Glossary definition

Boris Sobolev, Jillian Ashley Martin

Causal inference: the thought process to conclude whether a relation of cause to effect exists

Causality: the property of being related in causal fashion

Causation: the production of an effect by its cause

Conditional probability: the quotient of the joint probability of two events in the same space and the marginal probability of one of the events

Confounding: association between two variables due to a third variable

Counterfactual reasoning: comparing potential outcomes of treatment alternatives with observed outcomes

Directed acyclic graph: a graph in which nodes represent variables and arrows represent dependencies between the variables that the arrows connect without loops; an arrow drawn from variable X to variable Y represents a dependency of Y on X, and no dependency of X on Y

Causal diagram: a directed acyclic graph in which dependencies among connected nodes are linked to the observational data by the Markov assumption, so that the network of dependencies defines the set of the conditional probabilities of variables on their parents; the product of these conditional probabilities is said to be the factorization of the joint probability function of the variables

Health Services Research: A study of the use, organization and outcomes of health services.

Risk: probability of an event during a specified period of time.

Uncertainty: the lack of predictability of the outcome of an intervention at the level of an individual

Randomness: the lack of deterministic pattern in occurrence.

Interventional probability: the probability of outcome of a deliberately set intervention; a truncated form of factorization of the joint probability function in which the conditional probability for intervened variables is replaced with 1, and all other conditional probabilities are considered at set values of the intervened variables

Identifiability of causal effects: the causal effect of treatment is identifiable if all backdoor paths between treatment and outcome can be blocked. Stratifying on the sufficient set removes all confounding bias in estimates of the causal effect of X on Y.

Consistency assumption: the observed outcome for a subject equals to the subject’s potential outcome, as a function of intervention, when the intervention is set to the observed treatment.

Exchangeability: independence of the counterfactual outcome and the actual treatment

Back door criterion: identifies variables necessary to block all biasing paths that would result in association from X to Y; defines a set of variables S to be sufficient for the effect analysis of X on Y if no variable from S is a descendent of X and S blocks all backdoors paths from X to Y.

Attributable proportion: a proportion of cases that would not have occurred had the population not been exposed, relative to observed cases in the total population exposed (i.e. counterfactual risk when exposed with observed risk)

Excess proportion: a proportion of cases that would not have occurred had the population not been exposed, relative to cases when the entire population was exposed (i.e. counterfactual risk when exposed with counterfactual risk when unexposed)

Instrumental variable: a variable relative to the total effect of exposure on outcome if there is a set of measurements, unaffected by the exposure, such the instrument is conditionally independent of the outcome given the set and the instrument is not independent of the exposure conditional on the set

Propensity score: the probability of treatment given baseline characteristics

do(x) notation: represents an idealized experiment in which the variables in X is deliberately set to given values; in the causal diagram, this would correspond to removing all edges entering X, disconnecting directed influences of the parents of X and setting X at a value x.

Mediation: production of effect through a third variable

Indirect effect: the expected change in outcome in response to changing mediating variable to the level it would have attained had exposure changed by one unit, while holding exposure at initial level

Direct effect: the expected change in outcome in response to changing exposure while holding all other variables fixed

Selection bias: the lack of similarity of study groups in terms of other variables, resulting from choice of treatment on the basis of baseline characteristics

Marginal structural models: link the expected value of a counterfactual outcome to covariates; unlike in regression models, the expected value is not conditional on the covariates

Conditional independence: given value of a third variable, equivalence of the joint distribution of two variables to the product of distributions of each variable.

d-separation: variables X and Y are said to be d-separated, if all paths from X to Y are blocked. A path may be blocked by a set of variables S, if the path contains at least one arrow-emitting variable that is in S or if the path contains at least one collider that is not in S and has no descendent in S. A collider is defined as a variable on a path such that the path enters and leaves via arrow heads. X and Y are, therefore, d-separated conditional on S, if S blocks all paths between X and Y. X and Y are marginally independent when S contains the null set.

Markov assumption: conditional on its parents, a variable is independent of any other variable in the directed acyclic graph.