

Statistics Question Bank

First Paper

Abdullah Al Mahmud



www.statmania.info

Contents

Chapter 1	Statistics, Variable and Concepts of Different Symbols	1
1.1	Creative Questions	1
1.2	Short Questions	2
Chapter 2	Collection, Presentation, and Organization of Data	3
2.1	Creative Questions	3
2.2	Short Questions	3
Chapter 3	Measures of Central Tendency	4
3.1	Creative Questions	4
3.2	Short Questions	7
3.2.1	General Questions	7
3.2.2	Arithmetic Mean	7
3.2.3	Geometric Mean	8
3.2.4	Harmonic Mean	8
3.2.5	Median	8
3.2.6	Mode	8
3.2.7	Quadratic Mean	9
3.2.8	Partition Values	9
Chapter 4	Measures of Dispersion	10
4.1	Creative Questions	10
4.2	Short Questions	10
Chapter 5	Moments, SKewness, and Kurtosis	11
5.1	Creative Questions	11
5.2	Short Questions	12
Chapter 6	Correlation and Regression	13
6.1	Creative Questions	13
6.2	Short Questions	13
Chapter 7	Time Series	14
7.1	Creative Questions	14
7.2	Short Questions	14
Chapter 8	Published Statistics in Bangladesh	15
8.1	Creative Questions	15
8.2	Short Questions	15
Conclusion		16

Chapter 1

Statistics, Variable and Concepts of Different Symbols

1.1 Creative Questions

1. Height (in inches) of 10 cadets in a class are: 50, 60, 55, 65, 66, 70, 54, 64, 62, 72

- (a) What is population in statistics? 1
- (b) Is height discrete or continuous? 2
- (c) Find $\sum_{i=1}^{10} x_i^2$ 3
- (d) Find the square of mean and mean of square. Are they equal? 4

2. Income and expenditure (both in thousands) of some individuals in four successive months are collected:

Income (x)	20	30	25	10
Expenditure (y)	15	27	18	5

- (a) What is a discrete variable? 1
- (b) Can fractional numbers be discrete? Explain briefly. 2
- (c) Are, in the stem, $\sum_{i=1}^n \sum_{j=1}^n x_i y_j = \sum_{i=1}^n x_i y_i$? Show statistically. 3
- (d) Prove empirically that sum of square is unequal to square of sum of numbers. 4

3. Call duration of 6 calls in a customer care center are

2, 2.5, 1.5, 5, 6, 3

- (a) What is a sample? 1
 - (b) Are all quantitative variables continuous? 2
 - (c) Determine $\sum_{i=1}^7 (x_i - 3)^3$ 3
 - (d) Find the values of $\sum_{i=1}^7 (x_i - 5)^2$ and $\sum_{i=1}^7 x_i^2 + 5$. 4
- Explain mathematically why they are unequal.

4. Goals scored by Karim Benzema in five seasons are recorded to be the following:

- (a) What is a quantitative variable? 1
- (b) What is the notation to denote his total number of goals? 2
- (c) Compute $\sum_{i=1}^5 (y_i - 3)^2$

Season	La Liga (x)	Uefa Champions League (y)
2017-18	5	5
2018-19	21	4
2019-20	21	5
2020-21	23	6
2021-22	27	15

- (d) Find total number of goals using two different notations and examine whether they match.

4

5. Below are some information

$$x_1 = 3, x_2 = 4, x_3 = 1, x_4 = 0$$

$$y_1 = 1, y_2 = 5, y_3 = 0, y_4 = 2$$

- (a) What is a qualitative variable?

1

- (b) Find $\sum_{i=1}^4 x_i^2$

2

- (c) Prove that $\sum_{i=1}^4 (x_i + y_i) = \sum_{i=1}^4 x_i + \sum_{i=1}^4 y_i$

3

- (d) Find the value of $\sum_{i=1}^4 x_i y_i - \sum_{i=1}^4 x_i + 4$

4

6. Income of 100 individuals in the city of Rajshahi were analyzed and found to produce the the following distribution:

Income	40-50	50-60	60-70	70-80	80-90
Number of Individuals	15	20	35	20	10

- (a) What is Median?

1

- (b) Does median necessarily lie in the dataset?

2

- (c) Estimate Median and explain the result.

3

- (d) Find Arithmetic Mean and Mode. Which measure seems to be the best one?

4

1.2 Short Questions

1. $x_1 = 2, x_2 = -3, x_3 = 7, x_4 = 12$.

Find the values of the following:

$$4 \times 1.5 = 6$$

i) $\sum_{i=1}^3 x_i$ ii) $\sum_{i=1}^4 x_i^2$ iii) $\sum_{i=1}^4 (x_i + 3)$ iv) $\sum_{i=1}^4 (x_i - 4)^2$

2. Write down the scales of measurement of the following variables.

$$8 \times 0.5 = 4$$

Gender, Religion, Temperature, Income group (Lower class, Low, Middle, High)

Income, Distance of stars, Radius of screws, Room no.

3. Distinguish between the qualitative and quantitative variable.

$$6 \times 0.5 = 3$$

Diameter of trees, Color, Weight, Gender, Jersey Number, Family Size

4. Distinguish between the discrete and continuous variable.

$$8 \times 0.5 = 4$$

Number of vote cast for a particular candidate, Time required to run 100 m, Years of schooling, Number of goals in a soccer match, Body temperature, Gravity of stars, Absolute humidity, Atomic Number

Chapter 2

Collection, Presentation, and Organization of Data

2.1 Creative Questions

2.2 Short Questions

Chapter 3

Measures of Central Tendency

3.1 Creative Questions

1. The arithmetic and geometric means of ages two boys Abir and Abid are 10 and 8.

- (a) What is arithmetic mean? 1
- (b) When can we not calculate arithmetic mean? 2
- (c) Determine the ages of Abir and Abid. 3
- (d) Does the data comply with the theorem $AM \times = GM^2$? 4

2. Amount of rainfall in some cities around the world for a month were obtained as follows:

Rainfall (mm)	Frequency
20-30	5
30-40	6
40-50	4
50-60	3
60-70	5

- (a) When is Short-Cut method for Arithmetic Mean useful? 1
 - (b) Derive the formula of Short-cut Method 2
 - (c) Compute the Arithmetic Mean using the Short-cut method. 3
 - (d) Compute the Arithmetic Mean with a different value of origin (A). Do both the methods give same result? 4
3. A sports analyst collected ages of athletes having ages between 10 and 35. He then presented his findings as below:

Age	10-15	15-20	20-25	25-30	30-35
No. of Athlete	2	8	10	5	3

- (a) What is central tendency? 1
 - (b) When is geometric mean appropriate to measure? 2
 - (c) Compute median from the stem. 3
 - (d) Show that Arithmetic mean is greater than Harmonic mean. Which one of them is more suitable for this data? 4
4. Mean monthly salaries of employees of two companies A & B are tk. 65,000 and tk. 75,000. The combined arithmetic mean (AM) is tk. 71,000 and number of employees in the company A is 20.
- (a) Write down the formula of combined AM for k groups. 1
 - (b) What is the combined AM of two data sets with AM 35 and 45 and number of values equal? 2
 - (c) How many employees are there in the company B? 3
 - (d) Salary of an employee of company A was recorded as tk. 60,000 in place of 65,000. What is the new AM of company A. Also find the corrected combined AM. 4
5. A departmental store records their sales. An analysis of products with prices less than tk. 30 generates the following table.

Price	0-5	5-10	10-15	15-20	20-25	25-30
Frequency	1	0	2	3	8	12

- (a) What is relative frequency? 1
- (b) If $Y = a + bX$, $\bar{Y} = ?$ 2
- (c) Find 67th Percentile and 3rd Quartile and explain. 3
- (d) Is AM or Median more suitable for this data? Elucidate. 4

6. **Arithmetic and Harmonic Mean (HM) of two numbers are 25 and 9, respectively.**

- (a) When is HM useful? 1
- (b) Derive HM formula using the concept of average velocity. 2
- (c) Find the two values from the stem. 3
- (d) Show mathematically that $HM \leq AM$ (for $n=2$) 4

7. **Frequency distribution of marks in statistics of a college is given in the following table.**

Marks	Number of Students	Number of Students
	Group - A	Group - B
25-30	11	10
30-35	18	16
35-40	21	22
40-45	26	28
45-50	14	9

- (a) What is data? 1
- (b) What are the disadvantages of secondary data? 2
- (c) Calculate the arithmetic mean of Group - A 3
- (d) Compute the combined mean. Is it greater than the arithmetic mean of Group - B? Explain the possible reason(s). 4

8. **In the test examination, marks of 11 students in statistics are: 90, 92, 93, 49, 44, 88, 80, 58, 83, 71, 76.**

- (a) What is central tendency? 1
- (b) When is median better than arithmetic mean? Explain with an example. 2
- (c) Find the 3rd the quartile and 61st percentile from the data and explain. 3
- (d) Do quantiles depend on change of origin and scale. Prove using two examples. 4

9. **Scores of a batsman in the last 20 innings are**

28, 30, 16, 48, 50, 86, 105, 20, 10, 36,
12, 25, 20, 35, 65, 12, 10, 76, 55, 32

- (a) Write down the formula of weighted harmonic mean 1
- (b) Can median be a better measure of central tendency than arithmetic mean for this data? 2
- (c) Draw a stem and leaf plot from the data and explain. 3
- (d) Make a frequency distribution from the data and also find and interpret cumulative frequencies and percentages. 4

10. **In ODI cricket, two top batsmen are (as of 2nd Sept, 2022) Babar Azam and Rassie van der Dussen. Their average (arithmetic mean) scores are 59.79 and 69.32, appearing in 90 (including being not out in 12 occasions) and 33 (including being not out in 11 occasions) matches, respectively.**

- (a) When is arithmetic mean inappropriate to use? 1
- (b) Is arithmetic mean always suitable for comparison? 2
- (c) Find the combined arithmetic mean and explain. 3
- (d) How to compare two sets of data having significantly distinct ranges? 4

11. **A fridge manufacturing company observe temperatures of newly developed 8 deep fridges. The observed temperatures (in degree celsius are:**

-10, -8, -2, -4, -4, -1, -12, -3, -13

- (a) What is a Decile? 1
- (b) How many Deciles does a data set have? Why? 2
- (c) Compute the 8th Decile from the data and explain. 3
- (d) Find and compare arithmetic and geometric mean from the data. 4

12. Given below is a series of data.

5, 7, 9, \dots , 123

- (a) What is the summation of natural numbers up to nth value? 1
- (b) Find the arithmetic mean of natural numbers from 1 up to 20. 2
- (c) Find the arithmetic mean of the given series. 3
- (d) Prove that arithmetic mean is greater than geometric mean theoretically and empirically. 4

13. Grades of a an undergraduate student with major in statistics are given

Course	Grade	Credit
Probability	3.75	4
Simulation	3.50	3
Calculus	3.50	4
Linear Algebra	3.75	4
Econometrics	3.00	2
Programming	3.50	3

- (a) Write down the formula of weighted mean. 1
- (b) What is difference between weight and frequency? 2
- (c) Determine the GPA of the student. 3
- (d) Determine the geometric mean for the data and evaluate suitability. 4

14. A student walks 3 hours at 5 km per hour (kph), 4 hours at 4 kph, and 2 hours at 3 kph

- (a) When is harmonic mean suitable? 1
- (b) Which means could we use for the given data and why? 2
- (c) Find the average speed using weighted harmonic mean. 3
- (d) Find the average speed using another method and mathematically show their relationship. 4

15. A cyclist moves around a square-shaped lake with the speeds 20, 25, 30, and 16 km per hour.

- (a) What is grouped data? 1
- (b) Is arithmetic mean suitable for this data? 2
- (c) Find the average speed of the cyclist. 3
- (d) Can we use some other formula for finding the average? Demonstrate. 4

Rainfall (mm)	Frequency
20-30	5
30-40	6
40-50	4
50-60	3
60-70	5

3.2 Short Questions

3.2.1 General Questions

1. What is the primary goal of central tendency?
2. When is WAM equal to WHM? Show mathematically.
3. When is the equality true: $AM = GM = HM$? Prove mathematically and empirically.
4. When is Median a better measure of central tendency than Arithmetic Mean?
5. When is Harmonic Mean more suitable than Arithmetic Mean?
6. Write two primary uses of central tendency.
7. For two distinct non-zero values, what is the relationship among AM, GM, and HM?
8. what is the relationship among AM, GM, and HM?
9. What are the criteria of a good measure of central tendency?
10. For two non-zero positive numbers, prove $AM \geq GM \geq HM$
11. Find the AM, GM, and HM: $1, 2, 4, \dots, 2^n$

3.2.2 Arithmetic Mean

12. Does Arithmetic Mean depend on origin and scale? Prove with an example.
13. Prove with an example: $\sum_{i=1}^n (x_i - \bar{x})$
14. Prove mathematically: $\sum_{i=1}^n (x_i - \bar{x})$
15. Derive the formula of Arithmetic Mean for short-cut method.
16. Derive the formula of combined Arithmetic Mean for n number of observations.
17. Find Arithmetic Mean of first n natural numbers.
18. If $u_i = x_i + y_i$, find \bar{x} in terms of u . 1
19. For two numbers, $AM = 25$ and $GM = 15$. $HM = ?$ 1
20. $\bar{X} = 25$ and $Y_i = 5X_i + 20$. $\bar{Y} = ?$
21. Find Arithmetic Mean: $11, 13, 15, \dots, 57$
22. Find Arithmetic Mean: $115, 120, 125, \dots, 225$
23. Arithmetic Mean (\bar{X}) of five numbers is 40, and of three of them is 30. What is \bar{X} of the rest two?
24. AM of 200 values is found to be 50. Later it was seen two values were recorded as 92 and 8 in place of 192 and 88, respectively. What is the correct AM?
25. AM of 8 values of 20. If the 9th value is 0, what is the new AM?
26. AM of Income of City A is 1500 and of City B is 1200. If a person moves from city A to city B, can AM of both cities decrease?
27. Calculate Arithmetic Mean:

Renumeration (X)	No. of Employees (f)
30	5
35	8
40	10

28. Calculate Arithmetic Mean using short-cut method:
29. What is the relationship between changing origin & scale and short-cut method of Arithmetic Mean?

Rainfall (mm)	Frequency
20-30	5
30-40	6
40-50	4
50-60	3
60-70	5

3.2.3 Geometric Mean

30. Write down the formula of Geometric Mean for grouped data.
31. Find Geometric Mean for these values: 2, 4, 8
32. Derive the formula of Geometric Mean using logarithm.
33. When is Geometric Mean not calculable?
34. When is Geometric Mean appropriate?
35. Determine the formula of combined Geometric Mean when $n_1 = n_2 = n$
36. $Y_i = 3X_i$. If $G_y = 9, G_x = ?$ [G stands for Geometric Mean]
37. $n_1 = 15, G_1 = 75, n_2 = 10, \text{ and } G_2 = 80$ Find combined GM.

1

3.2.4 Harmonic Mean

38. Write down the formula of Weighted Harmonic Mean.
39. When is Weighted Harmonic Mean used instead of Unweighted one?
40. Calculate Harmonic Mean: 10, 15, 20
41. Show mathematically that Harmonic Mean of velocity for a fixed distance is equal to average velocity.
42. Find the average speed:

Path	Distance (km)	Speed (km/h)
Path 1	3	8
Path 2	2	9
Path 3	2	2

3.2.5 Median

43. Does Median depend on origin and scale? Prove with an example?
44. Median score of 50 students is 70. What does it mean?
45. Write down the formula of Median for even number of observations.
46. Write down the formula of Median for grouped data.
47. Find median: 5, 10, 15, 20, 25, 30
48. Is Median affected by outliers?
49. Does Median depend on origin and scale? Prove.
50. What is the greatest disadvantage of Median?
51. Does Median lie in the data set from which it is calculated?

1

3.2.6 Mode

52. What is the formula of Mode for grouped data?
53. Find the mode: 2, 2, 3, 3, 4, 4, 2, 2, 7
54. What is an Unimodal dataset?

3.2.7 Quadratic Mean

- 55. What is the formula of Quadratic Mean for grouped and ungrouped data?
- 56. When is Quadratic Mean used?

3.2.8 Partition Values

- 57. If a data is divided into four parts, how many partition values are created?
- 58. Write down the formula of Median for even and odd number of observations.
- 59. Derive the general formula of Quartiles using the concept of Median?
- 60. Third Quartile of score of 50 students is 76. What does it mean?
- 61. Which Quartile is equal to Median?
- 62. Which Percentile is equal to 3rd Quartile?
- 63. Find all the Quartiles: 2, 1, 0, 5, -6 , 7, -4

Chapter 4

Measures of Dispersion

4.1 Creative Questions

1. Temperatures of two cold regions for five days are as below:

City A: 2, 1, -1, 0, 3

City B: 3, 0, -2, 2, 3

- | | |
|---|---|
| (a) What is standard deviation?? | 1 |
| (b) Is standard deviation of a set of negative values negative? Justify mathematically. | 2 |
| (c) Find Mean Deviation about mean of the values of city A. | 3 |
| (d) Which city has more consistent weather? Verify statistically. | 4 |

4.2 Short Questions

Chapter 5

Moments, SKewness, and Kurtosis

5.1 Creative Questions

1. **There has been an increase in average lifetime of people of Bangladesh. To get more insight on this, a research was conducted, in which ages of retired government employees were recorded. A sample of 10 people is given below:**

75, 62, 63, 72, 66, 76, 59, 77, 70, 79

- (a) What is the 2nd central moment? 1
 - (b) Why is the first central moment zero? 2
 - (c) Find the variance of the data. 3
 - (d) Are the data symmetric? Justify. 4
2. **The arithmetic and geometric means of the first and third quartiles of a distribution are 10 and 8, respectively. The second quartile is 10.**
- (a) What is the formula suggested by Pearson to find skewness? 1
 - (b) Which moments are useful in measuring central tendency and dispersion? 2
 - (c) Find skewness from the stem using a suitable formula. 3
 - (d) Which method of finding skewness do you think is the best and why? 4
3. **For a particular data set, Median = 120, Mode = 110, Standard Deviation = 4, and Coefficient of Variation (CV) = 3.2**
- (a) Why is CV used? 1
 - (b) Find arithmetic mean.. 2
 - (c) Find skewness according to Pearson's method (SK_P) 3
 - (d) Does (SK_P) convey the proper idea about the data as to the given information? Justify. 4
1. **US Dollar exchange (to taka) in Bangladesh since 1980 to 2005 (after each 5 years) were: 16, 31, 36, 40, 52, 64**
- (a) What are moments? 1
 - (b) Which moment is equal to the variance? Show mathematically. 2
 - (c) Find, from the stem, the first and second raw moments about 1. 3
 - (d) Find skewness and kurtosis of and explain. 4
2. **A farmer in Dinajpur district produces seasonal crops. First four moments around 9 of his daily earnings are computed as -1, 8, -16, and 25.**
- (a) What is the Box and Whisker plot? 1
 - (b) Can Box and Whisker plot suggest symmetry? 2
 - (c) Find the arithmetic mean and variance of the farmer's earnings. 3
 - (d) Do the earnings produce a symmetric data? Analyze. 4
3. **The first four moments about 3 of a distribution are -1, 5, -10, and 120.**

- (a) What are moments used for? 1
- (b) Can the second central moment be greater than the third central moment? 2
- (c) Find the second and third moments about arithmetic mean of the distribution. 3
- (d) Find skewness and kurtosis and comment on the values. 4

4. Marks obtained by a student in 7 subjects are

70, 66, 55, 45, 80, 30, 82

- (a) What is negative skewness? 1
- (b) Draw graphs of positive and negative skewness showing the locations of mean and median. 2
- (c) Determine the five number summary from the stem and explain. 3
- (d) Are the data symmetric? If not, comment on the pattern of data. 4

5. United Nations Children's Fund (UNICEF) is an agency of the United Nations responsible for providing humanitarian and developmental aid to children worldwide. A UNICEF researcher collected heights of 7 children for a project, and the heights are

2.2, 2.15, 1.9, 3.1, 2.7, 3.0, 3.5

- (a) Which value are central moments estimated around? 1
- (b) Moments around origin (0) are central moments - Comment. 2
- (c) Find the first central moment. 3
- (d) Find the skewness of the data and interpret. 4

6. A researcher wants to compare average life time of people in Bangladesh and other countries. He collected life time of 10 people in Bangladesh.

75, 62, 63, 72, 66, 76, 59, 77, 70, 79

- (a) What is symmetry? 1
- (b) Mathematically show the theoretical value of the first central moment. 2
- (c) Compute the 2nd, 3rd, and 4th central moments of the data. 3
- (d) Estimate skewness and kurtosis and explain. 4

5.2 Short Questions

Chapter 6

Correlation and Regression

6.1 Creative Questions

6.2 Short Questions

Chapter 7

Time Series

7.1 Creative Questions

1. The yearly revenue (in hundred thousand) of shoe manufacturer company is given below

Year	2005	2006	2007	2008	2009	2010	2011
Revenue	35	22	40	35	50	42	60

- (a) What is general trend? 1
 - (b) Which method of determining trend gives only two values? 2
 - (c) Determine the trend using three-yearly moving average method. 3
 - (d) Find the trend using graphical method and extrapolate the approximate revenue earned in 2012. 4
2. **GDP (in bn. US\$ PPP) of Bangladesh since 1980 to 1985 according to an estimate of International Monetary Fund: 41.2, 47.4, 52.0, 56.5, 61.0, 65.3**
- (a) What is time series data? 1
 - (b) What are the components of a time series model? 2
 - (c) Determine the 3-yearly moving average from the data. 3
 - (d) Find trend of the data using another method (other than (c)), plot both, and comment which is better. 4

3. Annual sales of company are as given in the following

Year	2010	2011	2012	2013	2014	2015	2016
Profit (million)	40	45	46	53	65	70	73

- (a) What is a trend? 1
 - (b) Do the data in the stem seem to have a trend? 2
 - (c) Find the trend using semi-average method. 3
 - (d) Find the trend using 2-yearly moving average method. Would it better if we used 3-yearly method? 4
4. **Income of a freelancer in 6 successive months (from Jan to Jun) was found to be 46.0, 49.5, 51.5, 50.6, 56.5, and 60 (in thousands BDT.).**
- (a) What is time series data? 1
 - (b) What are the components of a time series model? 2
 - (c) Determine the 3-monthly moving average from the data. 3
 - (d) Draw the moving averages on a graph paper and interpret. 4

7.2 Short Questions

Chapter 8

Published Statistics in Bangladesh

8.1 Creative Questions

1. In 2015, tens of thousands of Rohingya people were forcibly displaced from their villages and IDP camps in Rakhine state, Myanmar. Many of them fled to neighboring countries, including Bangladesh, Malaysia, Indonesia. Many national and international agencies collect data on the issue.
 - (a) What is non-official statistics? 1
 - (b) Name five sources of official statistics. 2
 - (c) Shed some light on the limitations of official statistics. 3
 - (d) How can the quality of published statistics in Bangladesh be improved? 4

8.2 Short Questions

Conclusion

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetur at, consectetur sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Contents