

QA Theory Question bank

1. Introduction to Statistics

1. Define "Statistics". Explain its uses and limitations. # Discuss its use in business and trade.
2. Write short note on: Pie chart and its advantages and disadvantages. #
3. What is diagrammatic representation of data? Explain its advantages. #
4. Justify or contradict – 'Charts or graphs are more effective in attracting attention than any other method of presenting data'. #
5. Write a short note on: Meaning and importance of Tabulation. #
6. Define and explain the following terms with an example:
Grouped data, class interval, class limits, class boundaries, class mark, inclusive and exclusive series, frequency and tally marks.

2. Data Collection and Sampling methods

1. What are the various methods of collecting statistical data? Which of these is most reliable and why?
2. Explain primary data and secondary data in detail.
3. Distinguish between primary data and secondary data. #
4. What precautions should be taken in the use of secondary data. #
5. Explain Census method. Its merits and demerits. #
6. Why are personal interviews preferred to questionnaires? Under what conditions may a questionnaire prove as a personal interview? #
7. What do you mean by a questionnaire? What is the difference between a questionnaire and a schedule? State the essential points to be remembered in drafting a questionnaire.
8. Explain sampling and purpose of sampling. #
9. Differentiate between Probability sampling and non-probability sampling. #
10. Explain Simple Random Sampling. #
11. What is Stratified sampling? Explain the merits and demerits of Stratified sampling. #

3. Introduction to Regression

1. Explain the following methods to check the performance of regression model: #
a) MAE
b) MAPE
2. Justify or contradict ' b_{xy} and b_{yx} , must be either positive or negative'. #

3. Explain regression and its types. Also explain regression analysis and discuss its applications. How does it differ from correlation.
4. Write a detailed note on least square regression.

4. Introduction to Multiple Linear Regression

1. What do you mean by Partial correlation coefficients? Explain in detail.
2. Write a short note on: Significance of Overall fit of regression model. #
3. What are assumptions of Multiple Linear Regression. #
4. Write short note on: Multiple Regression. #

5. Statistical inference

1. Explain Point Estimation with characteristics.
2. Explain the following point Estimation Properties with Example:
 - a) Consistency
 - b) Unbiasedness
3. Explain the method of maximum likelihood estimation.
4. Show that sample variance (S^2) is an unbiased estimator of population variance (σ^2). Also illustrate with an example. #
5. Write short note on: Method of moments. #
6. Define a random variable and its mathematical expectation.

6. Test of hypothesis

1. What is Hypothesis testing? Explain
 - a) Z-test for single Mean
 - b) Z-test for difference of mean
2. Explain in detail Neyman Pearson lemma.
3. Write a short note on: MP and UMP tests. #
4. Differentiate between Null Hypothesis and Alternative Hypothesis. #
5. Differentiate between Critical Region and Region of acceptance.
6. What are the test of skewness.
7. Explain Type I and Type II error in detail.

	1	2	3	4	5	6
2024 Dec	20	20	25	25	15	15
2024 May	15	15	25	20	10	35
2023 Dec	25	10	20	15	20	30
2023 May	20	15	25	25	25	15
2022 Dec	25	10	25	10	20	15
Last 4 Avg	20	15	25	20	20	15-25
*2022 May	15	15	15	15	15	15
Total	120	85	135	110	100	125

QA Numerical questions

1. Introduction to Statistics

1. Problems on tabulation:

1. Out of a total number of 10,000 candidates who applied for jobs in government department, 6854 were males, 3146 were graduates and other non-graduates. The number of graduates with some experience was 2623 of whom 1860 were males. The number of male graduates was 2012. The number of graduates with experience was 1093 that includes 323 females.
2. A survey of 370 students from Commerce Faculty and 130 students from Science Faculty revealed that 180 students were studying for only C.A. Examinations, 140 for only Costing Examinations and 80 for both C.A. and Costing Examinations. The rest had offered part-time Management Courses. Of those studying for Costing only, 13 were girls and 90 boys belonged to Commerce Faculty. Out of 80 studying for both C.A. and Costing, 72 were from Commerce Faculty amongst which 70 were boys. Amongst those who offered part-time Management Courses, 50 boys were from Science Faculty and 30 boys and 10 girls from Commerce Faculty. In all there were 110 boys in Science faculty. Present the above information in a tabular form. Find the number of students from Science Faculty studying for part-time Management Courses.
3. In a simple study about coffee habits in two towns A and B the following information is given
Town A: Females were 40%, total coffee drinkers were 45% and female non coffee drinkers were 20%.
Town B: Males were 55%, male non coffee drinkers were 30% and female coffee drinkers were 15%.
Present the data into a table format.

4. Represent the following data by a percentage sub-divided bar diagram.

Item of Expenditure	Family A	Family B
	Income Rs 500	Income Rs 300
Food	150	150
Clothing	125	60
Education	25	50
Miscellaneous	190	70
Saving or Deficits	+10	-30

5. Explain Bar chart with following example:

The following table shows the number of books of different subjects in a library.

Subject	Phy.	Chem.	Bio.	Hist.	Gio.	Eng.	Math.	Comp
No. of Books	100	125	75	75	50	200	250	175

6. The following is the distribution of total household expenditure of 202 workers in a city.

Expenditure (in RS)	Number of workers	Expenditure (in RS)	Number of workers
100-150	25	300-350	30
150-200	40	350-400	22
200-250	33	400-450	16
250-300	28	450-500	8

Draw a suitable diagram to comment on distribution of data.

#

7. The following Table gives the frequency distribution of the weekly wages(in '00RS.) of 100 workers in factory. Draw the Histogram and frequency polygon of the distribution.

Weekly wages ('00 RS)	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	Total
No. of Workers	4	5	12	23	31	10	8	5	2	100

8. The frequency distribution of scores obtained by 250 candidates in an entrance tests is as follows. Draw a less than and more than frequency curve (ogive) to represent the given data. Also, what is the significance of the point of intersection of the two ogive curves?

Scores	Number of candidates
400 – 450	25
450 – 500	30
500 – 550	45
550 – 600	37
600 – 650	30
650 – 700	33
700 – 750	15
750 – 800	35

2. Data Collection and Sampling methods

3. Introduction to Regression

1. The equation of two lines of regression obtained in correlation analysis are given below:
 $2X=8-3Y$ and $2Y=5-X$
Obtain the value of the correlation coefficient.
2. Equation of the two lines of regression are: $x+6y=6$ and $3x+2y=10$. Find:
 - a) mean of x and mean of y
 - b) regression coefficients b_{yx} and b_{xy} .
 - c) correlation coefficient between x and y . #
3. In a laboratory experiment on correlation research study, the equations to the two regression lines were found to be $2x-y+1=0$ and $3x-2y+7=0$. Find the mean of x and y . Also work out the values of regression.
4. For 100 students of a class, the regression equation of marks Statistics(X) and Economics(Y) is $3Y-5X+180=0$. The mean marks in Economics is 50, and variance of marks in Statistics is $\frac{4}{9}$ of the mean in Economics. Find the mean marks in Statistics and the coefficient of correlation between them.
5. From the following data:

X	40	34	28	30	44	38	31
Y	32	39	26	30	38	34	28

- Find: a) Coefficient of regression
b) Lines of regression
c) Coefficient of correlation

6. The following data are given regarding expenditure on advertising and sales of a particular firm:

	Advertising expenditure X	Sales(in Lakhs) Y
Mean	10	90
Standard Deviation	3	12

Coefficient of correlation, $r = 0.8$

- a) Calculate the regression equation of Y on X
- b) Estimate the advertisement expenditure required to attain a sales target of 120 lakhs.

7. From the data given below find:
- The Two regression coefficients
 - The Two regression equations
 - The coefficient of correlation between the marks in Economics and Statistics
 - The most likely marks in Statistics if marks in Economics are 30.

Marks in Economics	25	28	35	32	31	36	29	38	34	32
Marks in Statistics	43	46	49	41	36	32	31	30	33	39

8. The following table gives the age of cars of a certain make and annual maintenance costs. Obtain the regression equation for Maintenance costs, taking age of the car as the independent variable. Also find the maintenance cost for Age of the car = 5 years.

Age of cars (in Years)	Maintenance cost (In thousands of rupees)
2	10
4	20
6	25
8	30

9. A departmental store gives in-service training to its salesmen which is followed by a test. It is considering whether it should terminate the service of any-salesmen who does not do well in the test. The following data give the test scores and sales made by nine salesmen during the certain period:

Test scores	14	19	24	21	26	22	15	20	19
Sales ('000 Rs.)	31	36	48	37	50	45	33	41	39

Calculate the coefficient of correlation between the test scores and the sales. Does it indicate that the termination of services of low-test scores is justified? If the firm wants a minimum sales volume of Rs. 30000, what is the minimum test score that will ensure continuation of service? Also estimate the most probable sales volume of a salesman making a score of 28.

4. Introduction to Multiple Linear Regression

1. Problems on partial correlation coefficient:

1. Explain order of coefficient regression. In a certain tri variate distribution: $r_{12} = 0.7$, $r_{23} = r_{13} = 0.6$, find all the partial correlation coefficients.
2. Given $r_{12} = 0.7$, $r_{13} = 0.61$ and $r_{23} = 0.4$. Compute a) $r_{23.1}$, b) $r_{13.2}$, c) $r_{12.3}$.
3. In a certain tri variate distribution: $r_{12} = 0.7$, $r_{23} = r_{31} = 0.6$. Find the partial correlation coefficient $r_{12.3}$. #
4. In a certain tri variate distribution, the simple coefficients of correlation are as follows: If $r_{12} = 0.86$, $r_{13} = 0.65$, $r_{23} = 0.72$, calculate the coefficient of partial correlation $r_{12.3}$. #

2. Problems on fitting regression equation, finding R^2 , and testing significance of regression.

5. For the following data

- i) Fit a regression $y = a + b_1x_1 + b_2x_2$
- ii) Find the coefficient of multiple determination (R^2).
- iii) Also test the significance of regression (Given the appropriate Table value, $F = 13.274$, for a significance level of $\alpha = 0.01$)

Sales Territory	Sales in (Lakh Rs) Y	Advt in '000 (x1)	Number of selling agents (x2)
1	190	80	40
2	80	35	13
3	75	35	7
4	100	50	20
5	125	60	19
6	90	40	13
7	70	20	20
8	130	60	28

6. For the given data:

- a) Fit a regression $\hat{y} = a + b_1x_1 + b_2x_2$
- b) Find the coefficient of multiple determination (R^2)
- c) Also test the significance of regression (Given the appropriate table value $F = 13.274$, for a significance level of $\alpha = 0.01$)

Y	x1	x2
150	65	24
155	62	25
159	67	24
179	70	20
192	71	15
200	72	14
212	75	14
221	76	12

7. The data with regard to the cost of production of 8 different drugs and cost of ingredients and packaging cost, are as given below:

Sr No	cost of production (Rs.) (Y)	cost of ingredients (in thousands of Rs) (X ₁)	packaging cost(Rs.) (X ₂)
1	100	17	19
2	79	50	54
3	100	90	75
4	129	30	36
5	158	15	16
6	106	20	25
7	58	20	24
8	78	50	53

- Fit a regression $\hat{y} = a + b_1X_1 + b_2X_2$.
- Find the coefficient of multiple determination (R^2).
- Also test the significance of regression. (Given $F = 5.786$, for a significance level of $\alpha = 0.05$)

8. Following is the data about the weights in Kgs of 10 Shipments (X_1), the distances they were moved (X_2) and the damage that was incurred (Y).

Shipment	Damage (thousands of Rs) (Y)	weights in Kgs (X ₁)	Distance moved in Km (X ₂)
1	12	17	10
2	15	15	6
3	14	15	10
4	19	10	21
5	8	13	8
6	16	15	13
7	15	11	9
8	25	6	25
9	10	15	10
10	11	7	8

- Fit a regression $\hat{y} = a + b_1X_1 + b_2X_2$
- Find the coefficient of multiple determination (R^2).
- Also test the significance of regression (Given the appropriate table value, $F = 9.55$, for a significance level of $\alpha = 0.01$)

9. The data with regard to the output of gram and cost of seed and labour per hectare at eight farmers' fields, are given as below:

Sr. No.	Cost of produce (Y) (Rs./hectare)	Cost of Seed (X1) (Rs./hectare)	Cost of Labour (X2) (Rs./hectare)
1	190	50	10
2	50	30	10
3	300	150	15
4	100	50	15
5	150	40	20
6	90	40	10
7	300	100	35
8	120	60	14

- Fit a regression $\hat{y} = a + b_1X_1 + b_2X_2$.
- Find the coefficient of multiple determination (R^2).
- Also test the significance of regression (Given the appropriate table value, $F = 13.27$, for a significance level of $\alpha = 0.01$)

10. Perform a simple linear regression, determine slope and intercept.

X	1	2	3	3	4	5
Y	8	4	5	2	2	0

5. Statistical inference

6. Test of hypothesis

- The mean lifetime of a sample of 400 fluorescent light tube produced by a company is found to be 1570 hours with a standard deviation of 150 hours. Test of hypothesis that the mean lifetime of the bulb produced by the company is 1600 hours at 1% level of significance. (From table 1% level of significance = 2.33)
- A random sample of size 100 has a standard deviation of 5. What can you say about the maximum error with 95% confidence is 1.96. #
- The manufacturer of a certain make of electric bulbs claims that his bulbs have a mean life of 25 months with standard deviation of 5 months. A random sample of 6 such bulbs gave the following values:
Life of bulb in months: 24, 26, 30, 20, 20, 18
Is the manufacturer's claim valid at 1% level of significance? (Given that the table values of the appropriate test statistics at said level are 4.032, 3.707 and 3.499 for 5, 6 and 7 degree of freedom respectively) #

Other uncategorized questions:

A survey conducted over the last 25 years indicated that in 10 years the winter was mild, in 8 years it was cold and in the remaining 7 years it was very cold.
A company sells 1000 woollen coats in a mild year, 1300 in a cold year and 2000 in a very cold year. You are required to find the yearly expected profit of the company if a woollen coat costs Rs. 1730 and it is sold to stores for Rs. 2480. #

Find the Mean Deviation from the Median for the following graph:

Age of Workers	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
No. of Workers	120	125	175	160	150	140	100	30