

Question 1

What is the fundamental problem of causal inference?

- A) Determining the accuracy of a model
- B) Predicting future outcomes
- C) Observing the counterfactual outcome
- D) Finding the best machine learning algorithm

Correct Answer: C) Observing the counterfactual outcome

Explanation: The fundamental problem of causal inference is the inability to observe the counterfactual outcome for any individual unit. This means we cannot see both the outcome that did happen and the outcome that would have happened under a different treatment for the same unit.

Question 2

What does a randomized control trial (RCT) help to ensure?

- A) Independence between treatment and outcome
- B) Independence between treatment and potential outcomes
- C) Increased bias in treatment assignment
- D) Correlation between treatment and outcome

Correct Answer: B) Independence between treatment and potential outcomes

Explanation: RCTs help to ensure independence between the treatment and potential outcomes, making the treated and control groups comparable except for the treatment they receive.

Question 3

What is the purpose of using control variables in causal inference studies?

- A) To increase the sample size
- B) To account for confounding variables
- C) To randomize the treatment
- D) To simplify the model

Correct Answer: B) To account for confounding variables

Explanation: Control variables are used to account for confounding variables, which are external factors that could influence both the treatment and the outcome, potentially biasing the causal effect estimation.

Question 4

Which of the following is NOT a common method for estimating causal effects from observational data?

- A) Linear regression
- B) Propensity score matching
- C) Random forest
- D) XGBoost

Correct Answer: D) XGBoost

Explanation: While XGBoost is a powerful gradient boosting algorithm used for prediction, it is not specifically designed for estimating causal effects from observational data. Propensity score matching and linear regression are commonly used for this purpose.

Question 5

What does the average treatment effect (ATE) measure?

- A) The effect of the treatment on the treated subjects only
- B) The effect of the treatment across the entire population
- C) The correlation between treatment and outcome
- D) The variance of the treatment effect

Correct Answer: B) The effect of the treatment across the entire population

Explanation: ATE measures the average effect of the treatment across the entire population, estimating the difference in outcome between subjects if all had received the treatment versus if none had received the treatment.

Question 6

In the context of causal inference, what is a confounder?

- A) A variable that is affected by the treatment
- B) A variable that mediates the relationship between treatment and outcome
- C) A variable that influences both the treatment and the outcome
- D) A variable that is irrelevant to the study

Correct Answer: C) A variable that influences both the treatment and the outcome

Explanation: A confounder is a variable that influences both the treatment and the outcome, potentially biasing the estimation of the causal effect if not properly controlled for in the analysis.

Question 7

Which of the following assumptions is crucial for causal inference?

- A) Large sample size
- B) No interference
- C) High model accuracy
- D) High correlation between variables

Correct Answer: B) No interference

Explanation: The no interference assumption (or SUTVA) is crucial for causal inference, stating that the treatment of one unit does not affect the outcome of another unit. This assumption allows for a clear attribution of observed effects to the treatment itself.

Question 8

What is the goal of propensity score matching?

- A) To increase the sample size
- B) To ensure balanced groups based on observed covariates
- C) To randomize the treatment assignment
- D) To predict the outcome variable

Correct Answer: B) To ensure balanced groups based on observed covariates

Explanation: Propensity score matching aims to ensure that treated and control groups are balanced based on observed covariates, reducing bias in the estimation of treatment effects from observational data.

Question 9

Which statement best describes the concept of exchangeability in causal inference?

- A) Treated and control units have the same distribution of covariates.
- B) Treated units are more likely to exhibit the outcome than control units.
- C) Control units can be easily exchanged with treated units in the model.
- D) The outcome variables are exchangeable between models.

Correct Answer: A) Treated and control units have the same distribution of covariates.

Explanation: Exchangeability means that treated and control units are comparable, having the same distribution of covariates. This condition supports the assumption that any observed differences in outcomes are attributable to the treatment rather than pre-existing differences.

Question 10

What does the term "counterfactual" refer to in causal inference?

- A) The factual outcome observed under the actual treatment received
- B) A hypothetical outcome that could have occurred under a different treatment
- C) The effect of a treatment on an outcome
- D) A variable that causes both the treatment and the outcome

Correct Answer: B

Explanation: Counterfactual refers to a hypothetical outcome that we could have observed if the unit had received a different treatment than it actually did. It's used to compare against the factual outcome to estimate the causal effect of the treatment.

Question 11

What is the fundamental problem of causal inference?

- A) Identifying appropriate data sources
- B) The inability to observe both the factual and counterfactual outcomes for the same unit
- C) Overfitting the causal model
- D) Choosing the right statistical model for analysis

Correct Answer: B

Explanation: The fundamental problem of causal inference is that for any given unit, we can observe the outcome under the treatment it received but not under the alternative treatment scenario. This makes it challenging to directly measure the causal effect.

Question 12

What is the purpose of randomization in a randomized controlled trial (RCT)?

- A) To ensure the study results are generalizable to the population
- B) To balance both observed and unobserved covariates across treatment groups

- C) To increase the statistical power of the study
- D) To reduce the cost of the study

Correct Answer: B

Explanation: Randomization in RCTs ensures that both observed and unobserved covariates are balanced across treatment groups, thereby reducing bias in the estimation of the treatment effect.

Question 13

Which of the following best describes a confounder?

- A) A variable that is affected by the treatment
- B) A variable that mediates the relationship between the treatment and outcome
- C) A variable that influences both the treatment and the outcome, potentially biasing the estimated effect
- D) A variable that is used to stratify the sample

Correct Answer: C

Explanation: A confounder is a variable that influences both the treatment and the outcome, creating a spurious association that can bias the estimated effect of the treatment on the outcome if not properly controlled for.

Question 14

What does the Average Treatment Effect (ATE) measure?

- A) The effect of the treatment on the treated group only
- B) The difference in outcomes between the treated and untreated groups in the sample
- C) The average effect of the treatment across all units in the population
- D) The median effect of the treatment across all units

Correct Answer: C

Explanation: The Average Treatment Effect (ATE) measures the average effect of the treatment across all units in the population, reflecting the difference in outcomes we would expect to see if all units received the treatment versus if no units received the treatment.

Question 15

Which method is commonly used to estimate causal effects from observational data?

- A) Linear regression
- B) Propensity score matching
- C) Pearson correlation
- D) Principal component analysis

Correct Answer: B

Explanation: Propensity score matching is a method commonly used to estimate causal effects from observational data by matching units with similar propensity scores (probability of receiving the treatment) to control for confounding variables.

Question 16

In causal inference, what does the term "heterogeneous treatment effects" refer to?

- A) Variability in treatment effects across different studies
- B) The effect of treatment varying across different levels of the treatment
- C) The effect of treatment varying across different subpopulations or units
- D) Differences in treatment effects observed in randomized trials versus observational studies

Correct Answer: C

Explanation: Heterogeneous treatment effects refer to the phenomenon where the effect of treatment varies across different subpopulations or units, highlighting that the impact of the treatment is not uniform for all individuals.