Learning from Causation

Fundamental of Casual Inference and its Applications

Huang Xiao

Chair of IT Security (I20)
Department of Informatics
Technische Universität München

November 7, 2012

Overview

What is Causality?

A definition from Wikipedia

Causality (also referred to as causation) is the relationship between an event (*the cause*) and a second event (*the effect*), where the second event is understood as a consequence of the first.

What is Causality?

A definition from Wikipedia

Causality (also referred to as causation) is the relationship between an event (*the cause*) and a second event (*the effect*), where the second event is understood as a consequence of the first.

An example in real life : Does smoking cause lung cancer?

What is Causality?

A definition from Wikipedia

Causality (also referred to as causation) is the relationship between an event (*the cause*) and a second event (*the effect*), where the second event is understood as a consequence of the first.

An example in real life : Does smoking cause lung cancer?

Yes, it might be!

From Probabilistic View

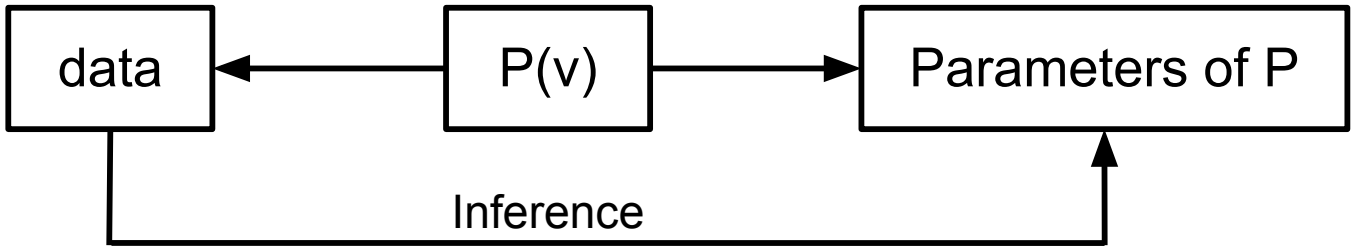
Problem: Does smoking cause lung cancer?

From Probabilistic View

Problem: Does smoking cause lung cancer?

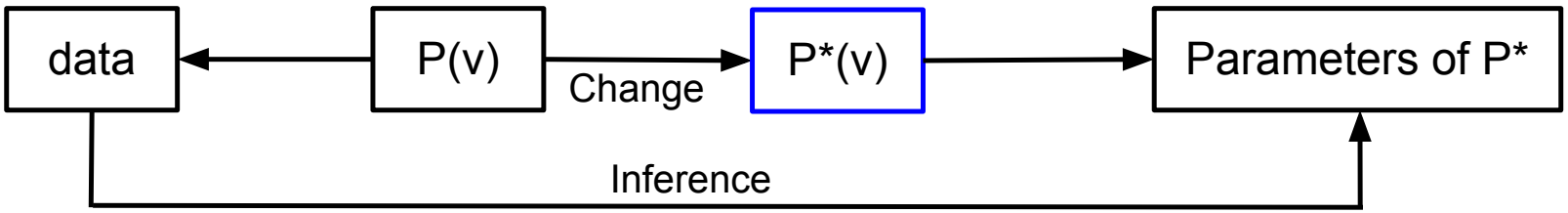
- Smoking does increase the probability of getting lung cancer.

Statistical Inference Overview



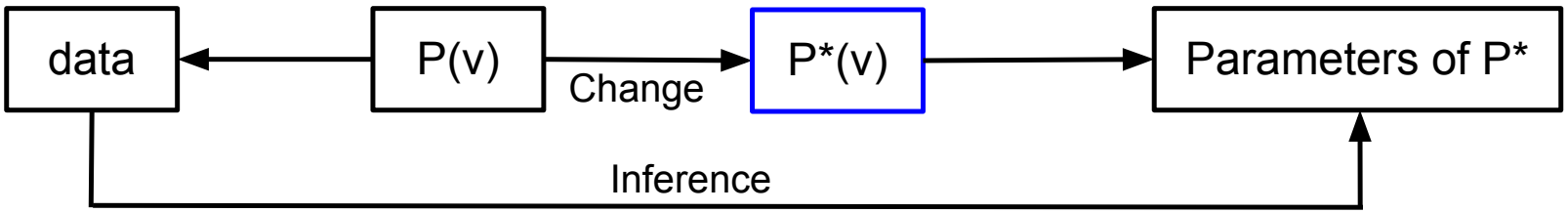
- Approximate an estimate of X given evidence e, namely, $Pr(X \mid e)$. E.g., Regression or Classification problems.
- Rejection of hypothesis, i.e., assert wether samples are from a certain distribution.
- Confidence interval, i.e., construct an interval based on dataset

Causal Inference Overview



- What if P has shifted itself to P^*?

Causal Inference Overview



- What if P has shifted itself to P^*?
- **Key factors:** Causes, Changes, and Invariants .
- Inference of P^* and reasoning of changes.

What makes Causal Inference interesting?

- Human understands the world in terms of causes and effects.
- Empirical science is about establishing causes.
- Causal inference gives a mathematical language for causal statements, and tools to solve causal problems formally.
- Alternative exercising to decision making, reasoning, etc.

Association

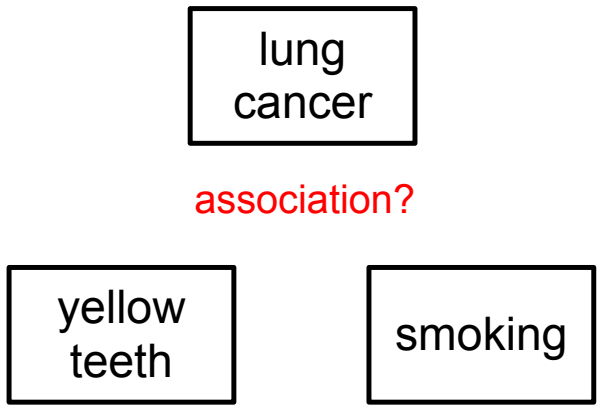
- Now we want to find out what causes lung cancer

Association

- Now we want to find out what causes lung cancer

		Lung cancer	
smoking	yellow teeth	yes	no
yes	yes	100	400
yes	no	100	400
no	yes	1	450
no	no	9	8540

Table: Data observations from 10000 people



Measurements of Association

To find out associations among variables

- Mutual information (Information theory)
- Pearson (linear) correlation
- Spearman's rho (rank correlation)
- Effect size between two variables
- Many others

Observations from Data

Obviously

- *yellow teeth* and *lung cancer* are associated.

Observations from Data

Obviously

- *yellow teeth* and *lung cancer* are associated.

But...

- Bleaching the teeth does not help reduce the probability of getting lung cancer.

Observations from Data

Obviously

- *yellow teeth* and *lung cancer* are associated.

But...

- Bleaching the teeth does not help reduce the probability of getting lung cancer.

Caution!

Correlation does not imply Causation

Overview

History

- Discuss the history of the problem.
- Describe context for the problem.
- Outline prior work on the problem.

Overview

Representation

Represent the problem in symbolic, graphic, or numeric format.

Mathematical formulas may be typeset:

$$\int_0^{\frac{\pi}{2}} \frac{1+\cos 2x}{2}\, dx$$

Overview

Methods and Tools

Discuss technical methods or tools required to formulate and solve the problem mathematically.

Theorem

If f is continuous on $[a, b]$, then

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

where F is any antiderivative of f, that is, a function such that $F' = f$.

Overview

Solution of the Problem

Present a solution of the problem, perhaps for a simple case, and indicate how the solution may be achieved in other cases.

Example

$$\int_0^{\frac{\pi}{2}} \frac{1+\cos 2x}{2}\, dx = \frac{\pi}{4}$$

Overview

Conclusion

Summarize the information presented in the talk.

- Problem statement
- Relevance
- Mathematical tools
- Solution

Overview

References


R. P. Boas, Can we make mathematics intelligible? *Amer. Math. Monthly*, **88** (1981), 727–731.


M. E. Page, A Brief Citation Guide for Internet Sources in History and the Humanities (Version 2.1), http://h-net.msu.edu/~africa/citation.html.