

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 | 3 |
| Chapter 2 | Collection, Presentation, and Organization of Data | 5 |
| 2.1 | Creative Questions | 5 |
| 2.2 | Short Questions | 6 |
| Chapter 3 | Measures of Central Tendency | 7 |
| 3.1 | Creative Questions | 7 |
| 3.2 | Short Questions | 11 |
| 3.2.1 | General Questions | 11 |
| 3.2.2 | Arithmetic Mean | 12 |
| 3.2.3 | Geometric Mean | 12 |
| 3.2.4 | Harmonic Mean | 13 |
| 3.2.5 | Median | 13 |
| 3.2.6 | Mode | 13 |
| 3.2.7 | Quadratic Mean | 13 |
| 3.2.8 | Partition Values | 13 |
| Chapter 4 | Measures of Dispersion | 14 |
| 4.1 | Creative Questions | 14 |
| 4.2 | Short Questions | 14 |
| Chapter 5 | Moments, SKewness, and Kurtosis | 15 |
| 5.1 | Creative Questions | 15 |
| 5.2 | Short Questions | 17 |
| 5.3 | Broad Questions | 18 |
| Chapter 6 | Correlation and Regression | 19 |
| 6.1 | Creative Questions | 19 |
| 6.2 | Short Questions | 19 |
| Chapter 7 | Time Series | 20 |
| 7.1 | Creative Questions | 20 |
| 7.2 | Short Questions | 21 |
| Chapter 8 | Published Statistics in Bangladesh | 23 |
| 8.1 | Creative Questions | 23 |
| 8.2 | Short Questions | 23 |
| Conclusion | | 24 |

Chapter 1

Statistics, Variable and Concepts of Different Symbols

1.1 Creative Questions

1. A system analyst collected frequencies of a signal at different times. Then he realized due to some unknown noise, 0.5 units got added to all the values. The recorded values are given below:

10, 12, 15, 14, 12, 16, 20, 16, 18, 11

- (a) What is change of origin? 1
- (b) Does change of origin have an effect on median? 2
- (c) Find $\sum_{i=1}^{10} (X_i - 5)$. 3
- (d) Determine the summation of the values discarding the noise. 4
2. Marks obtained by five students in statistics out of 15 were 4, 6, 10, 12, and 15. The examiner said, the square of the sum of the marks is greater than the sum of the squares of the marks.
- (a) What is finite population? 1
- (b) Explain quantitative variable with an example. 2
- (c) In the light of the available data, find $\sum_{i=1}^5 (x_i - 2x_i)^2$ 3
- (d) Verify the comment of the examiner. 4
3. The capital and profit (in million BDT) of some Bangladeshi industries are bgiven below:

| Industry | 1 | 2 | 3 | 4 | 5 |
|-------------|----|----|----|----|----|
| Capital (X) | 20 | 15 | 26 | 31 | 18 |
| Profit (Y) | 15 | 10 | 17 | 25 | 10 |

- (a) What is finite population? 1
- (b) What are the functions of statistics? 2
- (c) Find the value of $\sum_{i=1}^5 \sum_{j=1}^5 (x_i - y_j)$ 3
- (d) Analyze the statement theoretically and empirically: $\sum_{i=1}^5 (4x_i - 6y_j) = 4 \sum_{i=1}^5 x_i - 6 \sum_{i=1}^5 y_j$ 4
4. Goals scored by a footballer in 25 matches are summarized as shown below.

| Goals | 0 | 1 | 2 | 3 | 4 |
|-------|---|---|---|---|---|
| Times | 8 | 9 | 5 | 2 | 1 |

- (a) Is no. goals a discrete or continuous variable? 1

- (b) Verify theoretically: $\sum_{i=1}^2 X_i Y_i = \sum_{i=1}^2 X_i \times \sum_{i=1}^2 Y_i$ 2
- (c) Find the total number of goals using a suitable notation. 3
- (d) If he scores two (2) goals in the next match, will the scoring rate increase? Explain logically and empirically 4

5. **A researcher has collected some data:**

$$x_1 = 15, x_2 = -12, x_3 = 10, x_4 = 21, x_5 = 33$$

- (a) What is sample? 1
- (b) Briefly explain shift or origin and scale. 2
- (c) Compute the value of $\sum_{i=1}^5 (x_i - 20)^2$ 3
- (d) Find the value of $\sum_{i=1}^5 (3x_i^2 - 2x_i - 3)$ and examine its dependency on origin and scale. 4

6. **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

7. **Marks of 10 students in Statistics in a class were found to be the following:**

$$99, 88, 98, 85, 97, 71, 87, 79, 70, 84$$

Later it was discovered that all marks should be 5 less than the recorded marks.

- (a) What is change of origin? 1
- (b) Does summation of a variable depend on change of origin? 2
- (c) Considering the data in stem as X, find $\sum_{i=1}^{10} X_i$ and $\sum_{i=1}^{10} (X_i + 3)$ 3
- (d) Find the arithmetic mean of the corrected values, employing the concept of shift of origin. 4

8. **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, $\sum_{i=1}^n \sum_{j=1}^n x_i y_j = \sum_{i=1}^n x_i y_i$? Vindicate 3
- (d) Using data, prove that sum of square is unequal to square of sum of numbers. 4

9. **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

| 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 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.

10. **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 $\sum_{i=1}^5 (y_i - 3)^2$
(d) Find total number of goals using two different notations and examine whether they match. 4

11. **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

12. **An analyst obtains some data:**

$$x_1 = 15, x_2 = -12, x_3 = 17, x_4 = 11, x_5 = 23$$

- (a) What is sample? 1
(b) Briefly explain shift or origin and scale. 2
(c) Compute the value of $\sum_{i=1}^5 (x_i - 10)^2$ 3
(d) Find the value of $\sum_{i=1}^5 (5x_i^2 - 4x_i - 3)$ and examine its dependency on origin and scale. 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

5. What are the functions of statistics? 1
6. What is the origin of the word statistics? 1
7. What is finite population? 1
8. Give an example of an infinite population. 1
9. What is a variable? 1
10. What is a sample? 1
11. What is a qualitative variable? 1
12. What is a quantitative variable? 1
13. What is a discrete variable? 1
14. What is a continuous variable? 1
15. Differentiate between discrete and continuous variable. 2
16. What is univariate data? 1
17. What is bivariate data? 1
18. Differentiate between qualitative and quantitative variable. 2
19. What are the scales of measurement?
20. What is nominal data?
21. What is ordinal data?
22. What is ratio data?
23. What is interval data?
24. Give examples of four scales of measurement 2
25. Differentiate between ratio and interval scale of measurement.
26. What is change of origin?
27. What is change of scale?
28. What is another way to write $\sum_{i=1}^n bx_i$?
29. After expansion, what does $\sum_{i=1}^n (ax_i - b)$ become?
30. What is the expansion of $\left(\sum_{i=1}^n x_i\right)^2$

Chapter 2

Collection, Presentation, and Organization of Data

2.1 Creative Questions

1. Favorite colors of 30 individuals are noted down. There are five different colors. The recorded colors are given below:

Brown Red Pink Green Green Green Brown Pink Brown Red
Brown Red Green Pink White Red Brown Green White Brown
White Brown Pink Red White Brown Green Red Pink Red

- (a) What is nominal data? 1
- (b) What are the ways to deal with categorical data? 2
- (c) Draw a Pie Chart from the above data and explain. 3
- (d) Is Bar Diagram a better representation of this data? Justify. 4

2. The favorite fruits recorded from 30 individuals are as follows:

Apple, Banana, Cherry, Banana, Mango, Apple, Cherry, Grape, Banana, Mango,
Cherry, Apple, Grape, Banana, Mango, Apple, Grape, Cherry, Apple, Banana,
Cherry, Mango, Apple, Banana, Grape, Cherry, Mango, Apple, Banana, Grape

- (a) What kind of data is this? 1
- (b) How can you present the data in this dataset? List three methods. 2
- (c) Draw a Pie Chart from the above data and explain. 3
- (d) Is Bar Diagram a better representation of this data? Justify. 4

3. Hourly wages of 100 workers in an industry were collected by a market analyst. The analyst desires to mine a pattern and useful insights from the collected data about the industry. The obtained data are demonstrated below:

| Wage | 51-55 | 56-60 | 61-65 | 66-70 | 71-75 | 76-80 | 81-85 |
|-------------------|-------|-------|-------|-------|-------|-------|-------|
| Number of workers | 7 | 11 | 18 | 36 | 15 | 8 | 5 |

- (a) What is class interval? 1
- (b) How does a frequency distribution help us to find pattern in data? 2
- (c) Draw an Ogive from the data provided and explain. 3
- (d) Write five useful insights about the data combining information from Ogive and the table. 4

4. The heights of 20 adult males were collected and found to be as follows:

170, 172, 168, 175, 180, 178, 169, 173, 176, 175
171, 174, 177, 165, 179, 167, 181, 164, 182, 172,

- (a) What are discrete class intervals? 1
- (b) What are the problems with discrete intervals 2
- (c) Create a Histogram from the data and explain 3

| | | | | | | | |
|-----------|-------|-------|-------|-------|-------|-------|-------|
| Class | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 | 70-80 |
| Frequency | 5 | 8 | 10 | 13 | 11 | 7 | 3 |

- (d) Create an Ogive and compare the focus of these two graphs.4
5. The marks of 57 students in a village are given below.
- (a) Are the intervals discrete or continuous?1
- (b) Question2
- (c) Create an Ogive and interpret.3
- (d) Question4

2.2 Short Questions

1. How is a bar diagram different from a pie chart?1
2. What are discrete class intervals?1
3. What are the problems with discrete intervals2
4. What are the two main types of variables?1
5. How is a bar diagram different from a pie chart?1
6. What does a frequency distribution show?1
7. How do you define a categorical variable?1
8. What is the primary use of a histogram?1
9. Explain the significance of the height of bars in a bar diagram.1
10. What information can you gather from a pie chart?1
11. How is data grouped in a frequency distribution?1
12. What does the width of the bins represent in a histogram?1
13. Why is it important to label the axes in a bar diagram?1
14. Define nominal data.1
15. What is the purpose of a frequency distribution?1
16. Explain the difference between discrete and continuous data.2
17. What is a histogram used for?1

Chapter 3

Measures of Central Tendency

3.1 Creative Questions

1. A shrimp producer wanted to get an insight into his shrimp production. To do so, he randomly collected weights of different shrimps in his farm.

| Weight of shrimp (gm) | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 |
|-----------------------|-------|-------|-------|-------|-------|
| Frequency | 5 | 8 | 10 | 9 | 4 |

- (a) What is the primary goal of central tendency? 1
- (b) When is Median a better measure of central tendency than Arithmetic Mean and why? 2
- (c) From the stem, find 3rd quartile and explain. 3
- (d) Find harmonic mean (HM) and compare with the arithmetic mean (AM) 4

2. An arithmetic series is formed as follows:

$$a, a + c, a + 2c, \dots, a + 2nc$$

- (a) What is change of origin and scale? 1
- (b) Convert the series into a set of natural numbers. 2
- (c) Find the arithmetic mean of the series with the use of change of origin and scale. 3
- (d) Find the geometric mean of the series: $1, 2, 4, \dots, 2^n$. 4

3. Scores by Travis Head in the last two matches of ICC Men's Cricket World - 2023 are given. In Cricket, Strike Rate (SR) is computed by dividing Runs by Balls and then multiplying the quotient by 100.

| Match | Runs | Balls |
|-------|------|-------|
| 1 | 62 | 48 |
| 2 | 137 | 120 |

- (a) How many averages do you know of? 1
- (b) Give an example when arithmetic mean is appropriate instead of harmonic mean. 2
- (c) When is Weighted Harmonic mean is used. Show a numerical example. 3
- (d) Determine the average Strike Rate of the batter 4

4. Average height of the four tallest towers in Dhaka is 153.25 meters. The heights of first three towers is 171, 153 and 152 meters, respectively. A new tower has been built with height 150 meters.

- (a) Write two primary uses of central tendency. 1
- (b) Prove mathematically: $\sum_{i=1}^n (x_i - \bar{x}) = 0$ 2
- (c) Compute the height of the forth tower. 3
- (d) After the addition of the new tower, will the average increase or decrease? Explain logically and empirically (using data). 4

5. For two non-zero positive numbers, $GM = 4\sqrt{3}$ and $HM = 6$, where the quantities bear usual notations
- When is Harmonic mean suitable? 1
 - For two numbers, what is the relationship between AM, GM, and HM? 2
 - What is the Arithmetic mean? 3
 - Determine the numbers. 4

6. 12 is deducted from each value of a variable and then divided by 3. The new arithmetic mean (AM) is found to be 4.

- What is change of origin? 1
- Does AM depend on origin? Prove with an example. 2
- From the stem, find the original AM. 3
- Does the origin or the scale have greater impact on AM in this example? 4

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

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

8. 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 |

- What is Median? 1
- Does median necessarily lie in the dataset? 2
- Estimate Median and explain the result. 3
- Find Arithmetic Mean and Mode. Which measure seems to be the best one? 4

9. 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 |

- When is Short-Cut method for Arithmetic Mean useful? 1
- Derive the formula of Short-cut Method 2
- Compute the Arithmetic Mean using the Short-cut method. 3
- Compute the Arithmetic Mean with a different value of origin (A). Do both the methods give same result? 4

10. 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 |

- What is central tendency? 1
- When is geometric mean appropriate to measure? 2
- Compute median from the stem. 3
- Show that Arithmetic mean is greater than Harmonic mean. Which one of them is more suitable for this data? 4

11. 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.

- Write down the formula of combined AM for k groups. 1
- What is the combined AM of two data sets with AM 35 and 45 and number of values equal? 2
- How many employees are there in the company B? 3
- 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

12. 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 |

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

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

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

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

| Marks | 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 |

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

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

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

16. HSC exam result of two colleges in a village are given below:

| College | Students | Passed |
|---------|----------|--------|
| 1 | 10 | 8 |
| 2 | 20 | 5 |

- How many measures of central tendency are there? 1
- Which measure of central tendency is perfect? 2
- Find the pass rate of the colleges and, using them, of the village. 3

- (d) Show an alternative method which gives the same result.

4

17. 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

18. 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 occassions) and 33 (including being not out in 11 occassions) 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

19. 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

20. 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 gemetric mean theoretically and empricially. 4

21. Grades of a an undergraduate student with major in statistics are given below:

[Credits serve as weights]

| 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

22. 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 mean could we use for the given data and why?

2
- (c) Find the average speed using weighted harmonic mean.

3
- (d) Find the correct and suitable average speed using another method and mathematically show they are equivalent.

4
23. A passer-by walks 3 hours at 5 km per hour (kph), another 3 hours at 4 kph, and another 3 hours at 3 kph.

(a) When is harmonic mean suitable?

1

(b) Which mean could we use for the given data and why?

2

(c) Find the average speed of the passer-by usingt he proper method.

3

(d) Find the correct and suitable average speed using another method and mathematically show they are equivalent.

4
24. 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 same value of the average? Demonstrate.

4
25. There are only two students in a class IX in a college. In half-yearly exam, the arithmetic and the geometric mean of the marks of those two studnets are 25 and 15, respectively.

(a) Write down the formula of Geometric Mean for grouped data.

1

(b) Prove with an example: $\sum_{i=1}^n (x_i - \bar{x}) = 0$

2

(c) Determine the marks of the students.

3

(d) Is 20 a possible value of the harmonic mean of this data? Explain theoretically and empirically.

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?

1

2. When is WAM equal to WHM? Show mathematically.

2

3. When is the equality true: $AM = GM = HM$? Prove mathematically and empirically.

3

4. When is Median a better measure of central tendency than Arithmetic Mean?

1

5. When is Harmonic Mean more suitable than Arithmetic Mean?

1

6. Write two primary uses of central tendency.

1

7. For two distinct non-zero values, what is the relationship among AM, GM, and HM?

1

8. what is the relationship among AM, GM, and HM?

1

9. What are the criteria of a good measure of central tendency?

1

10. For two non-zero positive numbers, prove $AM \geq GM \geq HM$

3

11. Find the AM, GM, and HM: $1, 2, 4, \dots, 2^n$

4

3.2.2 Arithmetic Mean

12. Does Arithmetic Mean depend on change of origin and scale? Prove mathematically and with an example. 3
13. Prove with an example: $\sum_{i=1}^n (x_i - \bar{x}) = 0$
14. Prove mathematically: $\sum_{i=1}^n (x_i - \bar{x}) = 0$
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:

| Remuneration (X) | No. of Employees (f) |
|------------------|----------------------|
| 30 | 5 |
| 35 | 8 |
| 40 | 10 |

28. Calculate Arithmetic Mean using short-cut method:

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

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

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] 2
37. $n_1 = 15, G_1 = 75, n_2 = 10, \text{ and } G_2 = 80$ Find combined GM.

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?

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 and interpret: 2, 1, 0, 5, -6, 7, -4

1

1

3

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

2. Two companies A and B pay their workers on a weekly basis. The summary of wages paid by them is shown below:

| Factory | Wage (BDT) | Standard Deviation | Number of workers |
|---------|------------|--------------------|-------------------|
| A | 1560 | 90 | 200 |
| B | 1580 | 70 | 160 |

- (a) What is dispersion? 1
 - (b) Is variance always greater than stanard deviation? Justify. 2
 - (c) Which company is more consistent with their wages? 3
 - (d) Find the combined Coefficient of Variance (CV) and compare with individual companies. 4
3. Mean and Standard Deviation of 200 items are found to be 60 and 20. Later it was found that two items were recorded as 3 and 67 in place of 13 and 17.
- (a) Does standard deviation depend on change of origin? 1
 - (b) Prove $\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2$ from original formula. 2
 - (c) Should the correct mean be smaller or greater? Also find it and compare. 3
 - (d) Find the correct standard deviation. 4

4.2 Short Questions

Chapter 5

Moments, SKewness, and Kurtosis

5.1 Creative Questions

1. Duration of stays of a spy in foreign countries are obtained by a researcher. As part of an analysis, s/he starts with the following summary.

| Duration | 1-10 | 11-20 | 21-30 | 31-40 | 41-50 | 51-60 |
|-----------|------|-------|-------|-------|-------|-------|
| Frequency | 4 | 3 | 3 | 2 | 5 | 2 |

- (a) What is symmetry? 1
 - (b) What is implied by the value of coefficient of skewness 0.8 2
 - (c) Estimate the median of the data and interpret. 3
 - (d) Obtain coefficient of skewness from data and comment on the life of the spy based on it. 4
2. 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) Show that the first central moment is zero. 2
 - (c) Find the variance of the data. 3
 - (d) Are the data symmetric? Justify. 4
3. 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
 4. 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. **The first four moments of a distribution around 5 are 2, 20, 40, and 50, respectively.**
- (a) 1
- (b) Draw the shape of a left-skewed distribution. 1
- (c) 2
- (d) Derive the value of the first central moment. 2
- (e) Obtain the first four central moments. 3
- (f) Estimate and comment on the skewness and kurtosis. 4
5. **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
6. **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 (in feet) 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 three central moments. 3
- (d) Find the skewness of the data and interpret. 4
7. **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
8. **Exam marks of two students were summarized for the purpose of comparison. The summary is given below:**
- (a) What is kurtosis? 1
- (b) How much data are contained within Interquartile range? 2

| Measure | Student X | Student Y |
|-----------------|-----------|-----------|
| First Quartile | 28 | 27 |
| Second Quartile | 60 | 60 |
| Third Quartile | 75 | 73 |
| Minimum | 16 | 14 |
| Maximum | 89 | 86 |

- (c) For student A, estimate the Bowley's Coefficient of skewness and explain. 3
- (d) On the basis of skewness (and hence shape of the data), compare the students. 4

9. The first four moments around 2 of a dataset were the following:

-1, 5, 20, 90

- (a) What is raw moment? 1
- (b) What is the standard deviation of the data in the stem? 2
- (c) Determine the third central moment. 3
- (d) Comment on the kurtosis of the given data. 4
10. The heights of the trees of a certain species that were planted at around the same time in a park were examined by an analyst, hired by the park authority to check for any abnormality. He randomly observed 10 trees; the values (in cm) obtained are given below:

200, 190, 185, 210, 220, 200, 205, 207, 230, 195

- (a) What is five number summary? 1
- (b) Which measures are shown on a Box & Whisker plot? 2
- (c) Display the data on a box plot. 3
- (d) Taking a look at the drawn box plot, comment on the symmetry of the data. 4

5.2 Short Questions

- What are moments? 1
- How many types of moments are there? 1
- What is central moment? 1
- What is raw moment? 1
- What is the formula of raw moment for grouped data? 1
- Derive the value of the first central moment. 2
- Derive the relationship between the first central moment and the origin.
- What is the second central moment equal to?
- Write 3 uses of moments.
- Can moments be negative? Analyze. 3
- What is skewness? 1
- What is a symmetric distribution?
- In a symmetric distribution, what is the relationship between Mean, Median, and Mode?
- What is negative skewness?
- What is positive skewness?
- What is the pattern of in a left-skewed distribution?
- What is the pattern of in a right-skewed distribution?
- In a left-skewed distribution, what is the relationship between Mean, Median, and Mode?

19. In a right-skewed distribution, what is the relationship between Mean, Median, and Mode?
20. Draw the shape of a left-skewed distribution.
21. Draw a symmetrical distribution.
22. Draw the shape of a right-skewed distribution.
23. What is Pearson's measure of skewness?
24. How did Bowley measure skewness?
25. What is Kelly's method of coefficient of skewness?
26. How are moments used to measure skewness?
27. What is the relationship between β_1 and γ_1
28. Explain $\gamma_1 > 0$
29. What is kurtosis?
30. How many types of kurtosis are there?
31. Draw the curve of a Mesokurtic distribution
32. Draw the curve of a Platykurtic distribution
33. Draw the curve of a Leptokurtic distribution
34. Explain the Platykurtic distribution. 2
35. How can you use moments to estimate kurtosis?
36. What is the relationship between β_2 and γ_2
37. Explain $\gamma_2 > 0$
38. What is five number summary?
39. Find five number summary: 2, 1, 0, 5, -6, 7, -4 3
40. Write three uses of five number summary 2
41. Which measures are shown on a Box & Whisker plot?
42. How is inner fence found for a Box & Whisker plot
43. How is outer fence found for a Box & Whisker plot
44. What is IQR and how is it calculated? 2
45. Write four uses of the box plot. 2

5.3 Broad Questions

1. Analyze the effect of shift of origin & scale on central moments. 3
2. Prove that β_1 and β_2 do not depend on shift of origin and scale. 4
3. Express the 3rd central moment in terms of raw moments.
4. The first four moments of a distribution around 5 are 2, 20, 40, and 50, respectively. Obtain the first four central moments. Estimate and comment on the skewness and kurtosis.

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. Bangladesh foreign debt has been increasing rapidly in recent years. The Bangladesh bank provides the following data.

| Fiscal Year | 2015-16 | 2016-17 | 2017-18 | 2018-19 | 2019-20 | 2020-21 | 2021-22 | 2022-23 | 2023-24 |
|-------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| Debt | 41.17 | 45.81 | 56.01 | 62.63 | 68.55 | 81.62 | 95.45 | 98.94 | ~130.00 |

- (a) Name the components of time series. 1
- (b) What are linear and non-linear trends? 2
- (c) Find 3-yearly moving average from the data and plot. 3
- (d) Which components of time series may underlie the data? Analyze. 4

3. 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

4. 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 be better if we used 3-yearly method? 4

5. Daily expense of a certain individual in the month of April is tracked by himself (sorted from the first day of the month through the last day). He intends to analyze and see which part of the month is more expensive for him.

1430 777 4101 4840 3251 4035 2504 371 4326 2296
47 3207 1608 384 4705 3424 1168 1189 4276 4749
1117 156 572 1181 1031 4508 149 80 4475 2087

- (a) What is semi-average method? 1
 - (b) Make a line chart from the data? **TO BE REVIEWED** 2
 - (c) Find the trend of the data and explain. 3
 - (d) How can the person accomplish what he intends? 4
6. **US Dollar to Taka Exchange rate since 2016 to 2023 (for each year, the rate in January is picked) is provided below.**

| Year | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 |
|-------------------|-------|-------|-------|-------|-------|-------|-------|--------|
| USD Exchange Rate | 78.35 | 79.49 | 82.87 | 83.26 | 84.60 | 84.37 | 85.80 | 106.70 |

Table 7.1: Source–Investing.com

- (a) How many methods are there to measure the trend? 1
 - (b) Distinguish between seasonal and cyclic variation. 2
 - (c) Discuss the advantages and disadvantages of moving average method. 3
 - (d) What is the expected Exchange rate in 2024? Estimate using a suitable method. 4
7. **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
8. **Average monthly temperatures (in °C) in the city of Sylhet are collected by an analyst. The analyst assumes the next month will not follow the current trend.**

| Month | Jan | Feb | Mar | Apr | May | Jun | Jul |
|-------------|------|------|------|------|------|------|------|
| Temperature | 25.2 | 27.1 | 30.4 | 30.8 | 30.9 | 30.9 | 31.6 |

- (a) What is seasonal variation? 1
- (b) Differentiate between seasonal variation and cyclic variation. 2
- (c) Find the general trend using semi-average method. 3
- (d) Find the trend using moving average method and examine the assumption of the analyst. [the genuine next value is 31.2] 4

7.2 Short Questions

- 1. What is a time series? 1
- 2. Give an example of a time series data. 1
- 3. What are the components of Time Series?
- 4. What is trend?
- 5. How many types of trends are there?
- 6. What is linear trend?
- 7. What is non-linear trend?
- 8. What is seasonal variation?
- 9. Give an example of seasonal variation.
- 10. What is cyclic variation.

11. Distinguish between seasonal and cyclic variation.
12. Give an example of cyclic variation.
13. What is irregular variation?
14. What is an example of irregular variation?
15. How many models are there for time series?
16. Mention the additive model of time series.
17. What is the multiplicative model?
18. How many methods are there to measure the trend?
19. What are the weaknesses of graphical method of measuring the trend?
20. How does semi-average method work? 2
21. What are the limitations of semi-average method?
22. How is moving average method used?
23. Discuss the advantages and disadvantages of moving average method. 3
24. When are moving averages centered? 2

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
2. **Climate change is an alarming problem throughout the world. To determine what to do to solve the problem, many government and non-government organizations collect and analyze data to come to a consistent solution.**
 - (a) What is official statistics? 1
 - (b) What is the role of World Meteorological Organization? 2
 - (c) What are the limitations of published statistics in Bangladesh? 3
 - (d) How can the quality of published statistics in Bangladesh be improved? 4
3. **Every country has one or more agencies to deal with statistics of the country for proper management of its assets and population. Bangladesh Bureau of Statistics (BBS) serves as the centralized official bureau in Bangladesh for collecting and disseminating statistics in Bangladesh. USA has several such agencies, like Census Bureau or Bureau of Labor Statistics.**
 - (a) What is data? 1
 - (b) How is statistics important in planning? 2
 - (c) Differentiate between official and non-official statistics. 3
 - (d) Elucidate the classification of published statistics in Bangladesh. 4
4. **Globalization creates many opportunities. Many government and non-government organizations collect and analyze data to analyze different aspects of the trend.**
 - (a) What is official statistics? 1
 - (b) What is the role of the United Nations? 2
 - (c) What are the limitations of published statistics in Bangladesh? 3
 - (d) Compare the statistical analysis in Bangladesh with international standard. 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