

17.4 DEC 2022

o/c

**Deccan Education Society's
Fergusson College (Autonomous), Pune-4
S.Y.B.A. (Semester-IV)**

Applied Statistics

**STA2401: Continuous Univariate Distributions and Applications of
Statistics (2019 pattern)**

Time: $2\frac{1}{2}$ Hrs.]

[Max. Marks: 50

Instruction to candidates:

- i) All questions are compulsory
- ii) Figures to the right indicate full marks
- iii) Use of statistical table and scientific calculator is allowed.

Q.1 Attempt any five questions:

[2 x 5 = 10]

- a) Define the following
 - i) Multiple correlation coefficient
 - ii) partial correlation coefficient
- b) Write a note on multiple regression plane.
- c) State any two uses of time series analysis.
- d) Let $X \rightarrow N(15, 4)$ find (i) $P(13 \leq X \leq 17)$
(ii) $P(X \geq 15)$.
- e) State the probability density function (p.d.f) of exponential distribution.
Also state any two real life situations where it is applicable.
- f) Define distribution function of a continuous type r.v and state any two properties of it.
- g) Calculate CDR and age specific death rates for the population from the following data:

Age group	0 - 9	10 - 19	20 - 29	30 & above
Population (in thousands)	20	35	38	40
Number of deaths	410	180	195	220

Q.2 Attempt any five questions:

[3 x 5 =15]

- a) The following computation were made on 120 students
 $r_{xy_1} = 0.6$, $r_{yx_2} = 0.7$, $r_{x_1x_2} = 0.65$

Where symbols have their usual meaning. Find the value of multiple correlation coefficient $R_{y.x_1x_2}$.

- b) Let X_1 and X_2 be two independent normal variates. Their respective means are 2500 and 2000 in appropriate units. Their standard deviations are 500 and 400 respectively. Obtain the probabilities of the following events.
 i) $P(|X_1 - X_2| \leq 100)$
 ii) $P(X_2 > 2000)$

- c) Let X be a normal random variable with probability density function given by

$$f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} & -\infty < x < \infty \\ & \sigma > 0 \\ & -\infty < \mu < \infty \\ 0 & \text{otherwise} \end{cases}$$

Identify mean and standard deviation of X .

- d) Explain the method of semi-averages to estimate trend in time series analysis.
 e) Explain the need of vital statistics.
 f) Define (i) Sex ratio
 (ii) Dependency ratio
 g) Compute 'General Fertility Rate' for the following data:

Age group	15-19	20-24	25-29	30-34	35-39	40-44
Female population	9000	9200	8600	8800	8700	8100
No. of births	183	1202	1268	785	520	190

Q.3 Attempt any five questions:

[5 x 5 =25]

- a) Suppose the life of the expectancy X of each member of a certain group of people is a random variable having exponential distribution with parameter.
 $\alpha = 50$ years.
 Compute the probability that:
 i) Person will survive at least 70 years
 ii) Person will die before 65 years.

- b) Let $X_1 \sim N(4, 4^2)$ and $X_2 \sim N(5, 5^2)$ be independent random variables.
Find i) $P[X_1 + X_2 \geq 8]$
ii) $P[2X_1 + X_2 \leq 11]$
- c) Explain trend and irregular variations in time series analysis with one illustration of each.
- d) Define the following terms:
i) General Fertility Rate (GFR)
ii) Age – Specific Fertility Rate (ASFR)
iii) Net Reproduction Rate (NRR)
- e) The mean, standard deviation and correlation coefficients for three variables:
Y : yield of crop
X₁ : water quantity
X₂ : Fertilizer
(measured in proper units) are given below:
 $\bar{Y} = 15.9$ $\sigma_y = 1.7$ $r_{yx_1} = 0.66$
 $\bar{X}_1 = 3.7$ $\sigma_{x_1} = 1.3$ $r_{x_1x_2} = 0.60$
 $\bar{X}_2 = 5.8$ $\sigma_{x_2} = 3.1$ $r_{yx_2} = 0.8$
 Obtain the least squares equation of plane of regression of Y on X₁ and X₂.
- f) Compute seasonal indices for the following data using method of simple averages:

Quarter Year	I	II	III	IV
2000	60	66	79	86
2001	83	80	85	74
2002	67	65	94	65
2003	71	81	83	60

- g) Calculate (S.T.D.R) Standardized Death Rate for district A and B compare them:

Age group	District A		District B		Standard population (in'000)
	No. of persons (in'000)	No. of deaths	No. of persons (in'000)	No. of deaths	
0 - 10	2	50	1	20	200
10 - 55	7	49	3	30	580
55 & above	1	25	2	40	220

o/c

26/11/22

**Deccan Education Society's
Fergusson College (Autonomous), Pune-4
S.Y.B.A. (Semester - III)**

Applied Statistics

**STA2301: Probability Theory and Discrete Probability Distributions
(2019 pattern)**

Q1 Answer any five of the following Questions: (2x5=10)

- a. Define the events with one illustration each:
- i) Complementary Event,
 - ii) Sure event.
- b. Show that ${}^nC_r + {}^nC_{r-1} = {}^{n+1}C_r, r \geq 1$
- c. Give the classical definition of probability.
- d. Let A and B be the two events defined on sample space Ω , suppose $P(A) = 0.4, P(A \cup B) = 0.7$ and $P(B) = p$
- i) For what choice of p are A and B mutually exclusive,
 - ii) For what choice of p are A and B independent.
- e. Show that $P(A') = 1 - P(A)$
- f. State the probability mass function of binomial distribution also state its mean and variance.
- g. Asteroids with a diameter of at least 1 km collide with the earth at a rate of approximately 2 per million years. What is the probability that in a randomly selected million year period, there is exactly one collision?

Q2 Answer any five of the following Questions: (3x5=15)

- a. In a group of 10 men, 6 are graduates. If 3 men are selected at random, what is the probability that they consists of (i) all graduates (ii) at least one graduate.
- b. Distinguish between deterministic and non deterministic experiments.
- c. State the limitations of classical definition of probability.
- d. State the properties of Distribution Function.
- e. A radar system has a probability of 0.1 of detecting a certain target during a single scan. Find the probability that the target will be detected (i) at least twice in four scans (ii) at most once in four scans.
- f. State the additive property of Poisson distribution. Also if $X \rightarrow P(\lambda)$ such that $P(X=2) = P(X=3)$ then compute $E(X)$.

(5x5=25)

Q3 Answer any five of the following Questions:

- g. A discrete random variable has the following p.m.f;
- $$P(X=x) = \frac{x}{15} \text{ ; for } x = 1, 2, 3, 4, 5$$
- $$= 0 \text{ otherwise}$$
- Compute $E(X^2 + 4X - 5)$

- a. Let $P(A) = 0.38$, $P(A \cap B \cap C) = 0.10$, $P(A \cap B \cap C) = 0.17$ and $P(A \cap B \cap C) = 0.12$, find the probability of the occurrence of at least two of the three events A, B and C defined on sample space Ω

- b. Two fair dice are thrown independently. The three events A, B and C are defined as follows:
 A: odd number appears on the first dice
 B: odd number appears on the second dice
 C: sum of numbers on the two dice is odd
 Are the events A, B and C pair wise independent or mutually independent?

- c. The calculation of a home insurance premium involves seven consecutive computational steps. An auditor is reviewing a large number premium calculations that contains an error. Let X denotes the steps at which the error is made. Assume that X follows a discrete uniform distribution, calculate (i) $P(3 \leq X \leq 5)$ (ii) on an average, how many steps must the auditor check in each premium calculations to find the error?

- d. The probability mass function (p.m.f) of a discrete random variable is :
- $$P(X=x) = \frac{2x}{n(n+1)} \text{ ; } x = 1, 2, 3, \dots, n$$
- $$= 0 \text{ otherwise}$$
- Calculate $E(X)$ and $V(X)$.

- e. (i) State the axioms of probability
 (ii) State the additive property of binomial distribution. Also if $X \rightarrow B(5, 1/2)$ and $Y \rightarrow B(8, 1/2)$ and if X and Y are independent, then compute $P\left(\frac{X+Y}{2} \geq 1\right)$

- f. Define i) Conditional probability
 iii) Expectation of a discrete random variable
 iv) Distribution function

- g. i) Define Probability model.
 ii) Define Permutation and combination with one example each.
 iii) State the multiplication theorem.

25/11/2022 o/c

Deccan Education Society's
Fergusson College (Autonomous), Pune-4

F.Y.B.A. (Semester-I)

Applied Statistics

STA1101: Descriptive Statistics - I (2019 pattern)

Time: 2 $\frac{1}{2}$ Hrs.]

[Max. Marks: 50

Instruction to candidates:

- i) All questions are compulsory
- ii) Figures to the right indicate full marks
- iii) Use of statistical table and scientific calculator is allowed.

Q.1 Attempt any five questions:

[2 x 5 =10]

- a) State functions of central statistical office (CSO).
- b) Explain attributes and variables with one illustration each.
- c) Find the first quartile for the following data:

Class	5 - 10	10 - 15	15 - 20	20 - 25	25 - 30
frequency	7	3	10	4	2

- d) Find harmonic mean for the following data:
5, 7, 9, 13, 8, 12, 14
- e) Define raw moments and central moments.
- f) For certain frequency distribution first raw moment is 10, second central moment is 25 and mode is 15.
Find Karl - Pearson's coefficient of skewness and comment on the nature of distribution.
- g) Explain the method of construction of price index number based on geometric mean of price relatives.

Q.2 Attempt any five questions:

[3 x 5 =15]

- a) Explain the scope of statistics in the field of education and economics.
- b) Distinguish between time series data and cross sectional data.
- c) Find the mean deviation about median for two following data:

Class	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50
frequency	10	7	13	12	8

- d) Find the A.M. and geometric mean for the following data:
5, 7, 9, 13, 08, 12, 14

- e) The mean salary paid to 75 employees in a company is Rs. 1420. The mean salary of 25 of them is Rs. 1350 and that of 30 others is Rs. 1425. Find the mean salary of remaining employees.
- f) Find coefficient of variation for the following data:
Height of person: 155, 149, 170, 177, 179, 180
(in cms.)
- g) Discuss merits and demerits of arithmetic mean and median.

Q.3 Attempt any five questions:

[5 x 5 =25]

- a) i) Draw all possible systematic samples of size 2 from the following population of size 10.

Hemoglobin in body	11.5, 10.5, 9.5, 12.5, 13.5, 14, 14.9, 11.7, 12.6, 15.4
--------------------	---

- ii) Explain two stage random sampling with one real life example.

- b) i) Find combined standard deviation for the following data:

Group - I	Group - II
$n_1 = 150$	$n_2 = 200$
$\bar{X}_1 = 70$	$\bar{X}_2 = 80$
$\sigma_1^2 = 1.5$	$\sigma_2^2 = 2$

- ii) State the properties of good measures of central tendency.

- c) Find quartile deviation for the following data:

x_i	5	10	15	20	25	30	40
frequency	5	7	13	11	09	10	15

- d) Explain the following
i) measures of skewness
ii) measures of kurtosis

- e) Find Fisher's price index number based on the following data:

Commodity	Year 2010		Year 2015	
	price	quantity	price	quantity
L	200	10	210	9
M	400	15	450	7
N	500	5	750	6

- f) i) Show that first central moment is always zero.
ii) For certain frequency distribution median is 25, upper and lower quartiles are 40 and 10 respectively. Find Bowley's coefficient of skewness and comment on the value obtained.
- g) i) Define consumer price index number.
Also explain aggregate expenditure method of construction of consumer price index number.
ii) State the applications of index numbers.

26/11/2022 o/c

Deccan Education Society's
Fergusson College (Autonomous), Pune-4

T.Y.B.A. (Semester -V)

Applied Statistics

STA3501 :Testing of Hypotheses (General) (2019 pattern)

Time: 2 $\frac{1}{2}$ Hrs.]

[Max. Marks: 50

Instruction to candidates:

- i) All questions are compulsory
- ii) Figures to the right indicate full marks
- iii) Use of statistical tables and scientific calculators is allowed.

Q.1 Attempt any five questions:

[2 x 5 =10]

- a) Define the following:
 - i) Null hypothesis
 - ii) Critical region
- b) Explain the concept of sampling distribution of statistic $T(X_1, X_2, \dots, X_n)$.
- c) Define student's t – variate with n degrees of freedom.
- d) Let $X \rightarrow \chi_n^2$. State $E[X]$ and $\text{Var}[X]$. Further plot the probability curve of X .
- e) Explain the concept of analysis of variance.
- f) State the test statistic of Wilcoxon's signed rank test.
- g) State the test statistic for testing $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Q.2 Attempt any five questions:

[3 x 5 =15]

- a) Let $U \rightarrow N(0, 1)$ and $V \rightarrow \chi_n^2$. If U and V are independent r.v.s, find
$$P\left[\frac{U + \sqrt{V/n}}{\sqrt{V/n}} \leq 1\right].$$
- b) If $Y \rightarrow \chi_{16}^2$, find
 - i) $P[Y \leq 20.465]$
 - ii) median of Y .
- c) Explain the chi-square test for independence of attributes.

- d) Complete the following ANOVA table.

Source	d.f	s. s.	ANOVA	F comp
treatment	2	2.8	----	-----
Error	----	9.6	0.8	
Total	14			

- e) A die is rolled 180 times with the following results:

No. on the die	1	2	3	4	5	6
frequency	28	36	36	30	27	23

Test whether the die is balanced.

Use $\alpha = 0.01$

- f) Let $X \rightarrow F_{2,8}$. If $P[X \leq a] = 0.01$ and $P[X \leq b] = 0.99$, find the values of constants a and b .
- g) Describe one sample run test.

Q.3 Attempt any five questions:

[5 x 5 = 25]

- a) Of 150 democrates interviewed, 90 favour a proposition and of 120 republicans interviewed 80 favour it. Is the proportion in favour of the proposition significantly larger among republicans than among democrates? Use 1% l.o.s.
- b) A teacher wants to determine, whether a certain student has an average reading speed of 800 words per minute or whether his average speed is greater than 800 words per minute. What can be concluded at 5% l.o.s. if 6 students read 780, 800, 790, 750, 810 and 750 words per minute.
- c) The productivity of 12 workers were observed during one week when they under strict supervision and during another week when they were not. The following scores provide a measure of their productivity.

With supervision	81	59	78	85	60	50	77	79	53	66	89	67
Without supervision	72	52	83	88	58	30	82	83	40	66	76	63

Test whether the strict supervision affects the productivity. Use $\alpha = 0.02$

- d) Explain the test for testing $H_0 : P_1 = P_2$ against $H_1 : P_1 \neq P_2$, where P_1 and P_2 are populations of proportions of two populations. The sizes n_1 and n_2 of random samples, taken from these populations, are large.

- e) The following table is the life in hours of 4 batches of electric lamps:

Batches	Life in hours							
1	1600	1610	1650	1680	1700	1720	1800	
2	1580	1640	1640	1700	1750			
3	1640	1550	1600	1620	1640	1660	1740	1820
4	1510	1520	1530	1570	1600	1680		

Test whether the 4 batches are producing electric lamps with the same average life time. Use $\alpha = 0.01$.

- f) Explain the following:
- Statistic
 - Standard error of statistic
 - Level of significance
 - concept of distribution free statistic
- g) i) Explain break up of total sum of squares in case of one – way and two way classification.
- ii) Explain distinction between parameter and non-parametric problem.

17 DEC 2022

10.30 O/C

**Deccan Education Society's
Fergusson College (Autonomous), Pune-4
F.Y.B.A. (Semester-II)
Applied Statistics**

STA1201: Applied Statistics (Descriptive Statistics - II) (2019 pattern)

Time: 2½ Hrs.]

[Max. Marks: 50

Instruction to candidates:

- i) All questions are compulsory
- ii) Figures to the right indicate full marks
- iii) Use of statistical table and scientific calculator is allowed.

Q.1 Attempt any five questions:

[2 x 5 =10]

- a) With the help of scatter diagram explain types of correlation?
- b) Define Karl Pearson's correlation coefficient.
- c) If the fitted simple linear regression model for following bivariate data is
 $\hat{Y}_i = 24.9 + 21X_i$

Time (hours) (X)	5	7.5	15	21	22
Cost to complete job (Y)	124	174	325	430	478

Find the residuals.

- d) State formulae of regression coefficients for simple linear regression model. Also give interpretation of regression coefficients in simple regression model.
- e) State least squares normal equations to fit model of type $Y = aX^b$.
- f) Complete the 2X2 contingency table based on the following information:
 $(A) = 90, (AB) = 40, N = 150, (\beta) = 80.$
- g) Use method of semi-averages to determine the trend of following time series data:

Years	1977	1978	1979	1980	1981	1982
Profit in (Rs.000)	21	24	22	30	29	31

Q.2 Attempt any five questions:

[3 x 5 =15]

- a) Fit simple linear regression model for the following data:

Rainfall (X)	8.8	10.5	15.9	13.1	12.9
Yield of wheat (Y)	39.4	42.7	69.3	52.4	60.5

- b) i) State unbiased estimator of error variance σ^2 in simple linear regression model.
- ii) State least squares normal equations to fit a model of type $Y = ab^X$.
- c) Explain the plot of residuals in simple linear regression model with an illustration.
- d) Explain any two components of time series with one illustration each.
- e) Check whether the following data are consistent:
 $N = 500$, $(A) = 300$, $(B) = 250$, $(AB) = 40$.
- f) Explain least squares method to estimate trend in time series analysis.
- g) Find Spearman's rank correlation coefficient for the following data:

Ranks by judge A	1	2	3	4	5	6
Ranks by judge B	2	1	4	3	6	5

Q.3 Attempt any five questions:

[5 x 5 =25]

- a) Compute Karl Pearson's correlation coefficient for the following data:

Age	23	24	27	28	39	41	49
% of fat in body	9.5	27.9	7.8	17.8	32	25.9	26

Also comment on the result.

- b) Fit second degree curve $Y = a + bX + cX^2$ the following data:

Price of toy (X)	1	1.5	2	2.5	3
Total profit (Y)	8.6	9.8	10.2	8.8	6.9

- c) Use method of 4 yearly moving averages to estimate trend based on the following data:

Year	Production
1972	614
1973	615
1974	652
1975	678
1976	681
1977	655
1978	717
1979	719
1980	708
1981	779
1982	758

- d) Write note on multiplicative and additive models of time series.
- e) In a coeducational institution, out of 200 students 150 were boys. They gave an examination and it was found that 120 boys were passed, 10 girls had failed. Compute Yules coefficient of association to find association between sex and success in examination.
- f) i) Explain the concept of independence of two attributes with illustration.
ii) Explain dichotomy and manifold classification with one real life example each.
- g) i) Define coefficient of determination. Also state its interpretation.
ii) Find coefficient of determination based on the following data:

Age (X)	Blood pressure (Y)	Residuals
56	147	1.1597
42	125	-5.5803
36	118	-6.0403
47	128	-8.303
49	145	6.7897
42	140	9.4197
72	155	-8.2803
63	160	6.5297
55	149	4.2497
60	150	-0.2003
