Crash Course in Causality - Multiple Choice Quiz

Question 1: What is the main distinction between correlation and causation?

1. Correlation indicates a cause-and-effect relationship.
2. Causation implies that one event leads to another. ✅
3. Correlation always implies causation.
4. Correlation measures the strength of a relationship between two variables. ✅

Explanation:

- **Option 1**: Incorrect. Correlation only shows a statistical relationship, not cause-and-effect.
- **Option 2**: Correct. Causation establishes a direct link between one event and another.
- **Option 3**: Incorrect. Correlation does not necessarily imply causation without additional evidence.
- **Option 4**: Correct. Correlation quantifies the degree to which two variables are linearly related.

Question 2: Which of the following methods helps establish causality?

1. Randomized Controlled Trials (RCTs) ✅
2. Observational Studies
3. Directed Acyclic Graphs (DAGs) ✅
4. Correlation Analysis

Explanation:

- **Option 1**: Correct. RCTs are the gold standard for establishing causality by randomizing the treatment.
- **Option 2**: Incorrect. Observational studies can suggest causation but are prone to confounding.
- **Option 3**: Correct. DAGs visually represent causal relationships based on domain knowledge.
- **Option 4**: Incorrect. Correlation analysis only identifies statistical relationships, not causality.

Question 3: What is the purpose of a counterfactual in causal analysis?

1. To identify the actual outcome.
2. To estimate what would have happened if an event had not occurred. ✅
3. To calculate the correlation coefficient.
4. To measure the statistical significance of an event.

Explanation:

- **Option 1**: Incorrect. The actual outcome is observed, not counterfactual.
- **Option 2**: Correct. Counterfactuals model alternate scenarios to understand causal effects.
- **Option 3**: Incorrect. Correlation coefficients do not involve counterfactuals.
- **Option 4**: Incorrect. Statistical significance testing does not directly involve counterfactuals.

Question 4: Which statement about confounding variables is correct?

1. They are irrelevant to causal analysis.
2. They can create spurious relationships between variables. ✅
3. They enhance the reliability of causal estimates.
4. They must be controlled to isolate the true causal effect. ✅

Explanation:

- **Option 1**: Incorrect. Confounders are critical to identifying true causal relationships.
- **Option 2**: Correct. Confounders can obscure or mimic causal effects if not accounted for.
- **Option 3**: Incorrect. Confounders reduce reliability unless controlled.
- **Option 4**: Correct. Controlling for confounders isolates the actual causal effect.

Question 5: What is a Directed Acyclic Graph (DAG)?

1. A method for estimating correlations.
2. A graph that represents causal relationships between variables. ✅
3. A tool for identifying confounders. ✅
4. A network that can include feedback loops.

Explanation:

- **Option 1**: Incorrect. DAGs are used for causality, not correlation estimation.
- **Option 2**: Correct. DAGs visualize how variables causally influence one another.
- **Option 3**: Correct. DAGs can identify and control confounding variables.
- **Option 4**: Incorrect. DAGs do not allow feedback loops as they are acyclic.

Question 6: What is the significance of temporal precedence in causality?

1. It proves correlation.
2. It ensures that the cause precedes the effect. ✅
3. It determines statistical significance.
4. It eliminates all confounders.

Explanation:

- **Option 1**: Incorrect. Temporal precedence does not prove correlation.
- **Option 2**: Correct. Temporal precedence is necessary for establishing causation.
- **Option 3**: Incorrect. Temporal precedence is unrelated to statistical significance.
- **Option 4**: Incorrect. Temporal precedence alone cannot eliminate confounders.

Question 7: In RCTs, what is the primary purpose of randomization?

1. To create confounding variables.
2. To reduce selection bias. ✅
3. To guarantee causation.
4. To ensure a representative sample.

Explanation:

- **Option 1**: Incorrect. Randomization minimizes, not creates, confounding.
- **Option 2**: Correct. Randomization eliminates selection bias by distributing confounders equally.
- **Option 3**: Incorrect. Causation is inferred from the design and results, not randomization alone.
- **Option 4**: Incorrect. Randomization does not necessarily ensure representativeness.

Question 8: What is the "average treatment effect" (ATE) in causal analysis?

1. The average of all observed outcomes.
2. The difference in outcomes between treated and untreated groups. ✅
3. The average of the counterfactual outcomes.
4. The causal effect of an intervention.

Explanation:

- **Option 1**: Incorrect. ATE focuses on treatment and control differences.
- **Option 2**: Correct. ATE measures the impact of the treatment on average.
- **Option 3**: Incorrect. Counterfactuals are not directly observed but estimated.
- **Option 4**: Correct. ATE represents the causal effect at a population level.

Question 9: Why are observational studies prone to bias?

1. They rely on random assignment.
2. They cannot control for all confounders. ✅
3. They manipulate variables directly.
4. They provide accurate causal estimates.

Explanation:

- **Option 1**: Incorrect. Observational studies do not use randomization.
- **Option 2**: Correct. Observational studies may miss some confounders, introducing bias.
- **Option 3**: Incorrect. Observational studies do not involve manipulation.
- **Option 4**: Incorrect. They can be biased without robust methodologies.

Question 10: What is Simpson's Paradox in causality?

1. A statistical phenomenon where trends reverse when stratified by groups. ✅
2. A paradox caused by missing data.
3. A bias introduced by randomization.
4. A situation where causality and correlation are equivalent.

Explanation:

- **Option 1**: Correct. Simpson's Paradox highlights how aggregate trends can differ from group-specific trends.
- **Option 2**: Incorrect. The paradox is unrelated to missing data.
- **Option 3**: Incorrect. Randomization reduces bias, not causes it.
- **Option 4**: Incorrect. Simpson’s Paradox involves stratification, not correlation-causation equivalence.

Question 11: What is the purpose of sensitivity analysis in causal inference?

1. To validate causal assumptions. ✅
2. To estimate treatment effects.
3. To examine the robustness of results to unmeasured confounders. ✅
4. To identify the direction of causal effects.

Explanation:

- **Option 1**: Correct. Sensitivity analysis evaluates how results change with different assumptions.
- **Option 2**: Incorrect. While useful, sensitivity analysis primarily tests assumptions.
- **Option 3**: Correct. It checks robustness by introducing hypothetical unmeasured confounders.
- **Option 4**: Incorrect. Sensitivity analysis does not identify causal directions.

Question 12: What is the goal of propensity score matching?

1. To balance observed covariates between treatment and control groups. ✅
2. To randomize treatment allocation.
3. To estimate causal effects without confounding. ✅
4. To identify temporal precedence.

Explanation:

- **Option 1**: Correct. Propensity scores ensure covariate balance, mimicking randomization.
- **Option 2**: Incorrect. Propensity scores are used post hoc, not for randomization.
- **Option 3**: Correct. By reducing confounding, it estimates causal effects more reliably.
- **Option 4**: Incorrect. Temporal precedence is unrelated to propensity scores.

Question 13: Which assumption is critical for causal inference in observational studies?

1. Random assignment of treatments.
2. No hidden confounders. ✅
3. Perfect correlation between variables.
4. Large sample size.

Explanation:

- **Option 1**: Incorrect. Observational studies lack randomization.
- **Option 2**: Correct. Unmeasured confounders can bias causal estimates.
- **Option 3**: Incorrect. Perfect correlation is unnecessary and often impractical.
- **Option 4**: Incorrect. While desirable, sample size does not guarantee causality.

Question 14: How do instrumental variables help in causal analysis?

1. By identifying confounders.
2. By estimating causal effects when randomization is infeasible. ✅
3. By serving as a replacement for the outcome variable.
4. By directly measuring the treatment effect.

Explanation:

- **Option 1**: Incorrect. Instrumental variables do not directly identify confounders.
- **Option 2**: Correct. They help estimate causal effects in non-randomized settings.
- **Option 3**: Incorrect. Instrumental variables are linked to treatments, not outcomes.
- **Option 4**: Incorrect. They enable inference, not direct measurement.

Question 15: Why is double-blinding used in RCTs?

1. To reduce selection bias.
2. To prevent placebo effects. ✅
3. To ensure randomization.
4. To minimize experimenter bias. ✅

Explanation:

- **Option 1**: Incorrect. Randomization, not blinding, reduces selection bias.
- **Option 2**: Correct. Blinding participants prevents psychological biases.
- **Option 3**: Incorrect. Randomization occurs before blinding.
- **Option 4**: Correct. Blinding experimenters avoids influencing outcomes.