

Introduction to Probability and Statistics – Quiz

Probability & Statistics

Question 1

What are the two main branches of statistics?

- A. Theoretical and Applied
- B. Descriptive and Inferential
- C. Probability and Uncertainty
- D. Qualitative and Quantitative

Question 1

What are the two main branches of statistics?

- A. Theoretical and Applied
- B. Descriptive and Inferential
- C. Probability and Uncertainty
- D. Qualitative and Quantitative

Answer: B

Statistics has two main branches: Descriptive Statistics (summarizing and visualizing data) and Inferential Statistics (drawing conclusions about populations from samples).

Question 2

Which interpretation of probability views it as a long-run frequency?

- A. Bayesian
- B. Classical
- C. Frequentist
- D. Subjective

Question 2

Which interpretation of probability views it as a long-run frequency?

- A. Bayesian
- B. Classical
- C. Frequentist
- D. Subjective

Answer: C

The Frequentist interpretation views probability as the long-run frequency of an event. For example, if you flip a coin 1000 times, you expect approximately 500 heads.

Question 3

In probability theory, we reason from:

- A. Data to Model
- B. Sample to Population
- C. Model to Data
- D. Unknown to Known

Question 3

In probability theory, we reason from:

- A. Data to Model
- B. Sample to Population
- C. Model to Data
- D. Unknown to Known

Answer: C

Probability uses forward reasoning: given a known model (e.g., a fair coin), we predict what outcomes are likely. This is Model to Data reasoning.

Question 4

A researcher collects data from 500 customers to understand the buying behavior of all customers. The 500 customers represent the:

- A. Population
- B. Parameter
- C. Sample
- D. Statistic

Question 4

A researcher collects data from 500 customers to understand the buying behavior of all customers. The 500 customers represent the:

- A. Population
- B. Parameter
- C. Sample
- D. Statistic

Answer: C

A sample is a subset of the population that we actually observe. The 500 customers are a sample used to learn about the population (all customers).

Question 5

The true average return of all stocks in the market is denoted by which symbol?

- A. \bar{x}
- B. s
- C. μ
- D. $\hat{\sigma}$

Question 5

The true average return of all stocks in the market is denoted by which symbol?

- A. \bar{x}
- B. s
- C. μ
- D. $\hat{\sigma}$

Answer: C

Population parameters are denoted by Greek letters. Mu represents the true population mean. Latin letters like \bar{x} represent sample statistics.

Question 6

Credit ratings (AAA, AA, A, BBB) are an example of what type of data?

- A. Continuous
- B. Discrete
- C. Nominal
- D. Ordinal

Question 6

Credit ratings (AAA, AA, A, BBB) are an example of what type of data?

- A. Continuous
- B. Discrete
- C. Nominal
- D. Ordinal

Answer: D

Ordinal data are categories with a meaningful order. Credit ratings have a clear ranking (AAA \succ AA \succ A \succ BBB) but the intervals between them are not equal.

Question 7

Which scientist is credited with discovering the normal distribution (bell curve)?

- A. Blaise Pascal
- B. Thomas Bayes
- C. Carl Friedrich Gauss
- D. Ronald Fisher

Question 7

Which scientist is credited with discovering the normal distribution (bell curve)?

- A. Blaise Pascal
- B. Thomas Bayes
- C. Carl Friedrich Gauss
- D. Ronald Fisher

Answer: C

Carl Friedrich Gauss discovered the normal distribution in 1809, which is why it is also called the Gaussian distribution.

Question 8

If $P(\text{heads}) = 0.5$ for a fair coin, what does this probability value mean?

- A. Heads will occur exactly 50 times in 100 flips
- B. Heads is impossible
- C. Heads and tails are equally likely
- D. The coin is biased

Question 8

If $P(\text{heads}) = 0.5$ for a fair coin, what does this probability value mean?

- A. Heads will occur exactly 50 times in 100 flips
- B. Heads is impossible
- C. Heads and tails are equally likely
- D. The coin is biased

Answer: C

$P(\text{heads}) = 0.5$ means heads and tails are equally likely. In practice, 100 flips might yield 48 or 53 heads due to randomness, but on average it will be close to 50.

Question 9

What is volatility in finance?

- A. The average stock price
- B. How much prices jump around (big swings)
- C. The total return over a year
- D. The number of trades per day

Question 9

What is volatility in finance?

- A. The average stock price
- B. How much prices jump around (big swings)
- C. The total return over a year
- D. The number of trades per day

Answer: B

Volatility measures how much prices fluctuate. High volatility means big price swings (both up and down), while low volatility means more stable prices.

Question 10

Statistics uses which type of reasoning?

- A. Forward reasoning (Model to Data)
- B. Backward reasoning (Data to Model)
- C. Circular reasoning
- D. Deductive reasoning only

Question 10

Statistics uses which type of reasoning?

- A. Forward reasoning (Model to Data)
- B. Backward reasoning (Data to Model)
- C. Circular reasoning
- D. Deductive reasoning only

Answer: B

Statistics uses backward reasoning: given observed data, we try to infer what model or parameters explain it. This is Data to Model reasoning.

Question 11

The number of trades executed in a day is an example of what type of data?

- A. Continuous
- B. Discrete
- C. Nominal
- D. Ordinal

Question 11

The number of trades executed in a day is an example of what type of data?

- A. Continuous
- B. Discrete
- C. Nominal
- D. Ordinal

Answer: B

Discrete data can only take specific values (whole numbers). You can have 100 or 101 trades, but not 100.5 trades.

Question 12

Which of the following is a parameter?

- A. The average height of 50 sampled students
- B. The standard deviation calculated from a dataset
- C. The true proportion of defective items in all production
- D. The median income in a survey

Question 12

Which of the following is a parameter?

- A. The average height of 50 sampled students
- B. The standard deviation calculated from a dataset
- C. The true proportion of defective items in all production
- D. The median income in a survey

Answer: C

A parameter is a fixed value describing the entire population. The true proportion of defective items in ALL production is a parameter. The others are statistics calculated from samples.

Question 13

Why do we use samples instead of studying entire populations?

- A. Samples are always more accurate
- B. Populations are often too large or impossible to measure completely
- C. Parameters can only be calculated from samples
- D. Statistics require small datasets

Question 13

Why do we use samples instead of studying entire populations?

- A. Samples are always more accurate
- B. Populations are often too large or impossible to measure completely
- C. Parameters can only be calculated from samples
- D. Statistics require small datasets

Answer: B

Populations are often too large, expensive, or impossible to measure completely. We use representative samples to draw valid conclusions about the population.

Question 14

A return of +5% on a stock means:

- A. The stock price is \$5
- B. The stock price increased by 5 cents
- C. The stock price increased by 5% of its previous value
- D. 5% of investors made money

Question 14

A return of +5% on a stock means:

- A. The stock price is \$5
- B. The stock price increased by 5 cents
- C. The stock price increased by 5% of its previous value
- D. 5% of investors made money

Answer: C

A return is the percentage change in price. A +5% return means the price increased by 5% of its previous value (e.g., from \$100 to \$105).

Question 15

Which probability interpretation would say 'I am 70% confident it will rain tomorrow'?

- A. Classical
- B. Frequentist
- C. Bayesian
- D. Empirical

Question 15

Which probability interpretation would say 'I am 70% confident it will rain tomorrow'?

- A. Classical
- B. Frequentist
- C. Bayesian
- D. Empirical

Answer: C

The Bayesian interpretation views probability as a degree of belief. Saying 'I am 70% confident' reflects a personal belief about the likelihood of an event.

What is a random walk in finance?

- A. A guaranteed pattern of stock price movement
- B. A path where each step is unpredictable, like coin flips
- C. A strategy to beat the market
- D. The average path of all stocks

What is a random walk in finance?

- A. A guaranteed pattern of stock price movement
- B. A path where each step is unpredictable, like coin flips
- C. A strategy to beat the market
- D. The average path of all stocks

Answer: B

A random walk is a path where each step is unpredictable. In finance, this models stock prices where future movements cannot be predicted from past movements.

Question 17

Sector classifications (tech, finance, health) are an example of what type of data?

- A. Continuous
- B. Discrete
- C. Nominal
- D. Ordinal

Question 17

Sector classifications (tech, finance, health) are an example of what type of data?

- A. Continuous
- B. Discrete
- C. Nominal
- D. Ordinal

Answer: C

Nominal data are categories without any natural order. Sector names are labels with no inherent ranking (tech is not 'greater than' finance).

Question 18

The hat symbol ($\hat{}$) above a parameter indicates:

- A. The parameter is unknown
- B. An estimate of the parameter
- C. The parameter is negative
- D. The parameter is from a sample

Question 18

The hat symbol ($\hat{}$) above a parameter indicates:

- A. The parameter is unknown
- B. An estimate of the parameter
- C. The parameter is negative
- D. The parameter is from a sample

Answer: B

The hat symbol indicates an estimate of a parameter. For example, $\hat{\mu}$ is an estimate of the true population mean μ , calculated from sample data.

Question 19

In a return distribution, extreme gains and losses are found in the:

- A. Center of the distribution
- B. Mode
- C. Tails of the distribution
- D. Mean

Question 19

In a return distribution, extreme gains and losses are found in the:

- A. Center of the distribution
- B. Mode
- C. Tails of the distribution
- D. Mean

Answer: C

Extreme values (both gains and losses) are found in the tails of the distribution. Most returns cluster near the center, while tail events are rare but impactful.

Ronald Fisher is known for developing:

- A. The normal distribution
- B. Bayesian reasoning
- C. Modern hypothesis testing and experimental design
- D. Probability theory through gambling problems

Question 20

Ronald Fisher is known for developing:

- A. The normal distribution
- B. Bayesian reasoning
- C. Modern hypothesis testing and experimental design
- D. Probability theory through gambling problems

Answer: C

Ronald Fisher (1920s) created the foundations of modern hypothesis testing and experimental design, which are still widely used today in scientific research.