Causal Graphs

Marek J. Drużdżel

Politechnika Białostocka
Wydział Informatyki

m.druzdzel@pb.edu.pl
http://www.wi.pb.edu.pl/~druzdzel/

Overview

- Causal graphs
- Markov condition
- Intervention
- Experimentation
- Examples
- Concluding remarks

Causality and probability

The only reference to causality in a typical statistics textbook is: "correlation does not mean causation"
(if the textbook contains the word "causality" at all).

Many confusing substitute terms: "confounding factor," "latent variable," "intervening variable," etc.

What does correlation mean then (with respect to causality)?

The goal of experimental design is often to establish (or disprove) causation. We use statistics to interpret the results of experiments (i.e., to decide whether a manipulation of the independent variable caused a change in the dependent variable).

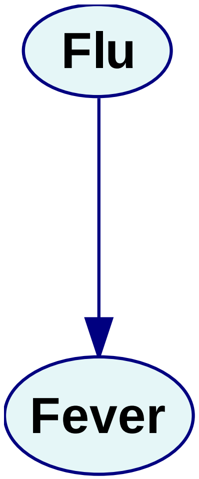
How are causality and probability actually related and what does one tell us about the other?

Not knowing this constitutes a handicap!

Causality and probability

Causality and probability are closely related and their relation should be made clear in statistics.

Probabilistic dependence is considered a necessary condition for establishing causation (is it sufficient?).



Flu and fever are correlated because flu may cause fever.
A cause can cause an effect but it does not have to. Causal connections result in probabilistic dependencies (or correlations in linear case).

This lecture introduces causal graphs as a tool for understanding research design.

We will use causal graphs to represent pictorially various experimental designs. Try to understand the basics!

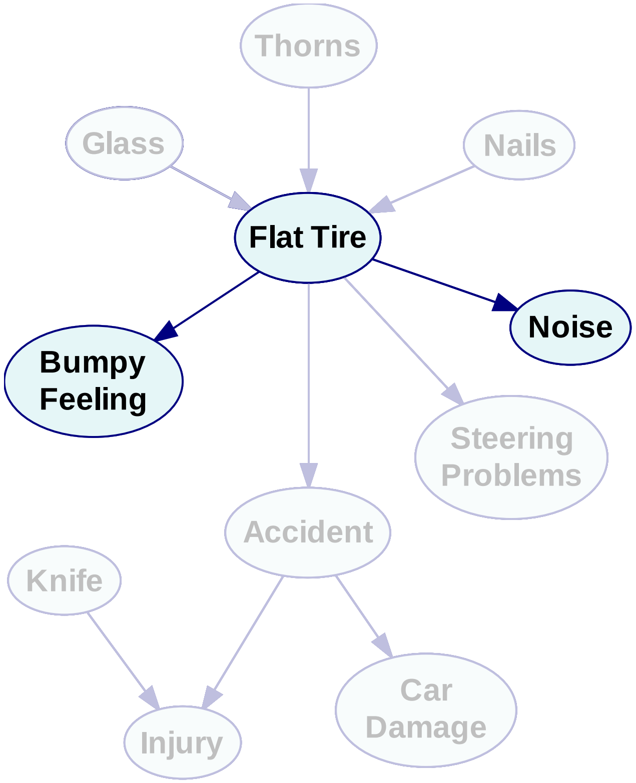
Causal graphs

Acyclic directed graphs (hence, no time and no dynamic reasoning) representing a snapshot of the world at a given time.

Nodes are random variables and arcs are direct causal dependencies between them.

Causal connections result in *correlation* (in general *probabilistic dependence*).

- glass on the road will be correlated with flat tire
- glass on the road will be correlated with noise
- bumpy feeling will be correlated with noise



Causal Markov condition

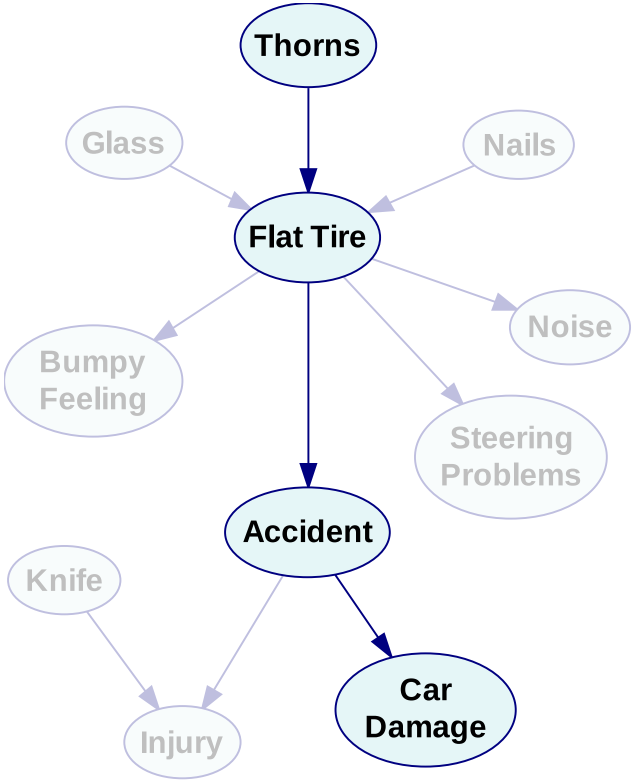
An axiomatic condition describing the relationship between causality and probability.

A variable in a causal graph is probabilistically independent of its non-descendants given its immediate predecessors.

Axiomatic, but used by almost everybody in practice and no convincing counter examples to it have been shown so far (at least outside the quantum world).

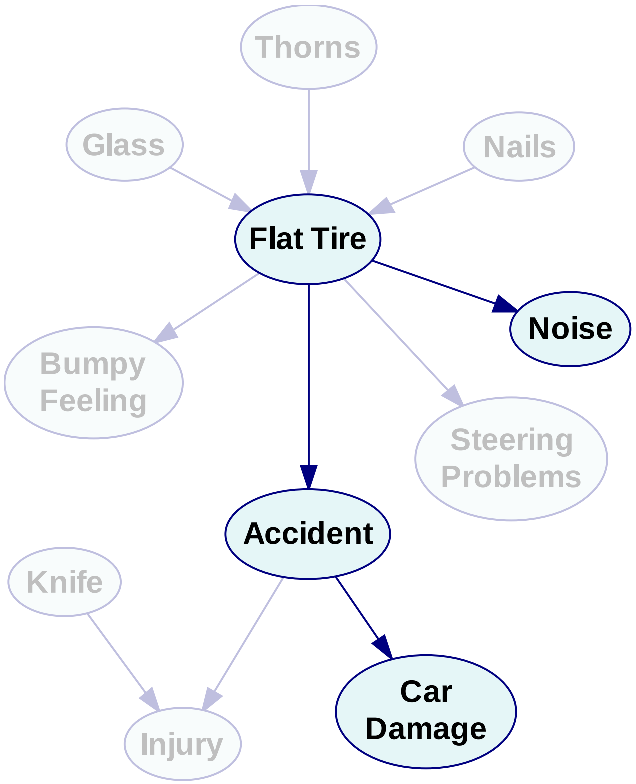
Markov condition: Implications

Variables A and B are probabilistically dependent if there exists a directed active path from A to B or from B to A:
Thorns on the road are correlated with car damage because there is a directed path from thorns to car damage.



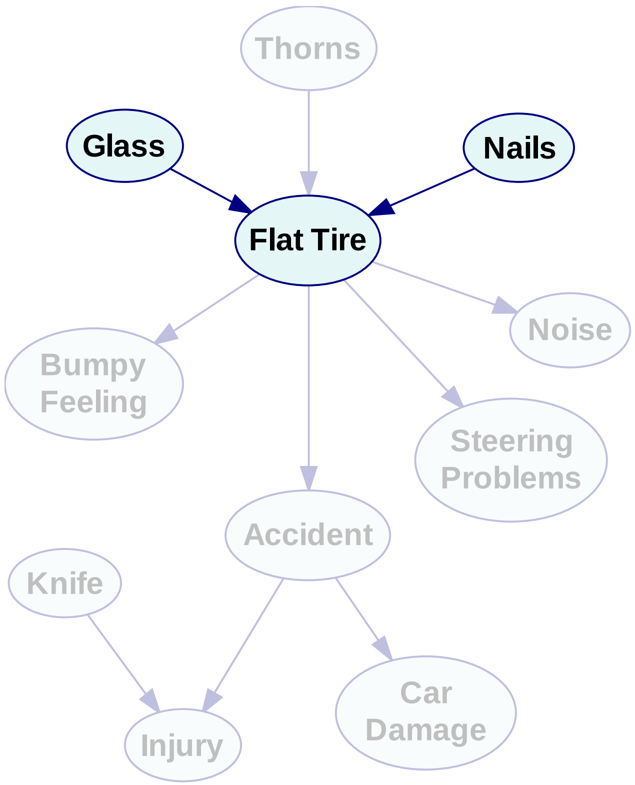
Markov condition: Implications

Variables A and B are probabilistically dependent if there exists a C such that there exists a directed active path from C to A and there exists a directed active path from C to B:
Car damage is correlated with noise because there is a directed path from flat tire to both (flat tire is a common cause of both).



Markov condition: Implications

Variables A and B are probabilistically dependent if there exists a D such that D is observed (conditioned upon) and there exists a C such that A is dependent on C and there exists a directed active path from C to D and there exists an E such that B is dependent on E and there exists a directed active path from E to D: Nails on the road are correlated with glass on the road given flat tire because there is a directed path from glass on the road to flat tire and from nails on the road to flat tire and flat tire is observed (conditioned upon).



Markov condition: Summary of implications

Variables A and B are probabilistically dependent if:

- **there exists a directed active path from A to B or there exists a directed active path from B to A**
- **there exists a C such that there exists a directed active path from C to A and there exists a directed active path from C to B**
- **there exists a D such that D is observed (conditioned upon) and there exists a C such that A is dependent on C and there exists a directed active path from C to D and there exists an E such that B is dependent on E and there exists a directed active path from E to D**

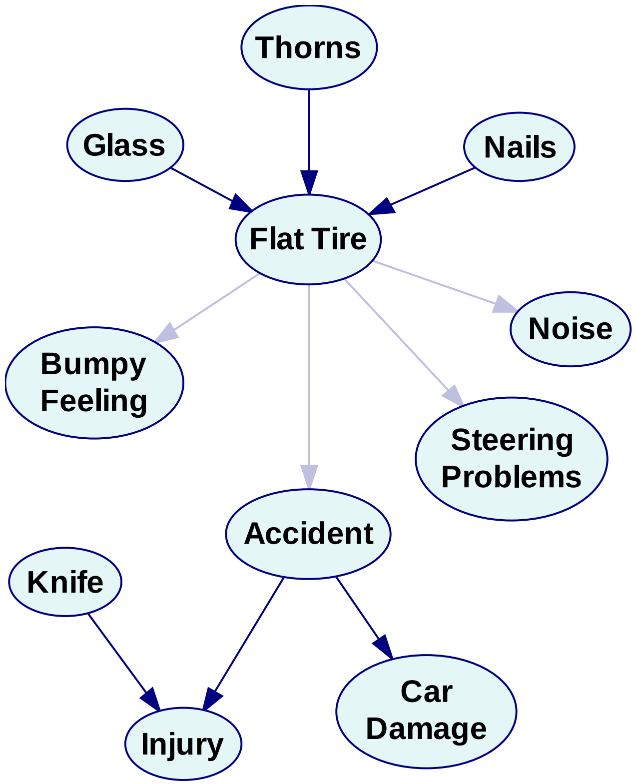
Markov condition: Conditional independence

Once we know all direct causes of an event E, the causes and effects of those causes do not tell anything new about E and its successors.

(also known as "screening off")

E.g.,

- **Glass and thorns on the road are independent of noise, bumpy feeling, and steering problems conditioned on flat tire.**
- **Noise, bumpy feeling, and steering problems become independent conditioned on flat tire.**

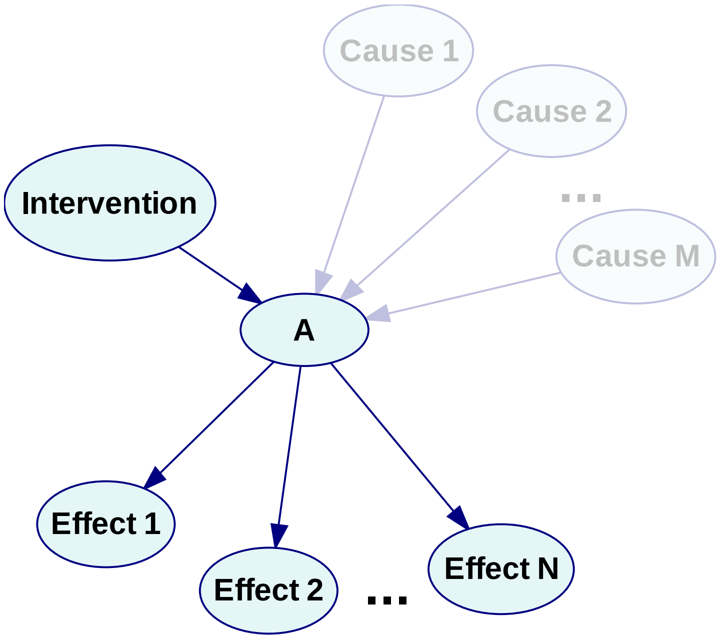


Intervention

Manipulation theorem [Spirtes, Glymour & Scheines 1993]:

Given an external intervention on a variable A in a causal graph, we can derive the posterior probability distribution over the entire graph by simply modifying the conditional probability distribution of A.

If this intervention is strong enough to set A to a specific value, we can view this intervention as the only cause of A and reflect this by removing all edges that are coming into A. Nothing else in the graph needs to be modified.



Intervention: Example

Suicide eliminates cancer as a cause of this brave samurai's death.

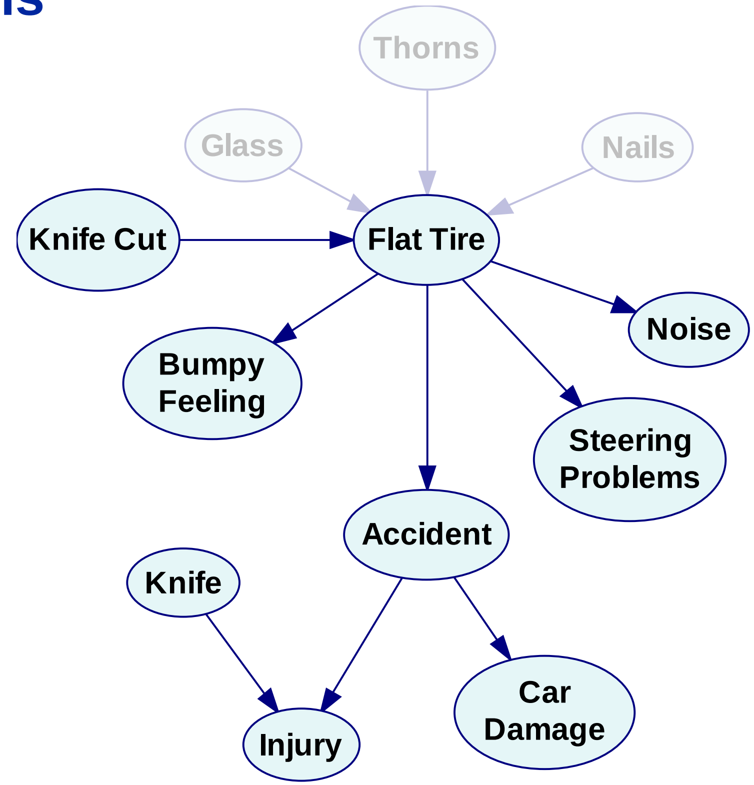
No?! Yes!!

Wooaah!

A cancer tumor!

Intervention: Example

Making the tire flat with a knife makes glass, thorns, nails, and what-have-you irrelevant to flat tire. The knife is the only cause of flat tire.



Experimentation

Empirical research is usually concerned with testing causal hypotheses.

Smoking and lung cancer are correlated.

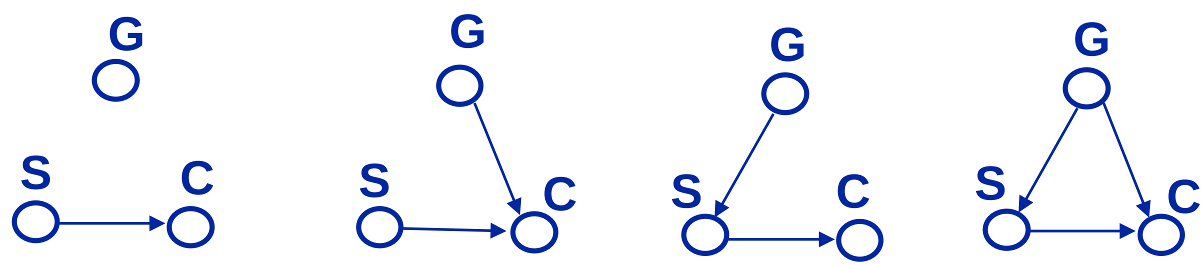
> **Can we reduce the incidence of lung cancer by reducing smoking?**
> **In other words: Is smoking a cause of lung cancer?**

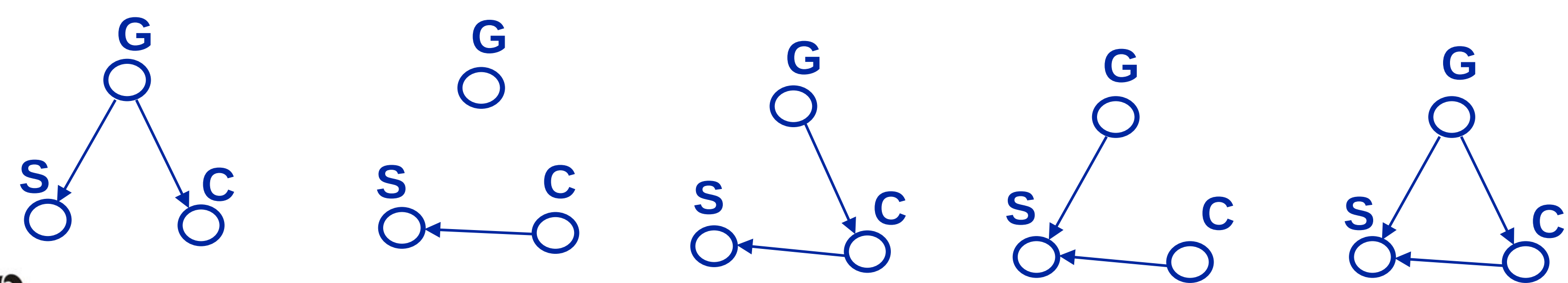
Each of the following causal structures is compatible with the observed correlation:

G = genetic factors

S = smoking

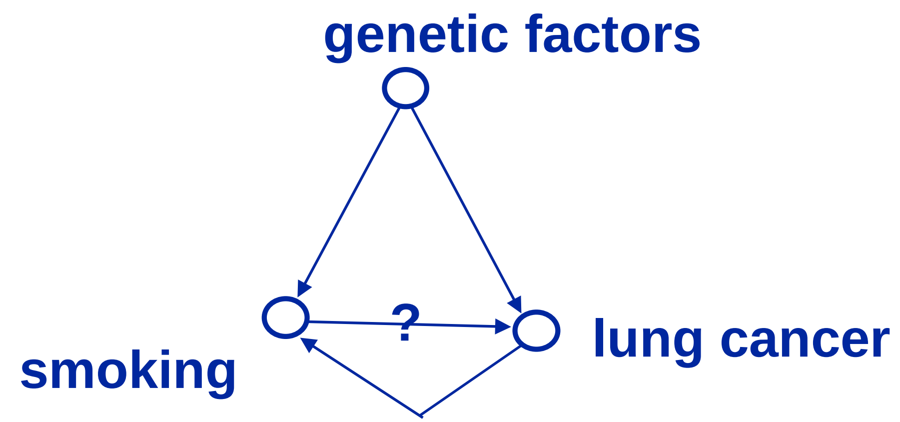
C = lung cancer





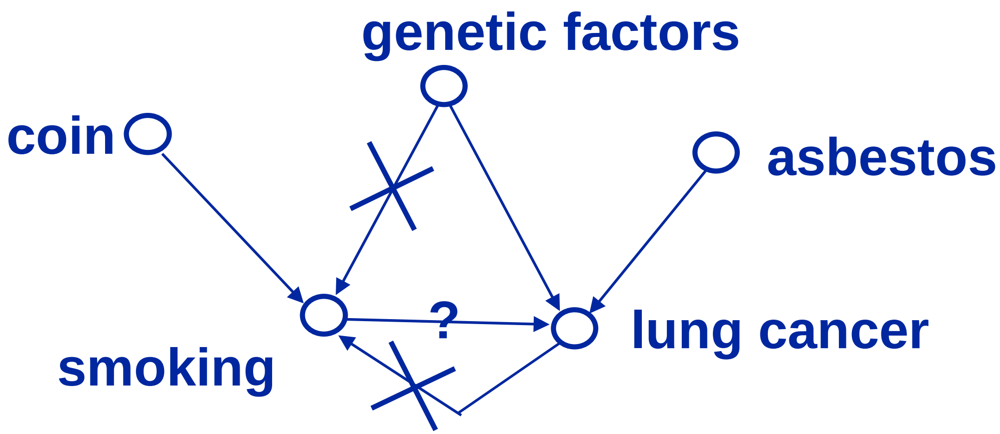
Selection bias

Observing correlation is in general not enough to establish causality.



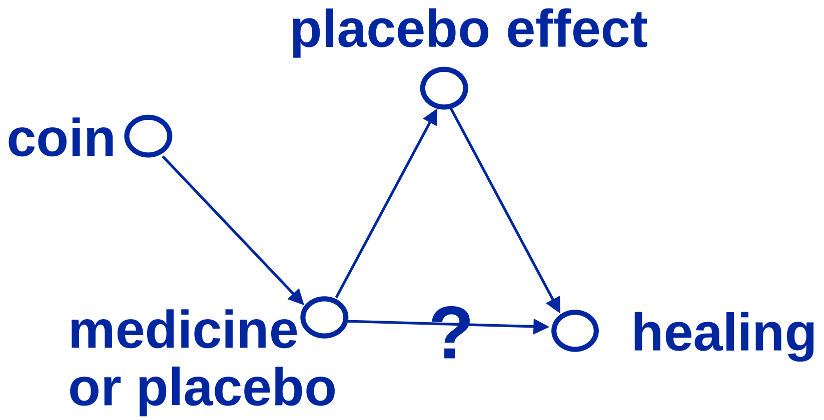
- If we do not randomize, we run the danger that there are common causes between smoking and lung cancer (for example genetic factors).
- These common causes will make smoking and lung cancer dependent.
- It may, in fact, also be the case that lung cancer causes smoking.
- This will also make them dependent without smoking causing lung cancer.

Experimentation



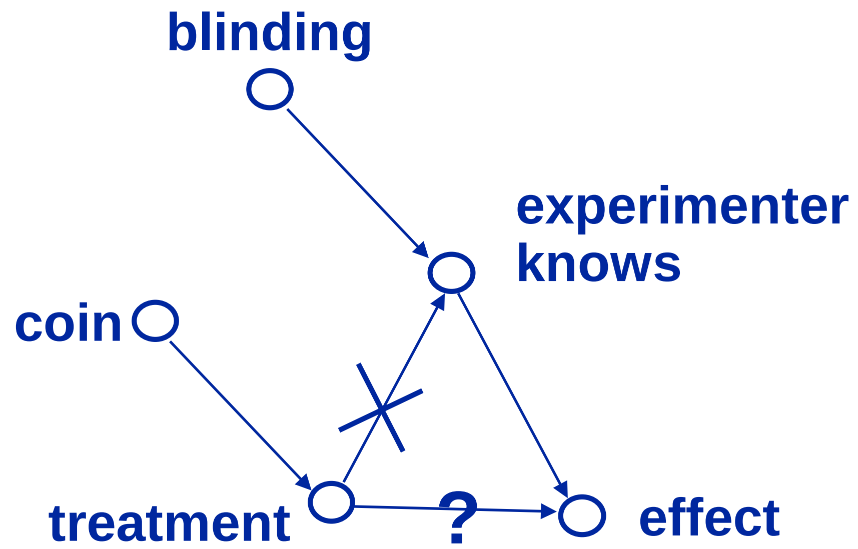
- In a randomized experiment, coin becomes the only cause of smoking.
- Smoking and lung cancer will be dependent only if there is a causal influence from smoking to lung cancer.
- If $Pr(C|S) \neq Pr(C|\sim S)$ then smoking is a cause of lung cancer.
- Asbestos will simply cause variability in lung cancer (add noise to the observations).

Placebo effect



By administering placebo to the control group we can compensate for an alternative path from medicine to healing (note that randomization will not break outgoing links).

Subject-experimenter effects



Blinding breaks the path from treatment to the subject-experimenter effects (note that randomization breaks only incoming links).

Concluding Remarks

Explicit representation of causal interactions in the design of empirical studies ...

- **makes the model clear and easier to comprehend**
- **explicates the idea that joint probability distribution from which the experimental measurements will be sampled is shaped by the mechanisms working in the system**
- **makes the goal consistent with the means: we want to discover causal mechanisms**
- **allows to understand the scope of experimental design**
- **exemplifies the importance of assumptions (they reduce the class of possible causal structures)**
- **shows that there are a variety of methods and inferences by which the structure of the system can be discovered (e.g., causal discovery from observation)**

Causality deserves a basic treatment in statistics.