SHRIRAM KARGAONKAR'S

QUESTION BANK ON "*STATISTICAL METHODS– I*"

1. Data Condensation & Graphical Methods

Q.1 Define the following terms.

a) Raw data, b) Attributes, c) Variable

d) discrete variables e) continuous variables.

Q.2 Explain Less than and more than cumulative frequencies.

Q.3 Explain the construction of ogive curves.

Q.4 Explain the construction of Histogram.

Q.5 Explain briefly, construction of stem–leaf chart.

Q.6 draw a less than ogive curve and more than ogive curve for the following frequency distribution.

Marks	0-10	10-20	20-30	30-40	40-50
No. of students	5	12	43	32	8

Q.7 Draw a Histogram to represent the following data of the earnings of workers.

Monthly Earnings (in Rs.)	80-120	120-160	160-200	200-240	240-280	280-320
No. of workers	4	7	13	8	5	2

Q.8 Marks Scored by 50 students in a test paper are given below.

30 45 48 55 39 25 31 12 18 21 54 59 51 33 43 44 10 38 19 26 41 35 37 41 46 33 51 37 58 58 17 19 23 26 29 38 57 36 35 44 43 27 19 43 22 31 47 34 31 15.

Construct a stem and leaf chart for the above data.

SHRIRAM KARGAONKAR'S

QUESTION BANK ON "*STATISTICAL METHODS– I*"

2. Measures of central Tendency

Q.1 Explain the concept of Central Tendency of a Data Set. What are the objectives & requisites of good measures of central tendency?

Q.2 Write a note i) Arithmetic mean.

ii) Weighted arithmetic mean

iii) Median.

iv) Mode.

v) Quartiles.

vi) Trimmed mean.

vii) Combined Mean.

Q.3 Compare mean & median in the light of requisites and usefulness.

Q.4 Explain briefly, the relative merits and demerits of mean, median & mode.

Q.5 Explain the concept of Percentile Ranks. Discuss its utility with the help of an example.

Q.6 Explain briefly, construction of Whisker box plot.

Q.7 Find the arithmetic mean of for the following values: 5,7,3,8,6,4,5,6,5,6. Also find 10% trimmed mean.

Q.8 A student scored 50,54,55,60 marks in four subjects Maths, Economics, Geography and English. Assigning weights 3,3,2,1 respectively, find the weighted A.M. of the scores of the student.

Q.9 Calculate arithmetic Mean, Median, Mode, Lower Quartile and Upper Quartile for the following data.

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Freq	4	6	9	6	5	4	4	2

SHRIRAM KARGAONKAR'S

QUESTION BANK ON "*STATISTICAL METHODS– I*"

Q.10 Calculate the arithmetic mean for the following data using step deviation method.

Class	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
Freq	50	70	100	180	150	120	70	60

Q.11 The mean salary of 50 employees was calculated to be Rs. 680/- per month. Later it was found that salary of Mr. A was wrongly taken as Rs. 270/- instead of Rs.720/- What will be the correct mean salary.

Q.12 Find the combined mean for the following data:
Mean(X_1) = 210, Mean of (X_2) = 150, n_1 = 150, n_2 = 100

Q.13 The mean weight of 150 students in a class is 60 kg. The mean weight of boys is 70 kg. and the mean weight of girls is 55 kg. Find the number of boys and girls in the class.

Q.14 Mean weight of 98 students as calculate from a frequency distribution is found to be 50 kg. It is later discovered that the frequency of the class interval 30-40 was wrongly taken as 8 instead of 10. Calculate the correct A.M.

Q.15 Find the missing frequency from the following data given that average number of tablets required to cure a person is 20. Also calculate mode and median of the completed table.

No. of tablets	4-8	8-12	12-16	16-20	20-24	24-28	28-32	32-36	36-38
No. of persons cured	11	13	16	14	---	9	17	6	4

Q.16 Find the missing frequency from the following data given that median is 126.

Class	100-110	110-120	120-130	130-140	140-150
Freq	5	---	20	10	7

Q.17 Draw a less than ogive for the following data and find the number of students getting 35 or more marks. Locate the median graphically.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of students	10	17	26	30	33	25	12	9

QUESTION BANK ON "*STATISTICAL METHODS– I*"

Q.18 The arithmetic mean, mode and median of a group of 100 observations were calculated to be 30,37 & 32 respectively. It was later discovered that one observation was wrongly noted as 56 instead of 65. Find the correct values of A.M.,mode & median.

Q.19 Find the quartiles from the following frequency distribution using formula. Also locate quartiles using graphical method.

Monthly salary	1400-1600	1600-1800	1800-2000	2000-2200	2200-2400	2400-2600
Freq	12	30	55	40	35	28

Q20) Calculate Simple and Weighted Arithmetic Mean of Price of wheat using following data:

Price of Wheat/Kg (Rs.)	**16**	**17.5**	**20**	**22**
Quantity purchased in Kg	**10**	**5**	**4**	**3.5**

Which of the two is preferable?

3. Measures of Dispersion

Q.1 Explain the concept of dispersion. List the various measures of dispersion. What are requisites of a good measures of dispersion? Also State various measures of dispersion.

Q.2 What are absolute and relative measures of dispersion?

Q.3 Define i) Range, ii) Coefficient of range? What are the merits and limitations of range?

Q.4 Define Quartile deviation and coefficient of quartile deviation. State its merits and demerits.

Q.5 Define Variance and standard deviation. State various properties of Standard Deviation.

Q.6 Write a note on the following:

i) Combined standard deviation.

ii)Coefficient of Variation.

Q.7 Discuss the effect of change of origin and scale on variance.

Q.8 Find the range & coefficient of range of the following series which gives the monthly expenditure of students in rupees.

22 35 32 45 42 48 39

Q.9 For a distribution Q_1=23.41, Q_2=25.3, Q_3=27.63. Find quartile deviation and coefficient of quartile deviation.

Q.10 Calculate Range, Quartile Deviation and Standard Deviation for the following data.

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Freq	18	16	15	12	10	5	2	2

Q.11 Calculate the standard deviation of the following observations on a certain variable.

240.12 240.13 240.15 240.12 240.17

240.15 240.17 240.16 240.22 240.21

Q.12 The mean of two samples of sizes 50 & 100 respectively are 54.1 & 50.3 and the standard deviations are 8 & 7. obtain the standard deviation of sample of size 150. obtained by combining the two samples.

Q.13 Find the missing information from the following.

	GroupI	GroupII	GroupIII	Combined
Number	50	?	90	200
Standard deviation	6	7	?	7.746
Mean	113	?	115	116

Q.14 A shareholder research centre of India has given the following results.

Share	Average price	Standard deviation
A	18.2	5.4
B	22.5	4.5
C	24.0	6.0

Above figures are in Rupees. Which share in your opinion appears to be more consistent?

Q.15 A sample of 50 cars each of 2 makes X and Y is taken and average running life in years is recorded.

Life (No. of years)	No. of Cars	
	Make X	Make Y
0-5	8	6
5-10	12	10
10-15	17	20
15-20	10	12
20-25	3	2

i) Which of these two makes gives higher average life?
ii) Which of these two makes shows greater consistency in performance?

Q.16 Scores of two golfers for 12 rounds were as follows.

Golfer A: 74 75 78 78 72 77 79 78 81 76 72 72

Golfer B: 86 84 80 88 89 85 86 82 82 79 86 80

Find which golfer may be considered to be a more consistent player.

Q.17 For the two groups, following results were obtained.

GroupI: $\sum(x_1-5) = 8$, $\sum(x_1-5)^2 = 40$, $n_1=20$

GroupII: $\sum(x_2-8) = -10$, $\sum(x_2-8)^2 = 70$, $n_2=25$.

Find the mean and standard deviation of the 45 observations obtained by combining the two groups.

Q.18 The mean & standard deviation of 20 observations are 10 & 2 respectively. Later it was discovered that item 8 taken was incorrect. Calculate Arithmetic mean and standard deviation if: i) The wrong item is omitted. ii) The wrong item is replaced by 12.

4. Moments

Q.1 Define the raw and central moments of a frequency distribution. obtain the relation between the central moments of order r in terms of raw moments.

Q.2 Express first four central moments in terms of moments about origin. What is the effect of change of origin and scale on moments.

Q.3 Find the first four central moments of the frequency distribution given below.

Class:	**100-105**	**105-110**	**110-115**	**115-120**	**120-125**
Freq:	**7**	**13**	**25**	**25**	**30**

Q.4 The first two moments of a distribution about value 4 are 3 and 34. Find the mean and variance.

Q.5 The first three moments of a distribution about 2 are 1, 22, 10. Find its mean, standard deviation, and third central moment.

5. Measures of Skewness & Kurtosis

Q.1 Explain the term Skewness, using suitable diagrams.
Explain the Various types of skewness.

Q.2 Explain the concept of Kurtosis? What are the types of kurtosis? Also state its measure.

Q.3 Explain the following measures of skewness.

i) Karl Pearson's coefficient skewness.

ii) Bowley's coefficient of skewness.

iii) Pearsonian coefficient of skewness based on moments.

Q.4 Given the following information, calculate Karl Pearson's coefficient of skewness.
$\Sigma X = 452$, $\Sigma X^2 = 24270$, mode = 43.7, n=10.

Q.5 From the following data, calculate the Measure of Skewness using mean, median and S.D.

X	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Freq	18	30	40	55	38	20	16

Q.6 In a frequency distribution, the coefficient of skewness based on the quartiles is 0.6. If the sum of the upper and lower quartile is 100 and median is 38. Find the value of upper and lower quartile.

Q.7 Find the coefficient of variance of a frequency distribution given that its mean is 120, mode is 123 and S_{kp} is -0.3.

Q.8 For the following distribution calculate S_{kB}.

Sales (Less than)	20	30	40	50	60	70	80	90	100
No. of firms	20	225	465	500	634	644	650	665	680

Q.9 Variance of a mesokurtuic distribution is 4. find μ_4

Q.10 Given that $\beta_1 = 0.19, \beta_2 = 2.6$, $\mu_2 = 1.2$. find μ_3 , μ_4

Q.11 Compute Karl Pearson's Coefficient of Skewness and Bowley's Coefficient of Skewness for following distribution.

Daily Wages	70-90	90-110	110-130	130-150	150-170
No. of Workers	16	22	36	18	8

Q.12 The first four moments about '4' of a certain distribution are 1.5, 17, -30 and 308. find kurtosis and interprete.

SHRIRAM KARGAONKAR'S

QUESTION BANK ON "*STATISTICAL METHODS– I*"

6. Correlation

Q.1 Write a note on: **i) Bivariate data**

ii) Correlation

iii) Scatter diagram.

iv) Covariance

Q.2 Prove that covariance is invariant to the change of origin.

Q.3 prove that, cov (aX+b, cY+d) = ac cov(X,Y), where a,b,c,d are constants.

Q.3 if X, Y, Z are three variables, then Prove that: Cov(X+Y, Z) = Cov(X, Z) + Cov(Y, Z)

Q.4 Define Karl Pearson's coefficient of correlation. State its properties.

Q.5 Prove that correlation coefficient does not change in magnitude under the change of origin and scale.

Q. 7 Define coefficient of determination? State its use.

Q.8 Define Rank correlation coefficient. State its merits over Karl Pearson's correlation coefficient.

Q.9 Calculate product moment correlation coefficient between income and expenditure from the following data.

Year	1981	1982	1983	1984	1985	1986	1987	1988
Daily income	100	110	115	120	152	130	132	140
Daily Expenses	85	90	92	100	110	125	125	130

Q.10 Calculate Karl Pearson's correlation coefficient between advt. cost and sales from the following data. Also find rank correlation coefficient.

Advt. Cost	41	67	65	92	84	77	27	100	38	80
Sales in Lakh Rs.	46	52	57	85	61	67	59	90	50	83

Q.11 Given,

No. of pairs of x & y series	**=**	**15**
Arithmatic mean of x	**=**	**25**
Arithmatic mean of y	**=**	**18**
Standard deviation of x	**=**	**3**
Standard deviation of y	**=**	**3**
Sum of product of x and y	**=**	**6870**

Find correlation coefficient between x and y.

Q.11 Given that, X : 1 2 …… n

Y : 1 2 …… n

Show that Cov(X,Y) = $(n^2-1)/12$.

Q.12 Calculate correlation coefficient from the following information.

n = 5, $\Sigma x = 20$, $\Sigma x^2 = 90$, $\Sigma y = 20$, $\Sigma y^2 = 90$, $\Sigma xy = 73$.

7. Regression Analysis

Q.1 Explain the term "Regression".

Q.2 Derive the equation for regression line of: i) Y on X ii) X on Y.

Q.3 Define regression coefficients. How will you interpret the coefficient of regression. State their properties.

Q.4 Explain the effect of change of origin and scale on regression coefficient.

Q.5 Prove that correlation coefficient is a square root of product of regression coefficients.

Q.6 Prove that regression coefficients can be expressed in terms of correlation coefficient as

$b_{yx} = r\,(\sigma_y/\sigma_x)$ and $b_{xy} = r(\sigma_x/\sigma_y)$

Q.7 prove that both the regression coefficients cannot exceed unity simultaneously.

Q.8 If $r = \pm 1$, then prove that regression coefficients are reciprocals of each other.

Q.9 If $\sigma_y = \sigma_x$,then prove that regression coefficients are equal.

Q.10 Prove that regression coefficients and correlation coefficients have same algebraic sign.

Q.12 Explain the procedure of fitting

i) a second degree curve

ii) a curve of the type $y= ab^x$

Q.13 Obtain the regression lines from the following information.

$N=8$, $\sum(X-45) =-45$, $\sum(X-45)^2=4400$, $\sum(X-45)(Y-150)=280$,

$\sum(Y-150)^2=167432$, $\sum(X-45)(Y-150)=21680$.

Also find correlation coefficient between X & Y.

Q.14 The regression equations are:

$8X-10Y+66=0$, $40X-18Y=214$

The value of variance of X is 9. Find,

i) The mean values of X & Y.

ii) The correlation coefficient between X & Y.

iii) The standard deviation of Y.

Q.15 Given the following information.

	Variable X	Variable Y
Arithmatic mean	7.6	14.8
Standard deviation	3.6	25

Coefficient of correlation between X & Y is 0.8.

Find the linear regression estimate of X, given Y=10.

Q.16 If the two lines of regression are:

$9X+Y-\lambda = 0$ and $4X+Y-\mu = 0$

Also the means of X & Y are 2 & -3 respectively. Find the values of λ,μ and the coefficient of correlation between X & Y.

Q.17 Find the most likely price in Mumbai Corresponding to the price of Rs. 70 at Delhi from the following.

	Delhi	Mumbai
Avg. Price	65	67
Standard deviation	2.5	3.5

Correlation coefficient between the prices of commodities in the two cities is 0.8.

Q.18 Calculate the correlation coefficient from the following data.

$N=100, \sum X = 12500, \sum Y = 8000, \sum X^2 = 1585000, \quad \sum Y^2 = 648100, \sum XY = 1007425.$

Also obtain the regression of Y on X.

Q.19 Explain the procedure for fitting a Second degree curve.

Q.20 Explain the procedure for fitting an Exponential curve of the type:

$$Y = a.b^x$$

Q.21 Explain the procedure for fitting a curve of the type

$$Y = a.e^{bx}$$

Q.22 Fit an exponential curve of the type $Y = a.b^x$ to the data given below. Estimate sales for the year 1999.

Year(X)	1989	1990	1991	1992	1993	1994	1995
Sales(Y)	32	47	65	92	132	190	275

8. Multiple and Partial regression and Correlation

Q.1 Explain the concept of multiple regression.

Q.2 Explain Yule's notation.

Q.3 Derive an equation of plane of regression of X_1 on X_2 and X_3 using trivariate sample data.

Q.4 State the expression for coefficient of multiple correlation.

Q.5 state the properties of Multiple correlation coefficient.

Q.6 State the expression for coefficient of partial correlation.

Q.7 State the multiple correlation in terms of total and partial correlations.

Q.8 If all total correlation coefficients in a set of three variables are equal to ρ, then show that,

i) $R^2_{1.23} = 2\rho^2/(1+\rho)$

ii) $r_{1.23} = \rho/(1+\rho)$, $\rho \neq 0$.

Q.9 Let Y_1, Y_2, and Y_3 be the heights in cm of son, mother and father respectively. A sample on X_1, X_2,X_3 showed following results.

Mean (Y_1) = 170, $r_{12} = 0.28$ $\sigma_1 = 2.4$

Mean (Y_2) = 160, $r_{13} = 0.49$ $\sigma_2 = 2.7$

Mean (Y_3) = 168, $r_{23} = 0.51$ $\sigma_3 = 0.51$

i) **Obtain the equation of least squares regression plane of Y_1 on Y_2 and Y_3.**

ii) **Compute the partial correlation coefficient $r_{12.3}$.**

iii) **Calculate the multiple correlation coefficient $R_{1.23}$**

Q.10 Compute $r_{13.2}$ and $R_{1.23}$ given that,

$R_{12} = 0.7$, $r_{13} = r_{23} = 0.5$.

9. Time Series Analysis

Q.1 Define a time series. Mention its important components with illustrations.

Q.2 Explain the trend, Seasonal variation, cyclical variations, irregular variations by giving illustrations.

Q.3 Describe the additive model, multiplicative model used in time series analysis.

Q.4 Distinguish between seasonal variation and cyclical variations.

Q.5 Describe moving average method, least square method used for the estimation of trend.

Q.6 Discuss merits and demerits of moving average method and least square method used for estimation of trend.

Q.7 Explain how to fit Straight line trend, parabolic trend a, exponential trend by method of least square method.

Q.8 Write a short note on business cycles, seasons in time series.

Q.9 what do you understand by the seasonal variations in a time series? Explain (i) link relative method of computing the indices of seasonal variation (ii) Ratio to trend method of computing the indices of seasonal variations

Q.10 Estimate trend by using 5 yearly moving average, 4 yearly centered moving average for the following time series.

Year	Gross Capital Assets (in crores Rs.)	Year	Gross Capital Assets (in crores Rs.)
1976	19.3	1985	19.3
1977	20.9	1986	18.1
1978	17.8	1987	19.5
1979	16.1	1988	19.2
1980	17.6	1989	22.2
1981	17.8	1990	20.9
1982	18.3	1991	21.5

1983 1984	17.3 21.4	1992	21.9

Q.11 Estimate trend using 4 yearly centered moving average

Year	1988	89	90	91	92	93	94	95	96	97
Production (in tones)	78	73	71	73	75	78	73	77	70	69

Q.12 Compute 5 yearly moving average and estimate trend.

Year	1	2	3	4	5	6	7	8	9	10	11	12
National Income (in crores)	260	270	275	300	310	315	300	290	310	320	335	380

Q.13 Compute 4 yearly centered moving average, 5 yearly moving average for the following data

Year	1977	78	79	80	81	82	83	84
Annual Sales (in lakhs)	3.6	4.3	4.3	3.4	4.4	5.4	3.4	2.4

Q.14 Fit a straight line trend to following data

Year	1989	90	91	92	93	94	95	96
Profit in 10,000 Rs.	90	100	102	93	104	109	102	114

Q.15 Estimate trend using parabolic trend to the following data.

Year	1992	93	94	95	96
Sales in 10,000 Rs.	20	22	23	20	18

Q.16 Fit $y = ab^x$ and estimate trend to the following time series.

Year	1	2	3	4	5	6	7
Expenditure	177.2	185	224.9	254	304.9	359.9	438.8

Q.17 Using ratio to trend method, determine the quarterly seasonal indices.

Years / Quarter	I	II	III	IV
1	65	60	61	63
2	70	58	56	60
3	68	63	68	67
4	65	59	56	62
5	60	55	51	58

Q.18 Compute the seasonal indices from the following time series data on production by link relative method.

Year / Month	2003	2004	2005	2006
Jan	226.7	194.7	185.2	221.1
Feb	208.1	176.2	175.1	223.2
March	237.1	201.7	202.8	267.7
Apr	243.3	201.1	203.2	259.0
May	248.3	197.4	205.8	261.5
June	228.4	191.1	190.5	259.3
July	212.3	174.9	177.9	243.1
Aug	217.1	182.4	202.9	275.3
Sept	222.7	189.6	213.3	265.6
Oct	235.5	218.1	236.9	292.2
Nov	222.3	211.6	236.1	291.5
Dec	218.4	206.0	225.4	294.8

10. Statistical Quality Control

Q.1 What do you understand by statistical quality control? Discuss briefly its need & utility in industry.

Q.2 What is control chart? Explain the basic principles underlying the control charts.

Q.3 Explain the construction of control charts for mean & range.

Q.4 Explain in detail X-bar & R-chart.

Q.5 Construct a control chart for mean and the range for the following data on the basis of fuses, samples of 5 being taken every hour (each set of 5 has been arranged in ascending order of magnitude.) Comment on whether the production seems to be under control, assuming that these are the first data:

42	**42**	**19**	**36**	**42**	**51**	**60**	**18**	**15**	**69**	**64**	**61**
65	**45**	**24**	**54**	**51**	**74**	**60**	**20**	**30**	**109**	**90**	**78**
75	**68**	**80**	**69**	**57**	**75**	**72**	**27**	**39**	**113**	**93**	**94**
78	**72**	**81**	**77**	**59**	**78**	**95**	**42**	**62**	**118**	**109**	**109**
87	**90**	**81**	**84**	**78**	**132**	**138**	**60**	**84**	**153**	**112**	**136**