2 0 1 7

STATISTICS

(Major)

Paper : 1.1

(Descriptive Statistics)

Full Marks : 60

Time : 3 hours

The figures in the margin indicate full marks for the questions

1. Answer the following questions as directed (reasoning is not necessary) : 1×7=7

(a) Write whether true or false :
"Data obtained from physical experiments are secondary data."

(b) Is it true that population is defined as a set of human beings?

(c) Is mode a partition value?

(d) Is coefficient of variation invariant of change of scale?

(e) State the range of multiple correlation coefficient.

(f) Define geometric mean of the values $x_1, x_2, \cdots, x_n$ in terms of arithmetic mean.

(g) State the values of β_1 and β_2 for a symmetric distribution.

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

(a) Mention two limitations of statistics.

(b) If x_i / f_i $(i = 1, 2, \cdots, n)$ is a frequency distribution and $u_i = \frac{x_i - a}{h}$, then show that $\bar{x} = a + h\bar{u}$. (Symbols have their usual meanings.)

(c) If r_{XY} is the coefficient of correlation between X and Y, then interpret the cases where—

(i) $r_{XY} = +1$;

(ii) $r_{XY} = -1$.

(d) State Sheppard's corrections for moments for grouped data.

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

(a) What is a statistical table? Mention with explanation the main parts of a statistical table.

(b) Define arithmetic mean of a discrete frequency distribution. Show that the algebraic sum of the deviations of observations for the frequency distribution is minimum when taken about mean.

(c) Define the following :

(i) Coefficient of correlation

(ii) Regression coefficients

(iii) Partial correlation coefficient

(iv) Multiple correlation coefficient

(v) Correlation index

(d) What are partition values of a distribution? Explain (with definition) median and quartiles as partition value.

(e) Obtain the normal equations for fitting of the 2nd-degree parabola $Y = a + bX^2$ on the basis of the n pairs

$$(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$$

of values of (X, Y).

4. Answer *either (a) or (b)* :

(a) Define the following with suitable examples : 2×5=10

(i) Qualitative and quantitative data

(ii) Normal and ordinal data

(iii) Cross-sectional and time-series data

(iv) Discrete and continuous data

(v) Frequency and non-frequency data

(b) What are primary data and secondary data? Clearly mention various sources of secondary data. 2+2+6=10

5. Answer *either (a) or (b)* :

(a) Define raw moments, standard moments and factorial moments of a set of non-frequency numerical data. Express the 4th-order standard moment in terms of raw moments. 2+2+2+4=10

(b) Define standard deviation of the observed values $x_1, x_2, \cdots, x_n$. If $\sigma_1^2, \sigma_2^2, \cdots, \sigma_p^2$ are the variances of p different sets containing $n_1, n_2, \cdots, n_p$ observed values respectively, then find out the variance of all the $n_1 + n_2 + \cdots + n_p$ observed values. 2+8=10

6. Answer *either (a) or (b)* :

(a) Explain the principle behind the method of least squares of fitting a mathematical curve $y = f(x)$ to a set of numerical data viz. $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ on (x, y). Find the normal equations for fitting of the mathematical curve

$$y = a + bx + cx^2 + dx^3$$

to the data on (x, y) mentioned above. 4+6=10

(b) Write notes on any *two* of the following : 5×2=10

(i) Skewness and Kurtosis

(ii) Orthogonal polynomials

(iii) Graphic representation of data

★★★

2017

STATISTICS

(Major)

Paper : 1·2

(Probability—I)

Full Marks : 60

Time : 3 hours

The figures in the margin indicate full marks for the questions

1. Answer the following as directed : 1×7=7

(a) If A and B are two events, then the probability of occurrence of at least one of them is given as

(i) $P(A)+P(B)$

(ii) $P(A \cap B)$

(iii) $P(A \cup B)$

(iv) $P(A)\,P(B)$

(Choose the correct option)

(b) With a pair of dice thrown at a time, the probability of getting a sum more than 9 is

(i) $\frac{5}{18}$

(ii) $\frac{7}{36}$

(iii) $\frac{5}{16}$

(iv) None of the above

(Choose the correct option)

(c) For two events

$$P(A) = P(A / B) = \tfrac{1}{4},\ P(B / A) = \tfrac{1}{2}$$

find the value of $P(B)$.

(d) For a continuous random variable X, the value of the probability $P(X = c)$, for all possible values of c is _____.

(Fill in the blank)

(e) If X assumes only positive values and $E(X)$ and $E(\frac{1}{X})$ exist, then $E(\frac{1}{X}) \leq \frac{1}{E(X)}$.

(State True or False)

(f) Define conditional expectation $E(X / Y)$ for two discrete random variables X and Y.

(g) A random variable may have no _____ although its moment-generating function exists.

(Fill in the blank)

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

(a) Define complement of an event. If $\overline{A}$ is the complement of event A, then show that $P(\overline{A}) = 1 - P(A)$.

(b) Explain the term 'conditional probability'. Find $P(B / A)$ if A and B are independent events.

(c) State the important properties of distribution function.

(d) Can $P(s) = \frac{2}{1+s}$ be the probability-generating function (pgf) of a random variable? Give reasons.

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

(a) Distinguish between mutually exclusive events and independent events. Show that two independent events each having non-zero probabilities cannot be mutually exclusive.

(b) Three persons A, B and C in order toss a fair coin. The first one who throws a 'head' wins. If A starts, find their respective chances of winning. (Assume that the game may continue indefinitely).

(c) In answering a question on a multiple choice test, a student either knows the answer or he guesses. Let p be the probability that he knows the answer and $1-p$ the probability that he guesses. Assume that a student who guesses at the answer will be correct with probability 1/5, where 5 is the number of multiple choice alternatives. What is the conditional probability that a student knew the answer to a question given that he answered it correctly?

(d) Show that for two continuous random variables X and Y, $E(X+Y) = E(X) + E(Y)$; provided the expectations exist.

(e) The joint probability distribution of two random variables X and Y is given by

$$f(x, y) = 4xye^{-(x^2+y^2)}, \; x \geq 0, \; y \geq 0$$

Find the marginal distributions and check whether X and Y are independent.

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

(a) If $A_1, A_2, \cdots, A_n$ are n events, prove that

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum\sum_{1 \leq i < j \leq n} P(A_i \cap A_j)$$

$$+ \sum\sum\sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) - \cdots$$

$$+ (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

What will happen to this relation if all the events $A_1, A_2, \cdots, A_n$ are mutually disjoint? 9+1=10

(b) *(i)* Define pairwise independence and mutually independence of events.

A balanced die is tossed twice. Let A_1 be the event that an even number comes on the first toss, A_2 is the event that an even number comes in the second toss and A_3 is the event that the same even number comes in both the tosses. Examine whether A_1, A_2 and A_3 are pairwise and mutually independent or not. 1½+1½+4=7

(ii) If A and B are two mutually exclusive events and $P(A \cup B) > 0$, then show that

$$P(A / A \cup B) = \frac{P(A)}{P(A) + P(B)}$$ 3

(c) (i) State Bayes' theorem. Explain 'a priori' and 'a posteriori' probabilities in the context of this theorem. 4

(ii) Suppose that event A can occur only along the event B which in turn can occur in n mutually exclusive ways $B_1, B_2, \cdots, B_n$. Show that

$$P(A) = \sum_{i=1}^{n} P(B_i)\, P(A / B_i)$$ 3

(iii) If n balls are placed at a random order into n cells, find the probability that exactly one cell remains empty. 3

(d) Let X be a continuous random variable with probability density function given by

$$f(x) = \begin{cases} ax &, \quad 0 \le x < 1 \\ a &, \quad 1 \le x < 2 \\ -ax + 3a, & \quad 2 \le x \le 3 \\ 0 &, \quad \text{elsewhere} \end{cases}$$

(i) Determine the constant a.

(ii) Determine $F(x)$.

(iii) Evaluate $P(\frac{1}{2} \le x \le \frac{3}{2})$.

(iv) Determine $E(x)$. 2+4+2+2=10

(e) (i) A coin is tossed until tail appears. What is the mathematical expectation of number of heads obtained? 4

(ii) Define conditional variance for discrete and continuous random variables.

For two discrete random variables X and Y, show that

$$V(X) = E[V(X / Y)] + V[E(X / Y)]$$ 2+4=6

(f) (i) Define moment-generating function (mgf). Show that the mgf of the sum of independent random variables is equal to the product of the mfg of the individual variables. 5

(ii) State the relation between the moments and cumulants. Are the cumulants independent of change of origin and scale of the variable? Explain. 5

Or

If $X_1, X_2, \cdots, X_n$ are independent random variables each assuming the values $0, 1, 2, \cdots, a-1$ with probability $\frac{1}{a}$, then find the probability-generating function of the sum $S_n = X_1 + X_2 + \cdots + X_n$. 5

★ ★ ★