Total No. of Questions—**5**] [Total No. of Printed Pages—**8**

| **Seat No.** | |
|---|---|

# [5116]-7

## F.Y. B.Sc. (Computer Science) EXAMINATION, 2017

## STATISTICS

## Paper I

## (Statistical Methods—I)

## (2013 PATTERN)

**Time : Three Hours** **Maximum Marks : 80**

***N.B.*** **:—** (*i*) *All* questions are compulsory

(*ii*) Figures to the right indicate full marks.

(*iii*) Use of non-programmable, scientific calculator and statistical tables is allowed.

(*iv*) Symbols have their usual meaning unless otherwise stated.

**1.** (A) Fill in the blanks : [1 each]

(*i*) For a negatively skewed distribution, the relationship between mean, media and mode is .....

(*ii*) The mean of Poisson distribution is ..... its variance.

(*iii*) Limits of multiple correlation coefficeint $R_{1.23}$ are .....

(*iv*) In Time series, the component having period of oscillation less than one year is called .....

P.T.O.

(B) Select the most appropriate option for each of the following : [1 each]

(*i*) For deciding most favorite actor, which is the most appropriate average ?

(*a*) mean

(*b*) median

(*c*) mode

(*d*) upper quartile

(*ii*) Variance is a measure of :

(*a*) dispersion

(*b*) central tendency

(*c*) kurtosis

(*d*) skewness

(*iii*) The probability distribution of a discrete random variable X is :

| $x_i$ | $P(X = x_i)$ |
|---|---|
| 1 | 0.1 |
| 2 | 0.3 |
| 3 | 0.4 |
| 4 | 0.2 |

What is $P(2 \leq x \leq 3)$ ?

(*a*) 0.7

(*b*) 0.3

(*c*) 0.4

(*d*) 1

(*iv*) The number of observations belonging to a class intervals is called as :

(*a*) cumulative frequency

(*b*) class width

(*c*) class mark

(*d*) frequency

(C) Attempt each of the following : [2 each]

(*i*) State AR(1) model.

(*ii*) If $b_{yx}$ = 0.7 and $b_{xy}$ = 0.9, find the value of $r$.

(*iii*) State recurrence relation for the Binomial distribution.

(*iv*) If X follows discrete uniform distribution with $n$ = 7, find the variance of X.

**2.** Attempt any *four* of the following : [4 each]

(*a*) Discuss median as a measure of central tendency. State merits and demerits of median.

(*b*) The daily expenditure of 100 people is as follows :

| **Expenditure** | **Number of Persons** |
|---|---|
| 20—30 | 14 |
| 30—40 | — |
| 40—50 | 27 |
| 50—60 | — |
| 60—70 | 15 |

If the mode of the distribution is 43, find the missing frequencies.

(*c*) Describe the procedure to plot less than ogive curve for a grouped frequency distribution.

(*d*) Consider the following data related to income in two villages :

| | **Village A** | **Village B** |
|---|---|---|
| Number of persons | 70 | 60 |
| Mean income (Rs.) | 280 | 310 |
| Variance of income | 144 | 169 |

(*i*) In which village the average income is more ? Justify your answer.

(*ii*) In which village the variation in income is more ? Justify your answer.

(*e*) Define quartiles. Describe procedure to compute third quartile for a grouped frequency distribution.

(*f*) The profits (in lakhs of Rs.) of 15 companies for financial year 2015-16 are as follows :

24, 21, 35, 48, 42, 27, 52, 43, 40, 47, 55, 25, 50, 33, 44.

Draw a stem and leaf chart.

**3.** Attempt any *four* of the following : [4 each]

(*a*) Explain the terms exclusive class interval and coefficient of variation.

(*b*) Consider the following data related to marks of students in division A and division B in statistics :

| | | | |
|---|---|---|---|
| Division A | $Q_1 = 23$ | $Q_2 = 52$ | $Q_3 = 78$ |
| Division B | $Q_1 = 34$ | $Q_2 = 52$ | $Q_3 = 68$ |

Determine marks of which division are more skewed ? Justify your answer.

(*c*) Explain concept of kurtosis. State its types with help of frequency curve.

(*d*) The standard deviation of a distribution is 5. What should be the value of fourth central moment so that distribution will be (*i*) mesokurtic (*ii*) leptokurtic ?

(*e*) Consider the function $P(x) = K(x^2 + 4)$, $x = 0, 1, 2, 3$ :

(*i*) Find the value of K for which P(*x*) will be valid p.m.f.

(*ii*) Find distribution function of X.

(*iii*) Find the value of mode of X.

(*f*) Describe in brief a Binomial experiment. State probability mass function (p.m.f.) of Binomial distribution. Also state expression for its mean.

**4.** Attempt any *two* of the following : [8 each]

(A) (*i*) What is regression ? State any *two* properties of regression coefficients.

(*ii*) For a trivariate data, $\sigma_1 = 4$, $\sigma_2 = 8$, $\sigma_3 = 7$, $r_{12} = 0.45$, $r_{13} = 0.55$, $r_{23} = 0.65$. Find the values of $b_{12.3}$ and $r_{23.1}$.

(B) (*i*) Explain in brief the procedure of fitting line of regression of X on Y for a bivariate data by method of least squares.

(*ii*) The following is the distribution function of a discrete random variable X :

| **X** | **F(*x*)** |
|---|---|
| 0 | 0.05 |
| 1 | 0.20 |
| 2 | 0.40 |
| 3 | 0.90 |
| 4 | 0.99 |
| 5 | 1.0 |

(1) Find p.m.f. of X

(2) Find P(X > 3)

(3) Find P(1 < X $\leq$ 5).

(C) (*i*) Explain concept of partial correlation in a trivariate data with help of an example.

(*ii*) Let X follows Poisson distribution with parameter 4 and Y follows Poisson distribution with parameter 6. X and Y are independent.

Find the distribution of (X + Y). Also find P[(X = 5)/(X + Y) = 9].

(D) (*i*) If the probability that a certain test gives a positive reaction is 0.4. What is the probability that less than 3 negative reactions occur before the first positive reaction.

(*ii*) Consider the following calculations for a bivariate data of size 10 :

$\Sigma x = 165$ $\Sigma y = 178$ $\Sigma x^2 = 3591$

$\Sigma y^2 = 3788$ $\Sigma xy = 3606$

Find the correlation coefficient between X and Y and interpret its value.

**5.** Attempt any *one* of the following : [16 each]

(A) (*i*) Describe the procedure of fitting equation $y = ax^b$ for a bivariate data.

(*ii*) In the regression analysis the equation of two lines of regression are 2X + 3Y = 8 and 2Y + X = 5 and the variance of X = 4.

Find :

(1) Mean values of X and Y

(2) Coefficient of correlation between X and Y

(3) The standard deviation of Y.

(B) (*i*) A teacher of mathematics wants to determine the relationship between grades in the final examination and two internal tests given during the semester. Let $X_1$, $X_2$ and $X_3$ denote the grades of a student in the final examination, first test and second test respectively. He obtained the following computation for a total of 120 students :

$\bar{X}_1 = 7.4$ $\bar{X}_2 = 6.8$ $\bar{X}_3 = 7.0$

$\sigma_1 = 0.9$ $\sigma_2 = 1$ $\sigma_3 = 0.8$

$r_{12} = 0.60$ $r_{13} = 0.70$ $r_{23} = 0.65$

(1) Find the equation of plane of regression of $X_1$ on $X_2$ and $X_3$.

(2) Estimate $X_1$ when $X_2 = 6$ and $X_3 = 6.5$.

(*ii*) Estimate trend value using method of moving averages with $m = 4$ for the following data on the number of students studying in a college during years 2001 to 2010 :

| **Year** | **Number of Students** |
|---|---|
| 2001 | 3320 |
| 2002 | 3170 |
| 2003 | 3570 |
| 2004 | 3920 |
| 2005 | 4020 |
| 2006 | 4050 |
| 2007 | 4100 |
| 2008 | 4270 |
| 2009 | 4050 |
| 2010 | 4380 |