

Descriptive Statistics – Quiz

Probability & Statistics

Question 1

What are the two main categories of descriptive statistics?

- A. Mean and Median
- B. Numerical summaries and Visual summaries
- C. Population and Sample
- D. Discrete and Continuous

Question 1

What are the two main categories of descriptive statistics?

- A. Mean and Median
- B. Numerical summaries and Visual summaries
- C. Population and Sample
- D. Discrete and Continuous

Answer: B

Descriptive statistics has two main categories: Numerical summaries (mean, variance, percentiles) and Visual summaries (histograms, boxplots, scatter plots).

Question 2

For the test scores 70, 80, 90, what is the mean?

- A. 70
- B. 75
- C. 80
- D. 85

Question 2

For the test scores 70, 80, 90, what is the mean?

- A. 70
- B. 75
- C. 80
- D. 85

Answer: C

The mean is calculated as $(70 + 80 + 90) / 3 = 240 / 3 = 80$.

Question 3

Which measure of central tendency is most robust to outliers?

- A. Mean
- B. Median
- C. Mode
- D. Range

Question 3

Which measure of central tendency is most robust to outliers?

- A. Mean
- B. Median
- C. Mode
- D. Range

Answer: B

The median is robust to outliers because it only considers the middle value, not the actual values of extreme observations.

Question 4

What does the mode represent?

- A. The middle value when data is sorted
- B. The average of all values
- C. The most frequently occurring value
- D. The difference between max and min

Question 4

What does the mode represent?

- A. The middle value when data is sorted
- B. The average of all values
- C. The most frequently occurring value
- D. The difference between max and min

Answer: C

The mode is the value that occurs most frequently in a dataset. A dataset can have no mode, one mode (unimodal), or multiple modes (bimodal, multimodal).

Question 5

For which type of data is the mean typically preferred over the median?

- A. Skewed data
- B. Data with extreme outliers
- C. Income data
- D. Symmetric data

Question 5

For which type of data is the mean typically preferred over the median?

- A. Skewed data
- B. Data with extreme outliers
- C. Income data
- D. Symmetric data

Answer: D

For symmetric data, the mean is preferred because it uses all information in the data. For skewed data or data with outliers, the median is more appropriate.

Question 6

What are the quartiles Q1, Q2, and Q3?

- A. The first three values in the dataset
- B. The 25th, 50th, and 75th percentiles
- C. The mean, median, and mode
- D. The minimum, middle, and maximum values

Question 6

What are the quartiles Q1, Q2, and Q3?

- A. The first three values in the dataset
- B. The 25th, 50th, and 75th percentiles
- C. The mean, median, and mode
- D. The minimum, middle, and maximum values

Answer: B

Quartiles divide the sorted data into four equal parts. Q1 is the 25th percentile, Q2 is the 50th percentile (median), and Q3 is the 75th percentile.

Question 7

The Interquartile Range (IQR) is defined as:

- A. Maximum minus Minimum
- B. Q3 minus Q1
- C. Mean minus Median
- D. Standard deviation squared

Question 7

The Interquartile Range (IQR) is defined as:

- A. Maximum minus Minimum
- B. Q3 minus Q1
- C. Mean minus Median
- D. Standard deviation squared

Answer: B

The IQR = $Q3 - Q1$, which represents the range containing the middle 50% of the data. It is robust to outliers.

Question 8

Why do we square the distances from the mean when calculating variance?

- A. To make the calculation faster
- B. To make all distances positive and penalize large deviations more
- C. To reduce the impact of outliers
- D. To convert to different units

Question 8

Why do we square the distances from the mean when calculating variance?

- A. To make the calculation faster
- B. To make all distances positive and penalize large deviations more
- C. To reduce the impact of outliers
- D. To convert to different units

Answer: B

Squaring ensures all distances are positive (preventing cancellation) and penalizes large deviations more heavily than small ones.

Question 9

Why do we divide by $(n-1)$ instead of n when calculating sample variance?

- A. To make the calculation simpler
- B. It is a technical correction for samples (degrees of freedom)
- C. To reduce variance
- D. To convert to standard deviation

Question 9

Why do we divide by $(n-1)$ instead of n when calculating sample variance?

- A. To make the calculation simpler
- B. It is a technical correction for samples (degrees of freedom)
- C. To reduce variance
- D. To convert to standard deviation

Answer: B

Dividing by $(n-1)$ is called Bessel's correction. It provides an unbiased estimate of the population variance when working with samples.

Question 10

What is the relationship between standard deviation and variance?

- A. Standard deviation = Variance squared
- B. Standard deviation = Square root of Variance
- C. Standard deviation = Variance / n
- D. Standard deviation = Variance + Mean

Question 10

What is the relationship between standard deviation and variance?

- A. Standard deviation = Variance squared
- B. Standard deviation = Square root of Variance
- C. Standard deviation = Variance / n
- D. Standard deviation = Variance + Mean

Answer: B

Standard deviation is the square root of variance. This brings the measure back to the original units of the data, making it more interpretable.

Question 11

In finance, what does the standard deviation of returns measure?

- A. Expected profit
- B. Average return
- C. Volatility (risk)
- D. Maximum loss

Question 11

In finance, what does the standard deviation of returns measure?

- A. Expected profit
- B. Average return
- C. Volatility (risk)
- D. Maximum loss

Answer: C

In finance, standard deviation of returns is called volatility. Higher volatility means higher risk because returns are more spread out and unpredictable.

Question 12

What is the purpose of the Coefficient of Variation (CV)?

- A. To calculate the median
- B. To compare variability across different scales
- C. To identify outliers
- D. To measure skewness

Question 12

What is the purpose of the Coefficient of Variation (CV)?

- A. To calculate the median
- B. To compare variability across different scales
- C. To identify outliers
- D. To measure skewness

Answer: B

CV = (Standard Deviation / Mean) × 100%. It allows comparison of variability across datasets with different units or scales.

Question 13

What does a histogram show?

- A. The relationship between two variables
- B. The distribution of a single variable
- C. The five-number summary
- D. The correlation between variables

Question 13

What does a histogram show?

- A. The relationship between two variables
- B. The distribution of a single variable
- C. The five-number summary
- D. The correlation between variables

Answer: B

A histogram shows the distribution of a single variable by grouping data into bins and displaying the frequency of values in each bin.

Question 14

In a boxplot, what do the points beyond the whiskers represent?

- A. The mean
- B. The mode
- C. Outliers
- D. The median

Question 14

In a boxplot, what do the points beyond the whiskers represent?

- A. The mean
- B. The mode
- C. Outliers
- D. The median

Answer: C

Points beyond the whiskers (which extend to 1.5 IQR from the box) are considered outliers and are plotted individually.

Question 15

A right-skewed distribution has:

- A. A long left tail, mean $<$ median
- B. A long right tail, mean $>$ median
- C. Equal tails, mean = median
- D. No tail, only a peak

Question 15

A right-skewed distribution has:

- A. A long left tail, mean $<$ median
- B. A long right tail, mean $>$ median
- C. Equal tails, mean = median
- D. No tail, only a peak

Answer: B

In a right-skewed distribution, the tail extends to the right. The mean is pulled toward the long tail, so mean $>$ median. Income is a classic example.

What does positive excess kurtosis indicate?

- A. The distribution is symmetric
- B. The distribution has fat tails (more extreme values than normal)
- C. The distribution has thin tails
- D. The mean equals the median

What does positive excess kurtosis indicate?

- A. The distribution is symmetric
- B. The distribution has fat tails (more extreme values than normal)
- C. The distribution has thin tails
- D. The mean equals the median

Answer: B

Positive excess kurtosis means fat tails - extreme values occur more often than in a normal distribution. This is common in financial returns.

Question 17

What are the components of the five-number summary?

- A. Mean, Median, Mode, Variance, Standard Deviation
- B. Minimum, Q1, Median, Q3, Maximum
- C. Q1, Q2, Q3, Q4, Q5
- D. Mean, Standard Deviation, Skewness, Kurtosis, Range

Question 17

What are the components of the five-number summary?

- A. Mean, Median, Mode, Variance, Standard Deviation
- B. Minimum, Q1, Median, Q3, Maximum
- C. Q1, Q2, Q3, Q4, Q5
- D. Mean, Standard Deviation, Skewness, Kurtosis, Range

Answer: B

The five-number summary consists of Minimum, Q1 (25th percentile), Median (Q2), Q3 (75th percentile), and Maximum.

Question 18

If a dataset has values 2, 4, 6, 8, 10, what is the median?

- A. 4
- B. 5
- C. 6
- D. 7

Question 18

If a dataset has values 2, 4, 6, 8, 10, what is the median?

- A. 4
- B. 5
- C. 6
- D. 7

Answer: C

With 5 values (odd n), the median is the middle value. When sorted: 2, 4, 6, 8, 10. The third value (6) is the median.

Question 19

A scatter plot is used to show:

- A. The distribution of one variable
- B. The five-number summary
- C. The relationship between two variables
- D. The mode of the data

Question 19

A scatter plot is used to show:

- A. The distribution of one variable
- B. The five-number summary
- C. The relationship between two variables
- D. The mode of the data

Answer: C

A scatter plot displays the relationship between two variables by plotting each observation as a point with x and y coordinates.

Stock returns typically exhibit which characteristic compared to a normal distribution?

- A. Thinner tails
- B. Fat tails (more extreme events)
- C. No tails
- D. Exactly normal tails

Stock returns typically exhibit which characteristic compared to a normal distribution?

- A. Thinner tails
- B. Fat tails (more extreme events)
- C. No tails
- D. Exactly normal tails

Answer: B

Stock returns have fat tails - crashes and booms occur more frequently than a normal distribution predicts. This is why kurtosis matters in finance.