	5	60

Calculate the price index number for 2007 taking 2005 as the base period adopting chain base formula and Paaches' formula at each stage. Also verify whether the circular test is satisfied by the Paaches' formula or not on the basis of the above data.

(10)

2. The population of a certain country, as recorded in each of the ten decennial censuses s shown below :

Year	Population (in million)
1911	238.3
1921	252.0
1931	251.2
1941	278.9
1951	318.5
1961	361.0
1971	439.1
1981	547.0
1991	683.8
2001	823.8

Fit a logistic curve to the data and comment on the nature of fit.

(10)

STATISTICS-HONOURS (2009)

1. What are the different sources of official statistics in India? Mention the name of two official publications containing information about industrial statistics.

2. What are the different uses of index number?

3. (a) Discuss the activities of National Sample Survey Organization (NSSO) in brief.

(b) Discuss the different steps in construction of a wholesale price index number.

(c) What is a chain index number?

4. (a) What is a value index?

(b) Derive Fisher's index number starting from Laspeyer's and Paaches' index numbers.

(c) What are the different tests for consistency that a price index number should satisfy?

(d) Why index is Fisher's number called an ideal index?

5. What do you mean by a vital event? Distinguish between rates and ratios of vital events.

6. Distinguish between Neonatal and Perinatal mortality rates.

7. Compare AP and GP methods for population estimation.

8. Suppose a population of size 1 lakh growing exponentially at annual rate r=0.02 . If r is compounded annually, determine the size of the population at the end of 3 years.

9. (a) Discuss the different sources of errors in census and registration data.

(b) What is age-specific death rate? How can one use such rates to compare the mortality situations of two states in India?

10.(a) What is a complete life-table? Describe how one can construct such a table from population and death statistics.

(b) Prove the following inequality:

$q_x < m_x < q_x \div (1 - q_x)$ (symbols have their usual significance)

11.(a) Define infant mortality rate. How can one compute it from available data?

(b) Discuss the different measures of population growth.

(c) Show that for any community NRR is necessarily less than the GRR

(3+5+2)

12.(a) Discuss the problem of comparing mortality situation of two different countries and describe a suitable measure for this situation.

(b) A population growing exponentially stood at 1 lakh in 1960 and 3 lakh in 2000,

(i) What is the rate of increase?

(ii) What is its doubling time?

(6+4)

STATISTICS-HONOURS-PRACTICAL (2009)

Consider the following data set for two countries.***Country 1***

Age group	***Population size***	***Number of males***	***Number of deaths***	***Number of deaths in males***
<1	***100***	***50***	***750***	***400***
1-4	***400***	***200***	***3000***	***1600***
5-20	***1500***	***750***	***12000***	***6500***
21-100	***4000***	***1500***	***31000***	***12000***

Country 2

Age group	***Population size***	***Number of males***	***Number of deaths***	***Number of deaths in males***
<1		7.5	12	6
1-4	60	30	45	24
5-20	200	15	160	
21-100	440	170	350	200

(i) Calculate the crude death rates for both the countries.

(ii) Compute the age-specific death rate of both the countries separately for Males, Females and the Total population.

(iii) In order to compare the mortality situations of the two countries propose a good measure and calculate its value for both the countries.

(10)

STATISTICS HONOURS (2008)

GROUP A

1) Briefly describe the nature of data usually collected through Indian Censuses with particular reference to 1991 Census.

2) Describe formula error, sampling error and homogeneity error in the context of index number.

3) (a) Discuss the activity of Central Statistical Organization (CSO) in brief.

(b) Identify the merits and demerits of Census and Sample survey.

(c) What do you mean by National income statistics?

4. Discuss how one can construct an Industrial Production for the industries of West Bengal.

GROUP B

5. Discuss how one can estimate the population of India for the year 2001 on the basis of 1991 Census and past registration data? What are the possible sources of errors in such an estimate?

6. Distinguish between population estimation projection and forecasting. Mention a method or population forecasting.

7. "The death in India due to the cause of HIV infection is more than that of United States"-How do you verify this statement with a proper measure?

8. (a) What do you mean by growth rate of a population?

(b) Explain the nature of growth of a country f which NRR is .2

(c)Show that for any country GRR is always greater than NRR

(d) Describe the AP method of population estimation.

(2+2+3+3)

9. Assume a constant force of mortality a given age interval,

$$\mu_{x+t}=\mu$$

Then prove that,

(i) $q_x= 1 - e^{-\mu t}$ $\quad 0 \leq t \leq 1$

(ii) $q_{x+t}=1- e^{-\mu(1-t)}$ $\quad 0 \leq t \leq 1$

(iii)

10. Consider a population containing two subgroups each initially of size 100,000. If the two subgroups grow exponentially at annual rates 0.2and 0.04 respectively, determine the total population after 5 years. What is the average rate of increase at this stage?

(5)

STATISTICS-HONOURS-PRACTICAL (2008)

From the following data relating to a particular community compute the Gross Reproduction Rate and the Net Reproduction Rate. Interpret your result.

Age of mother	**Number of women**	**Number of birth**	**Survival factor**
15-19	9000	140	0.920
20-24	9200	1312	0.914
25-29	8900	1067	0.918
30-34	8600	771	0.891
35-39	8400	468	0.878
40-44	8500	160	0.869

Assume that **48.7%** of the total are female births. Survival factor gives the rate of survival from birth to the mid-point of the corresponding age-group.

(10)

STATISTICS-HONOURS (2007)

1. What is expectation of life? Explain with reason how it varies from age.

(5)

2. Show that TFR/1000 > GRR > NRR.

(5)

3. Show that Crude Death Rate for a life table stationary population except for a multiplier 1000, equals $[e_0^0]^{-1}$.

(5)

4. On a life table with,

$l_x = (100 - x) / 190 \qquad 5 \leq x \leq 100$

Work out:

(i) The chance that a child who has reached the age 5 will up to age 60.

(ii) The chance that a man of age 30 will live up to age 80.

(iii)The probability of dying within 5 years for a man aged 40.

(6)

STATISTICS-HONOURS-PRACTICAL (2007)

1. The following table the average per capita monthly consumption of cereals and prices of cereals in rural areas of a given region for four different periods I,II,III,IV

Commodities	Consumption (in Kg.)				Price (Rs./Kg)			
	I	II	III	IV	I	II	III	IV
Rice	4.70	3.95	3.60	4.30	5.40	5.75	6.25	6.30
Wheat	5.50	6.75	6.40	6.87	4.95	5.20	5.00	5.10
Others	7.66	8.20	7.85	7.60	3.30	2.95	3.25	3.60

Calculate price index of cereals for the period IV taking period I as base adopting chain base formula and using Laspeyers' index number at each stage. If you find after calculation that the consumption figure in period IV are in error by 2%, do you think that your index is going to be affected by this? Give reason for your answer.

(10)

2. For a stationary population with cohort $\mathbf{l_o=100,000}$ out of the children born in 1990, the number of children decreased in 1990 was 20,000 and the number of children decreased in 1991 was 5,000. Given below are the age-specific mortality rates of the given population for the following ages *x* (last birthday):

X :	0	1	2	3	4
m_x :	0.4075	0.0123	0.0024	0.0018	0.0013

(b) Compute the complete expectation of life at age 4, given the same at birth is **65.90** years.

(c) What is the chance that two newborn babies would survive 4 years after their birth?

(10)

STATISTICS HONOURS (2006)

1. Distinguish between GDP and GNP.

(5)

2^{nd} year

Moment-Generating Function and Some Particular Probability Distribution

1) A. **Select the correct option:**

(i) A binomial distribution with parameters *n* and *p* is positively skew if

(a) $p>\frac{1}{2}$ (b) $p=\frac{1}{2}$ (c) $p<\frac{1}{2}$ (d) none of these

(ii) Poisson distribution is

(a) leptokurtic (b) platykurtic (c) mesokurtic (d) none of these

(iii) If a poisson distribution has double mode at X = 3 and 4, then coefficient of variation of the distribution is

(a) 100% (b) 200% (c) 50% (d) 25%

(iv) The variance of number of successes in a set of 20 Bernoullian trials cannot exceed

(a) 10 (b) 5 (c) 0.25 (d) 0.0125

(v) A binomial distribution with parameters 5 and $\frac{1}{3}$ is

(a) leptokurtic (b) platykurtic (c) mesokurtic (d) none of these

(vi) If X follows R(α, β) distribution, then E(X)

(a) $\frac{\alpha+\beta}{2}$ (b) $\frac{\alpha-\beta}{2}$ (c) $\frac{\beta-\alpha}{2}$ (d) $\frac{(\beta-\alpha)^2}{12}$

(vii) Mean deviation about mean of a N(0, 1) distribution is

(a) $\sqrt{\frac{\pi}{2}}$ (b) $\sqrt{2\pi}$ (c) $\frac{1}{\sqrt{2\pi}}$ (d) $\sqrt{\frac{2}{\pi}}$

(viii) The median of the standard exponential distribution is

(a) ln 2 (b) 1 + ln 2 (c) $\frac{1}{3}$ ln 2 (d) none of these

B. **Fill in the blanks:**

(i) If X follows binomial distribution with n = 70 and p= $\frac{1}{4}$, then P(X = r) is maximum when r = ____________.

(ii) Bernoullian trials are ____________.

(iii) In a throw of a fair die, the number of points on its upper face follows ____________.

(iv) If $f(x)$ is the p. m. f. of a discrete random variable X, then $\sum_x f(x)$ = ____________.

(v) A normal variable with mean 0 and s. d. 1 is called a normal ____________.

(vi) The points of inflection of a normal distribution with mean 55 and variance 25 are ____________ and ____________.

C. **Write 'true' of 'false':**

(i) In a binomial distribution, mean and standard deviation may be 5 and 2.

(ii) A binomial distribution with mean 9 is binomial.

(iii) In a Poisson distribution $\mu_2 = \mu_3$.

(iv) The p. d. f. of standard normal variable t is $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}$, $-\infty < t < \infty$.

(v) The coefficient of variation of standard exponential distribution is 100.

(vi) $0 \leq$ F(x) ≤ 1, where F(x) is the distribution function of a random variable X.

(vii) Expectation and standard deviation of a Poisson distribution can never be equal.

(viii)If mean, median and mode of a probability distribution are equal, the distribution must be normal.

D. **Answer the following questions:**

(i) The mean of a symmetrical binomial distribution is 5. What is its variance?

(ii) If X and Y follow a bivariate normal distribution with parameters 0, 0, 1, 1;ρ ; write down E(Y | X = x).

(iii) State for what value of p the hypergeometric distribution is symmetric.

(iv) Write the p. m. f. of the negative binomial distribution. When will this distribution be the geometric distribution.

(v) Write down the p. d. f. of bivariate normal (0, 0, 1, 1, ρ) distribution.

2) A random variable X follows binomial distribution with parameters n and p. prove that

(i) P(X is odd) = $\frac{1}{2} 1-(q-p)^n$

(ii) Cov(X, X - n) = Var(X) = npq, where $q = 1 - p$.

3) The probability that a Poisson variate X takes a positive value is 1- e^{-2}. Find the mean and mode of X and P(-1 < X < 1.6).

4) If a random variable X follows binomial distribution with mean 4 and E(X^2)=19, find P(X = at most 2) and P(X = at least 2).

5) Show that binomial distribution with parameters 3 and 0.5 is symmetric.

6) The chance of a person hitting a target is 1/3. How many times must he fire so that the probability of hitting the target at least once is more than 0.9?

7) If a random variable X follows Poisson distribution satisfying P(X = 1) = P(X = 2), determine the value of P(X > 1 | X < 3), expectation and standard deviation of X.

8) If a random variable X follows Poisson distribution so that E(X^2)= 30, find P(X assumes a non-zero value), P(X > 2) and P(X = at most 1).

9) Suppose 5% of the inhabitants of a city are consumers of tea. Find the probability that a sample of 100 inhabitants will contain at least 2 consumers of tea. (Given e^{-5}= 0.007)

10) For a binomially distributed random variable X with parameters n and p, what will be maximum variance and for what of p this maximum value will be attained?

11) For what value of k,

f(x) = k $\binom{n}{x} p^x q^{n-x}$, $x = 1, 2, \dots, n$; $p + q = 1$, $0<p<1$

= 0, elsewhere,

Is the probability mass function of a random variable X? Find P(X $\neq$ 1).

12) Determine *f(x)*, the probability mass function from *f(x)*= $\frac{\lambda}{x} f(x-1)$, x = 1, 2, …
where *f(x)* is non-zero for non-negative integral values of the random variable X.

13) Show that for the exponential distribution defined by that p. d. f.

$f(x) = \theta . e^{-\theta x}$, $x \geq 0$

= 0, elsewhere,

where $\theta > 0$, P(X > S + t | X>t) = P(X > S).

14) If X and Y are two independent normal variables with mean 5 and 7 and standard deviations 3 and 4 respectively, then find the mean and standard deviation of 3X – 2Y.

15) In a normal distribution, 465 of the terms are over 40 and 90% are under 75. Find the mean and standard deviation of the distribution. (Given that $\int_{-\infty}^{x} (2\pi)^{-1/2} e^{-t^2/2} dt$ = 0.54 or 0.90, according as x = 0.10 or 1.28).

16) See whether the following function can be accepted as a probability density function:

$f(x) = \frac{5}{\sqrt{\pi}} e^{-25x^2}$, $-\infty < x < \infty$.

If so, write down the important properties of this distribution.

17) What are the points of inflection of the normal distribution whose probability density function is

$f(x) = \frac{1}{\sqrt{18\pi}} e^{-\frac{(x-7)^2}{18}}$, $-\infty < x < \infty$.

18) TheBurdwan Municipality Corporation installs 10,000 electric lamps in the streets of Burdwan. If these lamps have an average life of 1850 burning hours with a standard deviation of 200 hours, what number of lamps may be expected to burn for (i) more than 2000 hours, and (ii) less than 1600 hours?
(Given that Φ(0.75) = 0.7734, Φ(1.25) = 0.8944)

19) If the random variables X and Y follow a bivariate normal distribution with parameters $\mu_X, \mu_Y, \sigma_X, \sigma_Y$ and ρ_{XY}, write down the conditional distribution of Y given X.

20) Show that the Poisson distribution can be obtained as a limiting form of a binomial distribution.

21) If a normal distribution is symmetrical about 5 with standard deviation 2, what are the points of inflexion of the distribution?

22) If a Poisson random variable X has two modes at X = 2 and X = 3, find the mean of X.

23) Suppose X is a normally distributed random variable with mean 10 and variance 9. Find $P(X < 7) + P(10 < X < 13)$.

24) Suppose the coefficient of variation of a binomial random variable X with parameters n and p is 100%. If $n = 3$, find the variance of X.

25) What is the variance of a symmetrical binomial distribution with $n = 32$?

26) Find the mean deviation about mean of X where X follows $N(0, \sigma^2)$.

27) A random variable X has the p. d. f.

$$f(x) = \begin{cases} x, & \text{for } 0 \le x \le 1 \\ k - x, & \text{for } 1 < x \le 2 \\ 0, & \text{for } x > 2 \end{cases}$$

Find the constant k.

28) If X follows B in (n, p), then find out $\text{Cov}\left(\frac{X}{n}, \frac{n-X}{n}\right)$.

29) Find the mean and variance of the random variable X with p. d. f.

$$f(x) = \frac{1}{B(2,3)} \cdot \frac{x}{(1+x)^5}, \quad 0 \le x < \infty.$$

2nd year

Moment-Generating Function and Some Particular Probability Distribution

1) A. **Select the correct option:**

(i) A binomial distribution with parameters n and p is positively skew if

(a) $p > \frac{1}{2}$ (b) $p = \frac{1}{2}$ (c) $p < \frac{1}{2}$ (d) none of these

(ii) Poisson distribution is

(a) leptokurtic (b) platykurtic (c) mesokurtic (d) none of these

(iii) If a poisson distribution has double mode at X = 3 and 4, then coefficient of variation of the distribution is

(a) 100% (b) 200% (c) 50% (d) 25%

(iv) The variance of number of successes in a set of 20 Bernoullian trials cannot exceed

(a) 10 (b) 5 (c) 0.25 (d) 0.0125

(v) A binomial distribution with parameters 5 and $\frac{1}{3}$ is

(a) leptokurtic (b) platykurtic (c) mesokurtic (d) none of these

(vi) If X follows R(α, β) distribution, then E(X)

(a) $\frac{\alpha+\beta}{2}$ (b) $\frac{\alpha-\beta}{2}$ (c) $\frac{\beta-\alpha}{2}$ (d) $\frac{(\beta-\alpha)^2}{12}$

(vii) Mean deviation about mean of a N(0, 1) distribution is

(a) $\sqrt{\frac{\pi}{2}}$ (b) $\sqrt{2\pi}$ (c) $\frac{1}{\sqrt{2\pi}}$ (d) $\sqrt{\frac{2}{\pi}}$

(viii) The median of the standard exponential distribution is

(a) ln 2 (b) 1 + ln 2 (c) $\frac{1}{3}$ ln 2 (d) none of these

B. **Fill in the blanks:**

(i) If X follows binomial distribution with n = 70 and $p=\frac{1}{4}$, then P(X = r) is maximum when r = ____________.

(ii) Bernoullian trials are ____________.

(iii) In a throw of a fair die, the number of points on its upper face follows ____________.

(iv) If $f(x)$ is the p. m. f. of a discrete random variable X, then $\sum_x f(x)$ = ____________.

(v) A normal variable with mean 0 and s. d. 1 is called a normal ____________.

(vi) The points of inflection of a normal distribution with mean 55 and variance 25 are ____________ and ____________.

C. **Write 'true' of 'false':**

(i) In a binomial distribution, mean and standard deviation may be 5 and 2.

(ii) A binomial distribution with mean 9 is binomial.

(iii) In a Poisson distribution $\mu_2 = \mu_3$.

(iv) The p. d. f. of standard normal variable t is $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}$, $-\infty < t < \infty$.

(v) The coefficient of variation of standard exponential distribution is 100.

(vi) $0 \leq F(x) \leq 1$, where F(x) is the distribution function of a random variable X.

(vii) Expectation and standard deviation of a Poisson distribution can never be equal.

(viii)If mean, median and mode of a probability distribution are equal, the distribution must be normal.

D. **Answer the following questions:**

(i) The mean of a symmetrical binomial distribution is 5. What is its variance?

(ii) If X and Y follow a bivariate normal distribution with parameters 0, 0, 1, 1; ρ ; write down $E(Y \mid X = x)$.

(iii) State for what value of p the hypergeometric distribution is symmetric.

(iv) Write the p. m. f. of the negative binomial distribution. When will this distribution be the geometric distribution.

(v) Write down the p. d. f. of bivariate normal (0, 0, 1, 1, ρ) distribution.

2) A random variable X follows binomial distribution with parameters n and p. prove that

(i) P(X is odd) = $\frac{1}{2} 1 - (q-p)^n$

(ii) Cov(X, X - n) = Var(X) = npq, where $q = 1 - p$.

3) The probability that a Poisson variate X takes a positive value is 1- e^{-2}. Find the mean and mode of X and P(-1 < X < 1.6).

4) If a random variable X follows binomial distribution with mean 4 and E(X^2)=19, find P(X = at most 2) and P(X = at least 2).

5) Show that binomial distribution with parameters 3 and 0.5 is symmetric.

6) The chance of a person hitting a target is 1/3. How many times must he fire so that the probability of hitting the target at least once is more than 0.9?

7) If a random variable X follows Poisson distribution satisfying P(X = 1) = P(X = 2), determine the value of P(X > 1 | X < 3), expectation and standard deviation of X.

8) If a random variable X follows Poisson distribution so that E(X^2)= 30, find P(X assumes a non-zero value), P(X > 2) and P(X = at most 1).

9) Suppose 5% of the inhabitants of a city are consumers of tea. Find the probability that a sample of 100 inhabitants will contain at least 2 consumers of tea.
(Given e^{-5}= 0.007)

10) For a binomially distributed random variable X with parameters n and p, what will be maximum variance and for what of p this maximum value will be attained?

11) For what value of k,

f(x) = k $\binom{n}{x} p^x q^{n-x}$, $x = 1, 2, \ldots, n$; $p + q = 1$, $0<p<1$

= 0, elsewhere,

Is the probability mass function of a random variable X? Find P(X $\neq$ 1).

12) Determine $f(x)$, the probability mass function from $f(x)= \frac{\lambda}{x} f(x-1)$, $x = 1, 2, \ldots$ where $f(x)$ is non-zero for non-negative integral values of the random variable X.

13) Show that for the exponential distribution defined by that p. d. f.

$f(x) = \theta . e^{-\theta x}$, $x \geq 0$

= 0, elsewhere,

where $\theta > 0$, P(X > S + t | X>t) = P(X > S).

14)If X and Y are two independent normal variables with mean 5 and 7 and standard deviations 3 and 4 respectively, then find the mean and standard deviation of 3X – 2Y.

15) In a normal distribution, 465 of the terms are over 40 and 90% are under 75. Find the mean and standard deviation of the distribution. (Given that $\int_{-\infty}^{x}(2\pi)^{-1/2}e^{-t^2/2}dt$ = 0.54 or 0.90, according as x = 0.10 or 1.28).

16) See whether the following function can be accepted as a probability density function:

$f(x) = \frac{5}{\sqrt{\pi}}e^{-25x^2}$, $-\infty < x < \infty$.

If so, write down the important properties of this distribution.

17) What are the points of inflection of the normal distribution whose probability density function is

$f(x) = \frac{1}{\sqrt{18\pi}}e^{-\frac{(x-7)^2}{18}}$, $-\infty < x < \infty$.

18)TheBurdwan Municipality Corporation installs 10,000 electric lamps in the streets of Burdwan. If these lamps have an average life of 1850 burning hours with a standard deviation of 200 hours, what number of lamps may be expected to burn for (i) more than 2000 hours, and (ii) less than 1600 hours?
(Given that $\Phi(0.75) = 0.7734$, $\Phi(1.25) = 0.8944$)

19) If the random variables X and Y follow a bivariate normal distribution with parameters $\mu_X, \mu_Y, \sigma_X, \sigma_Y$ and ρ_{XY}, write down the conditional distribution of Y given X.

20) Show that the Poisson distribution can be obtained as a limiting form of a binomial distribution.

21) If a normal distribution is symmetrical about 5 with standard deviation 2, what are the points of inflexion of the distribution?

22) If a Poisson random variable X has two modes at X = 2 and X = 3, find the mean of X.

23) Suppose X is a normally distributed random variable with mean 10 and variance 9. Find P(X < 7) + P(10 < X < 13).

24) Suppose the coefficient of variation of a binomial random variable X with parameters n and p is 100%. If n = 3, find the variance of X.

25) What is the variance of a symmetrical binomial distribution with n = 32?

26) Find the mean deviation about mean of X where X follows N (0, σ^2).

27) A random variable X has the p. d. f.

$$f(x) = \begin{cases} x, & \text{for } 0 \le x \le 1 \\ k - x, & \text{for } 1 < x \le 2 \\ 0, & \text{for } x > 2 \end{cases}$$

Find the constant k.

28) If X follows B in (n, p), then find out Cov$\left(\frac{X}{n}, \frac{n-X}{n}\right)$.

29) Find the mean and variance of the random variable X with p. d. f.

$f(x) = \frac{1}{B(2,3)} \cdot \frac{x}{(1+x)^5}$, $0 \le x < \infty$.

PROBLEMS ON INDEX NUMBER

1. A. CHOOSE THE CORRECT OPTION:

i) Fisher’s ideal index is

a) A.M. of Laspeyres’ and Paasche’s index

b) Median of Laspeyres’ and Paasche’s index

c) G.M. of Laspeyres’ and Paasche’s index

d) none of these

ii) Factor reversal test is satisfied by

a) laspeyres’ index

b) Paasche’s index

c) both a and b

d) Fisher’s index

iii) Fixed-base index and chain base index are equal when the formula satisfies

a) Circular test

b) time reversal test

c) factor reversal test

d) none of these

iv)Paasche’s index is based on

a) current year quantities

b) base year quantities

c) average of current and base year quantities

d) none of these

v) If the consumer price index for middle-class people of Kolkata in 1992 with 1985 as base period is 225, then the retail price have increased,on the average.

a) 225%

b) 125%

c) 25%

d) none of these

vi) Constructing one continuous series from two different index number series with a common base is

a) deflating

b) base shifting

c) splicing

d) none of these

vii) Time reversal is satisfied by

a) Laspeyres' index

b) simple G.M. of price relatives

c) Marshall-Edgeworth index

d) both b and c

viii) The quotient $\sum p1q1/\sum p0q0$ is called

a) price index

b) value index

c) quantity index

d) none of these

ix) If the prices of all items change in the same ratio, then

a) Laspeyres’ index = Paasche’s index

b) Laspeyres’ index <Paasche’s index

c) laspeyres’ index >Paasche’s index

d) none of these

x) If Laspeyres’, Paasche’s and F isher’s ideal index are respectively denoted by L,P,F ,then

a) P lies between L and F

b) F lies between L and P

c) L lies between F and P

d) none of these

xi) Weighted H.M. of price relatives using current year value as weights is

a) Paasche’s index

b) Fisher’s index

c) Laspeyres’ index

d) Bowley’s index

xii) The best average in the construction of index number is

a) arithmetic mean

b) geometric mean

c) median

d) mode

B. WRITE TRUE OR FALSE:

i) Marshall-Edgeworth price index number lies between Laspeyres' and Paasche's price index numbers.

ii) Factor reversal test is satisfied only by Fisher's index number formula.

iii) The index for the base period is arbitrarily taken as 100.

iv) Chain-base index numbers are easier to calculate than fixed-base index number.

v) Laspeyres' price index is based on current year quantities.

vi) Weighted A.M. of price relatives with base-period values as weights give Laspeyres' formula.

C. FILL IN THE BLANKS:

i) Factor reversal test states, Price index * Quantity index = ________ index

ii) Link indices are successively multiplied to obtain ________ indices.

iii) Fisher's index is called ________ index.

iv) Purchasing power of money is ________ proportional to the price index.

v) Cost of living index is also known as ____________ price index.

vi) Bowley's index is the ________ mean of Laspeyres' and Paasche's formula.

2. What is an index number? What are the uses of index number?

3. Discuss the problems in the construction of index numbers?

4. Show that Laspeyres' and Paasche's price index numbers can be obtained by the method of averaging price relatives as well as by aggregative method.

5. What are the tests proposed by Fisher for checking the goodness of an index number? Do the Laspeyres' and Paasche's price index numbers satisfy these tests? Mention an index number that satisfies them.

6. Show that Marshall-Edgeworth price index number is weighted arithmetic mean of price relatives.

7. a) Prove that Laspeyres' and Paasche's index numbers will be equal if the prices of all goocs change in the same ratio.

b) Write Paasche's quantity index formula.

8. What are errors in the construction of index number?

9. Discuss the relative merits and demerits of fixed-base and chain-base methods of constructing index number.

10. What is meant by cost of living index number? How is it constructed? Mention its uses.

11. Write note on :

i) Chain index number ii) Tests of index number iii) Wholesale price index number iv) Uses of price index number v) Base shifting

12. What is meant by i) splicing and ii) deflating of index number.

13. Using Paasche'sformula,find the price index number for the year 2000 with 1990 as base year.

COMMODITY	1990		2000	
	PRICE	QUANTITY	PRICE	QUANTITY
A	65	40	81	46
B	72	35	90	54
C	57	92	77	72

14.From the following data calculate price index numbers for 1986 with 1976 as base by i) Laspeyres' ii) Paasche's iii) Marshal-Edgeworth and iv) Fisher's formulae.

COMMODITY	UNIT	1976		1986	
		QUANTITY	MONEY VALUE	PRICE	MONEY VALUE
A	kg	10	40	5.50	66
B	litre	06	18	4.40	22

C	metre	05	75	18.20	91
D	kg	08	48	7.60	76

15. Calculate price index numbers for 1983 and 1984 with 1982 as base year using i) simple A.M. ii) weighted A.M. of price relatives.

COMMODITY	PRICE(Rs.)			WEIGHT
	1982	1983	1984	
A	160	200	240	15
B	90	120	108	25
C	60	150	120	20
D	50	120	80	40

16.Compute the index of food prices for January 1971 with 1961 = 100 on the basis of following data.

ITEM	PRICE PER KG		AVERAGE EXPENDITURE PER FAMILY PER MONTH IN 1961
	1961	1971	
RICE	0.75	1.40	40.2
WHEAT	0.46	0.98	15.8
SUGAR	1.25	1.80	7.2
MILK	1.20	1.85	19.4

MUSTARD OIL	3.85	6.20	14.2
FISH	3.55	6.50	30.8

17. If the ratio between Laspeyres' and Paasche's index number is 28:27, find the missing figure in the table:

COMMODITY	BASE YEAR		CURRENT YEAR	
	PRICE	QUANTITY	PRICE	QUANTITY
X	1	10	2	5
Y	1	5	-----	2

18. Calculate the price index number for 1985 with 1978 as base using a suitable formula on the basis of the following data.

COMMODITY	PRICE PER UNIT (in Rs.)		MONEY VALUE
	1978	1985	1978
A	4	8	40
B	6	9	42
C	5	11	60
D	3	6	24
E	2	4	32

19. Show the Fisher's "ideal" index number satisfies both the time reversal and the factor reversal tests, and verify this from the following data.

COMMODITY	1970		1972	
	PRICE	QUANTITY	PRICE	QUANTITY
A	6	50	10	56
B	2	100	2	120
C	4	60	6	60
D	10	30	12	24
E	8	40	12	36

20.In 1978 for working class people wheat was selling at an average price of Rs. 16 per 10 kg, cloth at Rs. 4 per metre, house-rent at Rs. 40 per house and other items at Rs. 10 per unit. By 1981 cost of wheat rose by 40 paisa per kg, house-rent by Rs. 20 per house, and other items doubled the price. The working class cost of living index for the year 1981 with 1978 as base was 160. By how much the cloth rose in price during the period 1978-81?

21. Group index numbers for 1991 with 1981 as base year and group weights of an average working class family's budget are given below:

GROUP	FOOD	FUEL &LIGHT	CLOTHING	RENT	MISCELLANEOUS
INDEX NO.	250	227	231	171	191
WEIGHT	48	10	14	12	16

Find the cost of living index number of 1991 with 1981 as base. A man was getting Rs. 500 in 1981 and Rs. 975 in 1991. Was he better or worse off in 1991 compared to 1981? If worse, how much extra allowance he ought to have received to maintain his 1981 standard of living?

22. In a working class budget enquiry in towns A and B, it was found that an average working class family's expenditure on food and other items are as follows:

	TOWN A	TOWN B
FOOD	64%	50%
OTHER ITEMS	36%	50%

In 1971 the consumer price index stood at 279 for town A and 265 for town B (Base year 1961 = 100). It was known that the rise in the prices of all articles consumed by the working class was the same for A and B. What were the indices of 1971 for food and other items?

23. The price relatives and weights of a set of commodities are given in the following table:

COMMODITY:	A	B	C	D
PRICE RELATIVE:	125	120	127	119
WEIGHT:	w1	2w1	w2	w2+3

If the sum of the weights is 40 and the index for the set is 122, find the numerical values of w1 and w2.

24. The following table gives the percentage increase in price for the different groups and the percentage of total expenditure for the middle-class people of Kolkata in 1981. Taking 1971 as base period, obtain a general cost of living index number for the class of people.

GROUP	FOOD	CLOTHING	FUEL&LIGHT	HOUSERENT	OTHER
PERCENTAGE INCREASE IN PRICE:	125	95	80	50	100
PERCENTAGE OF TOTAL EXPENDITURE:	10	15	30	20	25

25. The cost of living for a certain group of workers of a city has increased by 20% , while the rise in their wages is 50%, Find the percentage increase in their real wages.

26. A jute mill worker in Kolkata earns Rs. 750 per month. The cost of living index for January 1986 is given as 160. Using the following data find the following amount he spent on i) food ii) rent.

GROUP:	FOOD	CLOTHING	RENT	FUEL&LIGHT	MISCELLANEOUS
EXPENDITURE:	---	125	----	100	75
GROUP INDEX:	190	181	140	118	101

28. The following are index number of prices (1979 = 100)

YEAR:	1979	1980	1981	1982	1983	1984	1985	1986
INDEX NUMBER:	100	135	207	258	300	360	396	423

Shift the base from 1979 to 1983 and recast the data.

29. Monthly wages average in different years are given below:

YEAR:	1967	1968	1969	1970	1971	1972	1973
WAGES (Rs.):	200	240	350	360	360	380	400
PRICE INDEX:	100	150	200	220	230	250	250

Calculate the real wage index numbers.

30. Given below are two series of index numbers, one with 1971 as base and the other with 1976 as base:

YEAR:	1971	1972	1973	1974	1975	1976
(A) INDEX NO.:	100	122	126	122	127	132
YEAR:	1976	1977	1978	1979		
(B) INDEX NO.:	100	106	109	110		

Splice the index B to index A to prepare a continuous series with 1971 as base.

Inference

1. Show that the mean and the variance of S_n, the empirical distribution function, are

$$\Sigma[S_n(x)] = F(x),\ var[S_n(x)] = \frac{F(x)[1-F(x)]}{n}.$$

Hence show that $S_n(x)$ is a consistent estimator of $F(x)$.

2. Show that under the alternative $H{:}E(O_i) = np_i^{'}$, for the statistic defined in (3.),

$$E(\chi^2) = np_i^{'}(1-p_i^{'}) + (np_i^{'}-E_i)^2]/E_i.$$

Hence show that under H_0, i.e. when $np_i^{'} = E_i$, $E(\chi^2)=(k-1)$.

3. Show that the power function of the sign test with respect to the alternative $H{:}p>$ 0.5 is obtained by evaluating.

$$\sum_{r=r_\alpha}^{r+s}\binom{r+s}{r}p^r(1-p)^s.$$

for different values of p.

4. Show that the power of the sign test for $H_0 : {}_e = 45$ against $H : {}_e > 45$ for r+s = 18, under the assumption that the population is normal with mean 50 and standard deviation 5, and at the level of significance 0.05, is

$$\sum_{r=13}^{18}\binom{18}{r}(0.84)^r\,(0.16)^{18-r}.$$

5. Using the result of **Theorem 11.7**, show that the limiting distribution of $4n(D_n^{+})^2$ is the χ^2-distribution with $d.f.$=2.

6. Show that

$$D_n^- = \max\left\{ \max_{1\le r\le n} \left[F_0(X_{(r)}) - \frac{r-1}{n}\right], 0\right\}$$

and that the sampling distribution of D_n^- is identical with that of D_n^+.

7. For the Wilcoxon signed-ranked test, show that the null distribution of T^+ is symmetric about n(n+1)/4 and that T^+ and T^- are identically distributed.

8. (Test for randomness) Let $x_1 , x_2 , \ldots\ldots, x_n$ be the observed values of the n random variables $X_1 , X_2 , \ldots\ldots, X_n$. Suggest a test for the hypothesis that the random variables are IID.

9. Show that for the random variables $X_1, X_2, \ldots\ldots, X_n$ the signed-rank statistic for testing $H_0 :{}_e = 0$ is

$$T^+ - T^- = 2\sum_{1}^{n} i\, D_{(i)} - n(n+1)/2,$$

where $D_{(i)} = \begin{cases} 1 \text{ if } D_i > 0 \text{ for the } i\text{th smallest } |D_i| \\ 0 \text{ if } D_i < 0 \text{ for the } i\text{th smallest } |D_i| \\ \qquad i = 1,2,\ldots\ldots, n \end{cases}$

10. Let $r_{m,n}(u)$ be the number of distinguishable arrangements of mX and nY random variables such that the Mann-Whitney statistic $U=u$ under $H_0 : F_X(x) = F_Y(x)$, all x. Then show that $r_{m,n}(u) = r_{m,n}(mn - u)$ and hence that

$$P\left[\, U = \ + u\right] = P\left[U = - u\right],$$

i.e. prove that the null distribution of U is symmetrical about $mn/2$.

11. Deduce the following recurrence relation:

$$r_{m,n}(u) = r_{m,n-1}(u) + r_{m-1,n}(u-n).$$

Show that $(m+n)\, p_{m,n}(u) = np_{m,n-1}(u) + mp_{m-1,n}(u-n)$

where $p_{m,n}(u) = r_{m,n}(u) / \binom{m+n}{n}$, under H_0.

12. An alternative definition of Mann-Whitney statistic in case of ties is

$$U' = \sum_{i=1}^{m}\sum_{j=1}^{n} D'_{ij}, \text{ where } D'_{ij} = \begin{cases} 1 \text{ if } X_i > Y_j \\ 0 \text{ if } X_i = Y_j \\ -1 \text{ if } X_i < Y_j \end{cases}.$$

Then show that U' is an unbiased estimator of $mn\{P[Y<X] - P[Y>X]\}$.

13. Define the sign function as

$$\text{sgn}(u) = \begin{cases} -1 \text{ if } u < 0 \\ 0 \text{ if } u = 0 \\ 1 \text{ if } u > 0. \end{cases}$$

Then show that U' defined in **(12.)** is

$$U' = \sum_{i=1}^{m}\sum_{j=1}^{n} \text{sgn}\,(X_i - Y_j).$$

14. Show that the two-sample Wilcoxon and Mann-Whitney statistics are related as follows :

$$U = W - m(m+1)/2.$$

if there are no ties between X and Y observations.

15. Let $r_{m,n}(K)$ denote the number of arrangements of $mX$and nYrandom variables such that the sum of the X ranks is K. Then show that

$r_{m,n}(K)= r_{m,n-1}(K) + r_{m-1,n}(K-N), N=m+n$, and

$(m+n)\ p_{m,n}(K)=np_{m,n-1}(K) + mp_{m-1,n}(K-N)$ where $p_{m,n}(K)= r_{m,n}(K)/\binom{m+n}{n}$,under H_0.

16. Obtain one-sided confidence bands based on the statistics $D_n^+$and D_n^-.

17. If $X_{(r_1)}$and $X_{(r_2)}$, $r_1<r_2$, are order statistics for a random sample of size n from a continuous distribution, then prove that $(X_{(r_1)}, X_{(r_2)})$ is a confidence interval for the quantile$_p$, having confidence coefficient $I_p(r_1, n-r_1+1) – I_p(r_2, n-r_2+1)$, where $I_p(v_1,v_2)$ is the incomplete beta function.

18. Show that a 100(1-)% confidence interval for the unknown population median $_e$, based on the sign test, is given by

$$(X_{(r_{\alpha/2}+1)},X_{(r_{\alpha/2})}).$$

19. (Gibbons) For any continuous distribution function F, show that the interval $(X_{(r)}, X_{(n-r+1)})$, with $r<n/2$ is a 100(1-)% confidence interval for the median $_e$ of F, where

$$2n\binom{n-1}{r-1}\int_0^{.5} x^{n-r}(1-x)^{r-1}dx = \alpha.$$

Show that in the case of tolerance limits based on the extreme values, the equation $\frac{1}{\beta(s-r,n-s+r+1)}\int_\beta^1 y^{s-r-1}(1-y0^{n-s+r}dy = \gamma$ reduces to $n\beta^{n-1} - (n-1)\ \beta^n = 1-\gamma$.

STATISTICS QUALITY CONTROL

1. What do you mean by 'statistics quality control' and 'rational subgroups' ?

2. What is meant by assignable cause and chance cause of variation in statistics quality control ? Give the theory behind the control chart technique.

3. Distinguish between process control and lot control in Statistical Quality Control.

4. Give the reasons of using 3-σcontrol limits in control charts of SQC.

5. Describe the construction of control chart for mean from appropriate data when standards are given and also standards are not given.

6. Give the procedure of construction of control charts for range from appropriate data in both cases when standards are given and not given.

7. Write down the control charts of a fraction defective chart when standard is not given.

8. A machine is set to deliver packets of a given weight. Six samples of size 5 each were recorded. Mean and range of each sample are given below :-

Sample No.	1	2	3	4	5	6
Mean	14	18	16	15	17	16
Range	7	6	6	4	8	5

Draw mean chart and range chart, and comment on the state of control. (Given that for n=5, A_2=0.577, D_3=0, D_4=2.115 and for n=6, A_2=0.483, D_3=0, D_4=2.004)

9. 18 lots of electric bulbs, each lot containing 2500 bulbs, have 400, 352, 298, 348, 282, 450, 382, 312, 360, 326, 422, 310, 296, 350, 390, 412, 370, 332 defective bulbs. Draw control chart for number defectives in a graph paper and give your conclusion.

10. The followings are number of defective transistors in 10 lots of 100 transistors each.

17, 6, 10, 8, 10, 14, 7, 17, 2, 15

Construct a suitable control chart and write a brief report on the evidence of the chart.

11. What do you mean by control chart of a variable and control chart of an attribute ? Write down the uses of control chart.

12. Discuss the following terms in connection with sampling inspection plans :
Producer's risk, Consumer's risk, AOQL, ASN Curve and OC Curve.

13. Explain (i) single sampling plan (ii) double sampling plan.

14. Give a general idea about the techniques for determining the constants involved in a single sampling plan.

Sampling distribution

1. If X and Y are $N_2(\mu_X, \mu_Y; \sigma_X^2, \sigma_Y^2; \rho)$, show that X and Y$(\rho\sigma_Y/\sigma_X)$ X are independent and have univariate normal distribution.

   Hence obtain the distribution of of X/Y is case $\mu_X = \mu_Y = 0 = \sigma_X{}^2 = \sigma_Y{}^2 = 1.$

2. X and Y are distributed in the bivariate normal form with correlation coefficient ρ. If the marginal distribution functions of X and Y are respectively F and H, Show that

   (a) The correlation coefficient between X and F(X) is $\sqrt{3/\pi}$,

   (b) The correlation coefficient between X and H(Y) is $\rho\ \sqrt{3/\pi}$.

3. (Continuation) Show that the correlation coefficient between F(X) and H(Y) (called the grade correlation of X and Y) is $\rho_g = \frac{6}{\pi}\sin^{-1}(\rho/2).$ so that $\rho = 2\sin(\pi\rho_g/6).$

4. Suppose $\overline{X_i}$ and S_i^2 (i=1, 2,........, k) are the sample means and variances for independent random samples (of size n_i) from identical normal distributions $N(\mu, \sigma^2).$ Let n = $\sum_i n_i, \overline{X} = \sum_i n_i \overline{X}_I / n$ and

   U = $\frac{1}{\sigma^2}[(n_1-1)S_1{}^2 + (n_2-1)S_2{}^2 + ... + (n_k-1)S_k{}^2]$,

   V = $\frac{1}{\sigma^2}[n_1(\overline{X}_1 - \overline{X})S_1{}^2 + n_2(\overline{X}_2 - \overline{X})S_2{}^2 + ... + n_k(\overline{X}_k - \overline{X})S_k{}^2]$

   And W = $\frac{1}{\sigma^2} n(\overline{X} - \mu)^2$.

Show that U, V, and W are mutually independent χ^2 variables with n-k, k-1 and 1 degrees of freedom respectively.

5. If r be the sample correlation of a random sample of size n ($n \geq 3$) from a bivariate normal distribution with ρ = 0, show that

$$r\sqrt{n-2}/\sqrt{1-r^2}$$

Has the t distribution with n=2 degrees of freedom.

6. (Continuation) For n = 5, show that if

$$\mathrm{P}[|\mathrm{r}| \geq \mathrm{c}] = \alpha \ ,$$

Then c is a root of equation

$$\mathrm{c}\sqrt{1\text{-}\mathrm{c}^2} + \sin^{-1}\mathrm{c} - \frac{\pi}{2}(1-\alpha) = 0.$$

7. Consider a random sample of size 3 from the double exponentia distribution

$$f(x) = \tfrac{1}{2}\exp[-|s|], \ -\infty < s < \infty.$$

Prove that the distribution of the sum S of the sample observations has the p.d.f

$$\tfrac{1}{16}(3+3|s|+s^2)\exp[-|s|], \ -\infty < s < \infty.$$

8. If $\mathrm{X}_1, \mathrm{X}_2, \ldots\ldots, \mathrm{X}_n$ have the multivariate normal distribution such that $\mathrm{E}(\mathrm{X}_i) = \mu$ for each i, var(X_i) = σ^2 for all i and cov($\mathrm{X}_i \mathrm{Y}_j$) = $\sigma^2 \rho$ for all i , j ($i \neq j$), show that $\dfrac{\sqrt{\mathrm{n}}(\overline{\mathrm{X}}-\mu)}{S'}\sqrt{\dfrac{1-\rho}{1+(\mathrm{n}\text{-}1)\rho}}$, where $\overline{\mathrm{X}} = \sum_i \mathrm{X}_i / n$ and $\mathrm{S}'^2 = \sum_{i=1}^{n}(\mathrm{X}_i - \overline{\mathrm{X}})^2/(n-1)$, has the t-distribution with $df = n-1$.

9. If $X_1, X_2, \ldots\ldots\ldots X_p$ have multinormal distribution $N_p(\mu, \Sigma)$, show that x_1 and $x_2 + A_{22}^{-1}A_{21}x_1$ are independently distributed, each in mutinormal form.

10. If a random sample $(X_{1\alpha}, X_{2\alpha}, \ldots\ldots\ldots X_{p\alpha})$, α = 1, 2, ..., n (n > p), comes from the mutinormal distribution $N_p(\mu, (\sigma_{ij}, \delta_{ij}))$, where δ_{ij} is the Kronecker symbol, and if (S_{ij}) is the sample dispersion matrix, show that the p.d.f. of the the elements of the correlation matrix (r_{ij}), where $r_{ij} = S_{ij}$, has the value,

$$\frac{\Gamma^p\left[\frac{n-1}{2}\right]\left|r_{ij}\right|^{(n-p-2)/2}}{\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma\left[\frac{n-i}{2}\right]}$$

Over the part of the space of the r's for which (r_{ij}) is positive definite and the value 0 elsewhere.

11. Let $X_1, X_2, \ldots\ldots\ldots X_p$ have multivariate normal distribution $N_p(\mu, \Sigma)$. Let Y = c' X and Z = d' X, Where $X' = (X_1, X_2, \ldots\ldots\ldots X_p)$ while $c' = (c_1, c_2, \ldots\ldots., c_p), d' = (d_1, d_2, \ldots\ldots.., d_p)$ are vectors with real elements. Obtain the joint distribution of Y and Z and prove that Y and Z are mutually stochastically independent iff $c'\Sigma$d = 0.

12. Prove, by starting from the formula for a suitable correlation coefficient, that

$$a'Ga \le (a'Ta)^{\frac{1}{2}}(a'GT^{-1}Ga)^{\frac{1}{2}},$$

Where a is an arbitrary vector and T and G are positive definite symmetric matrices.

13. The scores of a student in three subjects, A, B and C, are supposed to be jointly

normally distributed, with mean μ_1 = 52.2, μ_2 = 57.6 and μ_3 = 43.5 ; variances $\sigma_{11} = 7.7, \sigma_{22} = 8.3$ and $\mu_{33} = 6.2$; and total correlation $\rho_{12} = 0.36, \rho_{13} = 0.57$ and $\rho_{23} = 0.48$. What percentage of student in a large group is expected to have a total score between 100 and 200 ?

14. Let $\bar{x} = (\bar{x}_1, \bar{x}_2, ... \bar{x}_p)'$ and S = $(s_{ij})(i, j = 1,2, ..., p)$ be based on a random sample of size n from $N_p(\mu, \Sigma)$. Let $x = (x_1, x_2, ..., x_p)'$ be an additional observation from $N_p(\mu, \Sigma)$. Show that $x - \bar{x}$ is distributed according to $N_p\left[0, \left(1 + \frac{1}{n}\right)\Sigma\right]$. . Verify that

$\frac{n}{n+1}(x - \bar{x})' S^{-1} (x - \bar{x})$ has T^2-distribution with n-1 degrees of freedom.

15. If $\bar{x}$ is the vector of sample means for a random sample of size n from $N_p(\mu, \Sigma)$, show that

$$n\bar{x}'\Sigma^{-1}\bar{x}$$

Is distributed as a non-central χ^2 with p degrees of freedom and non-centrality parameter

$$n\mu'\Sigma^{-1}\mu$$

SAMPLING DISTRIBUTION (II)

1. Show that if X and Y are independent Poisson random variables with the same mean λ, then the conditional distribution of X, for any given value of X+Y, is binomial.

2. Also, show that the conditional distribution of X, for any given value of X+Y, is hypergeometric incase X and Y are independent binomial random variables with the same parameter p.

3. Let X and Y be mutually independent negative binomially distributed random variables with parameter (r_1 , p) and (r_2 , p), respectively. Show that X+Y also has the negative binomial distribution with parameter (r_1+r_2 , p). What is the conditional distribution of X, for any given value of X+Y?

4. Let X and X be independent binomially distributed random variables, with parameters (n_1 ,$\frac{1}{2}$) and (n_2 ,$\frac{1}{2}$), respectively. Show that X_1-X_2+n has the binomial distribution with parameters (n_1+n_2 ,$\frac{1}{2}$).

5. Show that if X and Y are normal variables with zero means, unit variance and correlation coefficient ρ, then

$$E\,[\max(X, Y)] = \sqrt{(1-\rho)/\pi}$$

6. Show that the sample variance S^2 obeys the equation

$$(n-1)S^2 = \sum_{i=2}^{n} i(X_i - \bar{X})/(i-1)$$

Where $\bar{X}_i = (X_1 + X_2 + X_3 + \ldots + X_i) / i$.

7. If X_i (i= 1, 2,…, n) are a random sample from $N(\mu, \sigma^2)$, show that

$$\sqrt{\frac{n}{n-1}}(X_1 - \bar{X}) \;/\sqrt{\left\{(n-1)S^2 - \frac{n}{n-1}(X_1 - \overline{X_2})\right\}/(n-2)}$$

Follows the t distribution with df= n-2.

8. If $X_1-\mu$, $X_2- X_1$, X_3-X_2 , … , $X_p - X_{p-1}$ are independently distributed, each as N(0,1), show that the random variables X_1 , X_2 , … , X_p have a p-variate normal distribution with common mean μ and dispersion matrix

9. Let X_1, X_2 be a random sample from R (0, 1). Obtain the distribution function, and hence the p.d.f of X_1+X_2

How should the above result modified in case the population distribution is R (a, b)?

10. Derive the sampling distribution of the geometric mean

$$X_g = \left(\prod_{i-1}^{n} X_i\right)^{\frac{1}{n}}$$

For sampling from the distribution R (0, 1). Indicate how the result will be modified in case the population distribution is R (a, b).

11. Let X and X be distributed as independent gamma variables with parameters () and (), respectively. Show that U= X+X is itself distributed in the gamma form with parameters (), while V=   has the beta distribution. Also obtain the distribution of W=

Use the above result to prove the reproductive property of the χ^2 distribution and to derive the F distribution.

12. A random variable X is said to have a Weibull distribution if its p.d.f is

$f(x) = k\, x^{\beta-1} \exp(-\alpha x^{\beta})$ if x>0

= 0 otherwise, where $\alpha>0$, $\beta>0$

Show that $Y = X^{\beta}$ has, in fact, a gamma distribution.

13.

14. According to the Maxwell-Boltzmann law of theoretical physics, the p.d.f of V, the velocity of a gas molecule, is

$f(v) = kv^2 \exp(-\beta v^2)$ for $v>0$

$= 0$ otherwise

where $\beta>0$ depends on the mass and absolute temperature of the molecule and k is the appropriate constant. Show that the kinetic energy $E=\frac{1}{2}mV^2$ is a gamma random variable.

15. Let the variables X_i = (i=1, 2, 3) be independently and identically distributed in the form N (0, 1). Obtain, from appropriate statistical tables, the probabilities

(a) $P[2X_1 + 3X_2 + 5X_3 \geq 5]$

(b) $P[X_1^2 + X_2^2 + X_3^2 \geq 5]$

(c) $P[X_1^2 - X_2^2 - X_3^2 \leq 0]$

(d) $P[X_1^2 - 2X_2^2 - 2X_3^2 \leq 0]$

(e) $P[4X_1^2 + 4X_2^2 - 7X_3^2 \leq 0]$

16. Evaluate the probability that for independent random samples of size n=5 each from two normal distributions with the same variance, the ratio S_1^2/S_2^2 of the sample variance will be between 1/3 and 3.

17. Let X and Y be independently distributed, each in the form N (0, 1). Show that Z= X/Y has the Cauchy distribution with p.d.f

$f(z) = \frac{1}{\pi[1+z^2]}$, $-\infty < z < \infty$

What would the distribution for $W = X/|Y|$?

18. If X_i (i= 1, 2, 3, 4) are independent N (0, 1) variables, Show that $U = X_1X_2 - X_3X4$ has the p.d.f

$f(u) = \frac{1}{2}\exp(-|u|)$.

Let X_1 and X_2 be independent observations from the distribution with p.d.f

$f(x) = 2x$, $0 < x < 1$.

Evaluate the probability: $P[X_1 < X_2 \mid X_1 < 2X_2]$.

19. Let X_1 and X_2 be a random sample from a rectangular distribution R (0, 1). Show that $V = |X_1 - X_2|$ has the p.d.ff(v) = 2(1-v), $0 \le v \le 1$

20. If X and Y are independent random variables each distributed uniformly over (0, 1), find the distribution of (i) X/Y, (ii) XY and (iii) $\sqrt{(X^2 + Y^2)}$.

21.

22.

23. If X_1 and X_2 are stochastically independent with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively, show that

$E(X_1 X_2) = \mu_1 \mu_2$

$\mathrm{Var}(X_1 X_2) = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2$

24.

25. If X and Y are independent normal standard variables, show that the mean value of max (|X|, |Y|) is $2/\sqrt{\pi}$.

26. Let X_1 and X_2 be independently and identically distributed random variables. Show that $Y = X_1 - X_2$ has a symmetrical distribution.

27.

28. Let X_i (i= I, 2, ..., 2n) be independent random variables each distributed as $N(\mu, \sigma^2)$. Find the distribution of

(i) $(X_1 + \ldots. + X_n - X_{n+1} - \ldots. - X_{2n})/ 2n$,

(ii) $(X_1 - X_2)^2 + (X_3 - X_4)^2 + \ldots + (X_{2n-1} - X_{2n})^2$.

29. If X_i (i = 1, 2, …, n) are a random sample from $N(\mu, \sigma^2)$ and $V = \frac{1}{n}\sum_i |X_i - u|$ then what are E(V) and var(V)?

30. Let X_1 and X_2 be independent random variables each distributed as a χ^2 with 2 degrees of freedom. Find the distribution of

$U = \frac{a_{1X_1}+a_{2X_2}}{X_{1+}X_2}$, where a_1 and a_2 are constants such that $a_1 > a_2$.

31. If Y_1 and Y_2 are independent variables, each with n degrees of freedom, show that

$$\frac{\sqrt{n}(Y_1 - Y_2)}{2\sqrt{Y_1 Y_2}}$$

Has the t distribution with n degrees of freedom and is independent of $Y_1 + Y_2$.

32.

33. If X and Y are independent N (0, 1) variables, show that $XY/\sqrt{X^2 + Y^2}$ is distributed as N $(0, \frac{1}{4})$.

34. Let X and Y be independent random variables, having the distributions B (a, b) and b (c, d), where a= c+d, show that XY has the distribution B (c, b+d).

35. Let X_1 and X_2 be independently distributed random variables, each in the form R (0, 1), show that $U_1 = \sqrt{-2\ln X_1}\cos 2\pi X_2$ and $U_2 = \sqrt{-2\ln X_1}\sin 2\pi X_2$ are independently distributed N (0, 1) variables.

36. If X follows the beta (n_1, n_2) distribution, Y follows the beta $(n_1+\frac{1}{2}, n_2)$ distribution and the two are independent, then what is the distribution of $Z = \sqrt{XY}$?

37. Let X be a discrete random variable with uniform distribution on {0, 1, 2, …, r-1}, where r = ab and a and b are positive integers such that 1<a, b<r. Show that the distribution of X is same as that of U + V where U and V are independent random variables both with uniform distribution on appropriate subsets of {0, 1, 2, …, r-1}.

38. Let X have the N(0, 1) distribution and c be a non-negative number, and let Y = -X if |X|≤ c and Y = X if |X|> c. Show that Y also has the N(0, 1) distribution. Find the value of c for which X and Y are uncorrelated.