STAT C-101
DESCRIPTIVE STATISTICS

Descriptive Statistics Question Bank

1. In a frequency table, the upper boundary of each class interval has a constant ratio to the lower boundary. Show that the geometric mean G may be expressed by the following formula:

$$LogG = x_0 + \frac{c}{N}\sum_i f_i\,(i-1)$$

where x_0 is the logarithm of the mid value of the first interval and c is the logarithm of the ratio between upper and lower boundaries.

2. Prove that the sum of the squares of the deviations of a set of observations is minimum when taken about mean.

3. A variate takes values $a, ar, ar^2, \dots, ar^{n-1}$ each with frequency unity. Compute arithmetic mean (A), geometric mean (G) and harmonic mean (H) and show that $AH = G^2$.

4. If the deviations $X_i = x_i - M$ are small compared with the value of the mean M so that $(X_i/M)^3$ and higher powers are neglected, prove that

$$G = M\left(1 - \frac{1}{2}\frac{\sigma^2}{M^2}\right),$$

hence prove that

$$V = \sqrt{\frac{2(M-G)}{M}},$$

where G is the geometric mean of the values $x_1, x_2, \dots, x_n$ and V is the coefficient of dispersion.

5. Show that in a discrete series if deviations $x_i = X_i - M$, are small compared with the value of the mean M so that $(x/M)^3$ and higher powers of (x/M) are neglected,

(i) $H = M\left(1 - \frac{\sigma^2}{M^2}\right)$

(ii) $Mean\left(\frac{1}{\sqrt{X}}\right) = \frac{1}{\sqrt{M}}\left(1 + \frac{3\sigma^2}{8M^2}\right)$ approx.

where H is the harmonic mean of the values $x_1, x_2, \dots, x_n$ and σ^2 is the variance.

6. Find the mean deviation about mean and standard deviation of A.P. a, a+d, a+2d, ..., a+2nd and verify that standard deviation is greater than mean deviation about mean.

7. Let r be the range and s be the standard deviation of a set of observations $x_1, x_2, \dots x_n$. Prove that $s \leq r$. Also prove that $S \leq r\left(\frac{n}{n-1}\right)^{\frac{1}{2}}$ where

$$S = \left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{\frac{1}{2}}$$

8. If n_1, n_2 are the sizes $\overline{x_1}$, $\overline{x_2}$ the means and σ_1, σ_2 the standard deviations of two series, then obtain the mean and variance of the combined series of size $n_1 + n_2$.
9. Find mean square deviation and variance if the variable takes values 0, 1, 2, ..., n with frequencies given by the terms of binomial series $\binom{n}{0}$, $\binom{n}{1}$, ..., $\binom{n}{n}$ respectively.
10. For a random variable x, central moments of all order exist, then show that

$$\mu_j^2 \leq \mu_{j-1}.\mu_{j+1},$$

where μ_j is the j^{th} central moment.

11. If ∂_r is the rth absolute moment about zero, then the mean value of $\{u|x|^{(r-1)/2} + v|x|^{(r+1)/2}\}^2$.
Show that (i) $\partial_r^{r+1} \leq \partial_{r+1}^{r}$, and
(ii) $\partial_r^{1/r} \leq \partial_{r+1}^{1/(r+1)}$.
12. Establish the relationship between the moments about mean, in terms of moments about any arbitrary point A. What are the Sheppard's corrections to central moments.

13. Discuss the principle of least squares. Derive the normal equations for fitting the curve Y=a*exp(bX + cX^2) to the given set of n points $\{(x_i, y_i), i = 1,2 \ldots\ldots n\}$.

14. Two variables X and Y are known to be related to each other by the relation $Y = \frac{X}{aX+b}$. Derive the normal equations for fitting the given curve and estimate the constants 'a' and 'b'from a given set of n points $\{(x_i, y_i), i = 1,2 \ldots\ldots n\}$.

15. Discuss the principle of least squares. Derive the normal equations for fitting the curve Y=a*exp(cX^2) to the given set of n points $\{(x_i, y_i), i = 1,2 \ldots\ldots n\}$

16. If X and Y are independent random variables, show that

$$r(X + Y, X - Y) = r^2(X, X + Y) - r^2(Y, X + Y)$$

where, r(X + Y, X - Y) is the correlation coefficient between X + Y and X - Y, r(X, X + Y) is the correlation coefficient between X and X + Y and r(Y, X + Y) is the correlation coefficient between Y and X + Y.

17. Define Spearman's rank correlation coefficient. Let x_1, x_2, …, x_n be the ranks of n individuals according to a character A and y_1, y_2, …, y_n be the corresponding ranks of the individuals according to another character B. Obtain the rank correlation coefficient between them.

18. X and Y ate two random variables with variances σ_x^2 and σ_y^2 respectively and r is the coefficient of correlation between them. If $U = X + kY$ and $V = X + \frac{\sigma_x}{\sigma_y} Y$, find the value of k so that U and V are uncorrelated.

19. Given that $Y = kX + 4$ and $X = 4Y + 5$ are the lines of regression of Y on X and X on Y respectively, show that $0 \leq k \leq 1/4$. If $k = \frac{1}{16}$, find mean of two variables and the coefficient of correlation between them.

20. Show that $1 - R_{1.23}^2 = (1 - r_{12}^2)(1 - r_{13.2}^2)$

Also deduce that (i) $R_{1.23} \geq r_{12}$

(ii) $R_{1.23}^2 = r_{12}^2 + r_{13}^2, if\ r_{23} = 0$

(iii) $1 - R_{1.23}^2 = \frac{(1-\rho)(1+2\rho)}{(1+\rho)}$, provided all coefficients of zero order are equal to ρ.

21. X and Y are two random variables with same mean and the two regression equations are $Y = aX + b$ and $X = \alpha Y + \beta$. Show that $\frac{b}{\beta} = \frac{1-a}{1-\alpha}$. Also find the common mean.

22. If r_{12} and r_{13} are given, show that r_{23} must lie in the range:

$$r_{12}r_{13} \pm (1 - {r_{12}}^2 - {r_{13}}^2 + {r_{12}}^2\,{r_{13}}^2)^{1/2}$$

If $r_{12} = k$ and $r_{13} = -k$, show that r_{23} will lie between -1 and $1 - 2k^2$.

23. What do you mean by independence of attributes? Give a criterion of independence of attributes. If $\delta = (AB) - (AB)_0$, where $(AB)_0$ is the value of (AB) under the hypothesis that attributes A and B are independent, prove that

$$[(A) - (\alpha)]\,[(B) - (\beta)] + 2N\delta = (AB)^2 + (\alpha\beta)^2 - (A\beta)^2 - (\alpha B)^2$$

24. If $\frac{(A)}{N} = x$, $\frac{(B)}{N} = 2x$, $\frac{(C)}{N} = 3x$ and $\frac{(AB)}{N} = \frac{(BC)}{N} = \frac{(CA)}{N} = y$,
then, using the conditions of consistency of attributes show that
$0 < y \leq x \leq ¼$.

25. Given that $(A) = (\alpha) = (B) = (\beta) = (C) = (\gamma) = N/2$ and $(ABC) = (\alpha\beta\gamma)$, then show that

$$2(ABC) = (AB) + (AC) + (BC) - N/2.$$

26. Show that for n attributes $A_1, A_2, \ldots, A_n$

$$(A_1 A_2 A_3 \ldots A_n) \geq (A_1) + (A_2) + (A_3) + \ldots + (A_n) - (n-1)N$$

where N is the total number of observations.

27. Let $A_1, A_2, \ldots, A_n$ be the events in the domain of probability function P, such that $P\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} P[A_i]$. Using this relationship, prove that:

(i) $P\left[\bigcap_{i=1}^{n} A_i\right] \geq 1 - \sum_{i=1}^{n} P[\overline{A}_i]$, and

(ii) $P\left[\bigcap_{i=1}^{n} A_i\right] \geq \sum_{i=1}^{n} P[A_i] - (n-1)$.

28. Three newspapers A, B, and C are published in a certain city. It is estimated from the survey that of the adult population 20% read A, 16% read B, 14% read C, 8% read both A and B, 5% read both A and C, 4% read both B and C, 2% read all three. Find the percentage that read at least one of the papers? Also find the percentage that read both A and B but does both read C?

29. A sportsman's chance of shooting an animal at a distance r ($>a$) is a^2/r^2. He fires when $r = 2a$, and if he misses he reloads and fires when $r = 3a, 4a, \ldots$. If he misses at distance na, the animal escapes. Find the probability that the animal escapes.

30. It is 8: 5 against the wife who is 40 years old living till she is 70 and 4: 3 against her husband now 50 living till he is 80. Find the probability that (i) both will be alive, (ii) only one will be alive and (iii) at least one will be alive 30 years hence.

31. A and B alternately cut a pack of cards and the pack is shuffled after each cut. If A starts and the game is continued until one cuts a diamond, what are the respective chances of A and B first cutting the diamond?

32. In a class of 75 students, 15 were considered to be very intelligent, 45 as medium and the rest below average. The probability that a very intelligent student fails in a viva examination is 0.005; the medium student failing has a probability 0.05; and the corresponding probability for a below average student is 0.15. If a student is known to have passed the viva-voce examination, what is the probability that he is below average?

33. A problem in Statistics is given to three students A, B and C, whose chance of solving it are ½, ¾ and ¼ respectively. What is the probability that the problem will be solved if all of them try independently.

34. State Bayes' theorem.

In answering a multiple choice test, an examinee either knows the answer or he guesses or he copies. Suppose each question has four choices. Let the probability that examinee copies the answer is 1/6 and the probability that he guesses is 1/3. The probability that his answer is correct given that he copied the answer is 1/8. Suppose an examinee answers a question correctly, what is the probability that he really knows the answer?

STAT -C102
CALCULUS

Question Bank

Paper Name: Calculus

B.Sc. (Hon.) – 1st year (1st Sem.)

Set-A

1. (a) Discuss the continuity of the function

$$f(x) = \begin{cases} 1, & x \le 0 \\ 3 - x, & 0 < x \le 1 \\ \dfrac{4}{x+1}, & 1 < x \end{cases}$$

at $x = 0$ and $x = 1$.

(b) Verify that $\dfrac{\partial^2 z}{\partial x \partial y} = \dfrac{\partial^2 z}{\partial y \partial x}$ for $z = x^2 \tan^{-1} \dfrac{y}{x}$.

(c) If $p^2 = a^2 \cos^2\theta + b^2 \sin^2 \theta$, prove that

$$p + \frac{d^2p}{d\theta^2} = \frac{a^2b^2}{p^3}$$

2. (a) Find the n^{th} derivative of $y = \sin^{-1} \dfrac{2x}{x^2+1}$.

(b) Prove that if $y^3 - 3ax^2 + x^3 = 0$ then $\dfrac{d^2y}{dx^2} + \dfrac{2a^2x^2}{y^5} = 0$.

(c) Find the maxima and minima of the function

$$f(t) = \sin t + \frac{1}{2}\sin 2t + \frac{1}{3}\sin 3t \quad \forall\, t \in [0, \pi]$$

3. (a) Find the asymptotes of the curve

$(y-a)^2(x^2-a^2) = x^4 + a^4.$

(b) Find the position and nature of the multiple points on the follow

$$(y^2 - a^2)^3 + x^4(2x + 3a)^2 = 0.$$

(c) Find the points of inflexion of the curve
$x = a \tan t,\ y = a \sin t \cos t.$

4. (a) Find the asymptotes of the curve
$r \cos 2\theta = a \sin 3\theta.$

(b) Trace the curve
$y^2(x + a) - x^2(3a - x) = 0.$

SECTION – II

5. Solve the following differential equations :

(a) $\dfrac{dy}{dx} + \dfrac{x-2y+5}{2x+y-1} = 0$

(b) $(x^3y^2 + xy)dx = dy$

(c) $x^3\dfrac{dy}{dx} - x^2y + y^4 \cos x = 0$

6. Solve the following differential equations :

(a) $p^3y^2 - 2xp + y = 0$

(b) $4p^3 + 3xp = y$

(c) $4y = x^2 + p^2$

7. Solve the following differential equations :

(a) $(D^2 - 2D + 1)y = (1 + e^{-x})^2$

(b) $(D^3 + D^2 + D + 1)y = x^5 - 2x^2 + x$

(c) $(D^4 - 1)y = x^2 \sin x$

Set- B

1. Attempt any **five** parts :

(a) Show that $\lim_{x \to 0} \frac{e^{\frac{1}{x}} - 1}{e^{\frac{1}{x}} + 1}$ does not exist.

(b) Evaluate $\lim_{x \to 0} \frac{x \cos x - \log(1+x)}{x^2}$.

(c) Prove that $\int_0^\infty \frac{x^c}{c^x} dx = \frac{\Gamma(c+1)}{(\log c)^{c+1}}$, $c > 1$.

(d) Evaluate $\int_0^1 \int_0^1 \frac{1}{\sqrt{(1-x^2)(1-y^2)}} dx\, dy$.

(e) Solve $(1 + y^2)dx = (\tan^{-1} y - x)dy$.

(f) $(D^2 - 3D + 2)y = 3\sin x$.

(g) Find partial differential equation of all planes a distance of a units from origin.

(h) Solve partial differential equation $(y - z)p + (z - x)q = x - y$. (5×3)

SECTION – I

2. (a) Determine the minimum value of $x^2 + y^2 + z^2$ subject to the condition

$$x + 2y - 4z = 5.$$

(b) If $y = \cos(m(\sin^{-1} x))$, show that $(1 - x^2)y_{n+2} - (2n + 1)xy_{n+1} + (m^2 - n^2)y_n = 0$ and hence evaluate $y_n(0)$. (6,6)

3. (a) If A, B and C are the angles of a triangle such that

$\sin^2 A + \sin^2 B + \sin^2 C = \sqrt{3}$, prove that $\frac{dA}{dB} = \frac{\tan B - \tan C}{\tan C - \tan A}$.

(b) Find the position and nature of the double points on the curve

$$(y - 2)^2 = x(x - 1)^2.$$ (6,6)

SECTION – II

4. (a) Prove that $\int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{a\cos^4\theta + b\sin^4\theta}} = \frac{\Gamma^2\left(\frac{1}{4}\right)}{4\sqrt{\pi}(ab)^{\frac{1}{4}}}$.

(b) Assuming the validity of differentiation under integral sign, prove that

$$\int_0^\infty e^{-x^2} \cos \alpha x \, dx = \frac{\sqrt{\pi}}{2} e^{-\frac{1}{4}\alpha^2}. \qquad (6,6)$$

5. (a) Find the limit, when n trends to infinity, of the sum : $\sum_{r=1}^{n-1} \frac{1}{n}\sqrt{\frac{n+r}{n-r}}$.

(b) Change the order of integration in $\int_0^{3a}\int_{x^2/4a}^{3a-x} F(x,y)\,dy\,dx$ and hence evaluate when $F(x,y) = x + y$. (6,6)

SECTION – III

6. Solve the following differential equations :

(i) $(x^2y - 2xy^2)dx - (x^3 - 3x^2y)dy = 0$

(ii) $\left(1-x^2\right)\frac{dy}{dx}+2xy = x\left(1-x^2\right)^{\frac{1}{2}}$ (6,6)

7. Solve any two of the following differential equations :

(i) $x^2\frac{d^2y}{dx^2} + 7x\frac{dy}{dx}+13y = \log x$

(ii) $\left(1+2x\right)^2\frac{d^2y}{dx^2}-6\left(1+2x\right)\frac{dy}{dx}+16y=8\left(1+2x\right)^2$

(iii) $(D^2 + 2D + 1)y = 2x + x^2$ (6,6)

SECTION – IV

8. Solve any two of the following partial differential equations :

   (i) $(x^2 - y^2 - z^2)p + 2xyq = 2xz$

   (ii) $x^2p^2 + y^2q^2 = z^2$

   (iii) $(D^2 + DD' - 6D'^2)z = y \cos x$

9. (a) Solve $p^2 + q^2 = 1$ using variable separation method.

   (b) Solve the partial differential equation $\frac{\partial^2 z}{\partial x^2} + x^2 \frac{\partial^2 z}{\partial y^2} = 0$.

GENERIC ELECTIVE: 1
STATISTICAL METHODS

Question Bank for GE 1 – Statistical Methods

1. The mean and standard deviation of 20 items is found to be 10 and 2 respectively. At the time of checking it was found that one item 8 was incorrect. Calculate the mean and standard deviation if (i) the wrong item is omitted, and (ii) it is replaced by 12.
2. A motor when travelling from rest travels the first twentieth of a mile at 6 miles per hour and the next three twentieths at respectively 8, 12, 24 miles per hour. Use the most appropriate measure of central tendency to find the average speed of motor?
3. A person travels 900 miles by train at an average speed of 60 miles per hour; 300 miles by boat at an average of 25 miles per hour; 400 miles by plane at 350 miles per hour and finally 15 miles by taxi at 25 miles per hour. Find the average speed of the person for the entire distance using the most appropriate measure of central tendency.
4. For the following frequency distribution:

Class Interval	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70
Frequency	8	7	12	28	20	10	10

Represent the frequency distribution by suitable graph, compute mode and locate it graphically.

5. The mean weight of 150 students in a certain class is 60 kilograms. The mean weight of boys in the class is 70 kilograms and that of girls is 55 kilograms. Find the number of boys and number of girls in the class.
6. For a group of 50 male workers the mean and standard deviation of their daily wages are Rs 63 and Rs 9 respectively. For a group of 40 female workers these values are Rs 54 and Rs 6 respectively. Find the mean and variance of the combined group of 90 workers.
7. Mean monthly salary of 12 workers and 3 managers in a factory was Rs. 26000. When one of the managers whose salary was Rs. 27500 was replaced with a new manager, the mean salary of the team went down to Rs. 25000. What is the salary of the new manager?
8. Establish the relationship between the moments about mean, in terms of moments about any arbitrary point A.
9. The first three moments about point A = 5 are 2, 20 and 40. Determine mean and first three central moments.
10. Explain with the help of a suitable example the term "dispersion". Discuss different measures of dispersion in details.
11. Karl Pearson's coefficient of skewness of a distribution is 0.32, the standard deviation is 6 and mean is 29. Find the mode and hence median of the distribution.
12. Explain what are skewness and kurtosis. Also explain the methods of measuring skewness and kurtosis of a frequency distribution.

13. Obtain correlation coefficient r(X, Y) based on the following data:

$$n = 30, \sum_{i=1}^{30} x_i = 150, \sum_{i=1}^{30} y_i = 90, \sum_{i=1}^{30} y_i^{\,2} = 450, \sum_{i=1}^{30} x_i^2 = 1200,$$

$$\sum_{i=1}^{30} x_i y_i = 630.$$

Also interpret the result.

14. The coefficient of Spearman's rank correlation coefficient between microeconomics and statistics marks of 10 students was found to be 0.5. It was later discovered that the difference in ranks in two subjects obtained by one of the students was wrongly taken as 3 instead of 7. Find the correct value of coefficient of rank correlation coefficient.
15. Calculate the Spearman's rank correlation coefficient from the following data:

X	5	10	6	3	19	5	6	12	8	2	10	19
Y	8	3	2	9	12	3	17	18	22	12	17	20

16. If r_{12} and r_{13} are given, show that r_{23} must lie in the range:

$$r_{12}r_{13} \pm (1 - r_{12}{}^2 - r_{13}{}^2 + r_{12}{}^2 r_{13}{}^2)^{\frac{1}{2}}$$

If $r_{12} = k$ and $r_{13} = -k$, show that r_{23} will lie between -1 and $1 - 2k^2$.

17. If r_{12}=0.6, r_{13}= 0.7 and r_{23}= 0.8, then obtain $R_{3.12}^2$ and $r_{23.1}$. Interpret the results.
18. State the principle of least squares and use it to fit a curve of the form $Y = ae^{bX}$.
19. State the principle of least square and use it to obtain the curve of best fit of the form $Y = a + bX + cX^2$.
20. The following results were obtained in the analysis of data on X and Y.

	X	Y
Average	4.9	28
Variance	4	25

Correlation coefficient between X and Y is 0.8.

(i) Obtain the two lines of regression.
(ii) Estimate value of Y when X = 10.
(iii) Estimate the value of X when Y = 20.

21. From the following data:

X	15	20	25	30	35	40	45	50
Y	10	15	20	20	22	25	26	28

(i) Obtain the two regression coefficients.
(ii) Also obtain the two lines of regression.

22. In a survey of population of 12000, information is gathered regarding three attributes A, B and C. In the usual notations,

(A) = 980, (AB) = 450, (ABC)= 130, (B) = 1190, (AC) = 280, (C) = 600 and (BC) = 250

(i) Compute (αβγ), (Aβγ), (αBC) and (ABγ).
(ii) Obtain the coefficient of association between A and B. Interpret the result.

23. The following table gives the distribution of students and also of regular players among them, according to age completed in years:

Age in years	15	16	17	18	19	20
No. of students	250	200	150	120	100	80
Regular Players	200	150	90	48	30	12

Calculate the coefficient of association between majority and playing habit, on the assumption that majority is attained in 18^{th} year.

24. 1000 candidates of both sexes appeared at an examination. The girls outnumbered the boys by 20% of the total. The number of candidates, who passed exceed the number failed by 480. Equal number of girls and boys failed in the examination. Prepare a 2 X 2 table and find coefficient of association. Comment.

25. State the principle of least squares. Obtain the normal equations for fitting the exponential curve $y=ab^x$. Given the following data N = 10, $\sum U_i = 5.56$, $\sum X_i = 55$, $\sum U_i X_i = 39.83$ and $\sum X_i^2 = 385$ where $U_i = \log Y_i$. Obtain the estimates of A = log a and B = log b.