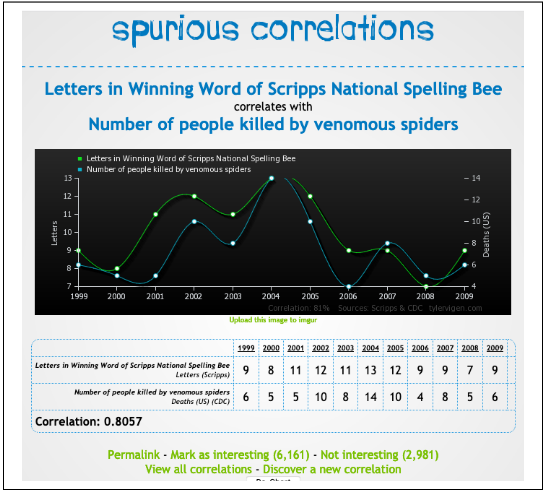
Correlation, Causality and the do-calculus

Dr. Paul Larsen

June 7, 2022

Why causality?

To avoid spurious correlations



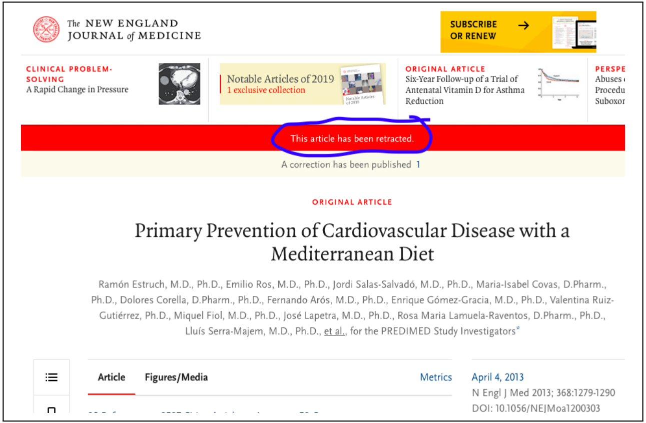
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Letters in Winning Word of Scripps National Spelling Bee *Letters (Scripps)*	9	8	11	12	11	13	12	9	9	7	9
Number of people killed by venomous spiders *Deaths (US) (CDC)*	6	5	5	10	8	14	10	4	8	5	6

Correlation: 0.8057

Tyler Vigen's Spurious Correlations

Why causality?

To estimate effects of interventions



The NEW ENGLAND JOURNAL of MEDICINE

SUBSCRIBE OR RENEW

CLINICAL PROBLEM-SOLVING
A Rapid Change in Pressure

Notable Articles of 2019
1 exclusive collection

ORIGINAL ARTICLE
Six-Year Follow-up of a Trial of Antenatal Vitamin D for Asthma Reduction

PERSPE
Abuses
Procedu
Suboxor

This article has been retracted.

A correction has been published 1

ORIGINAL ARTICLE

Primary Prevention of Cardiovascular Disease with a Mediterranean Diet

Ramón Estruch, M.D., Ph.D., Emilio Ros, M.D., Ph.D., Jordi Salas-Salvadó, M.D., Ph.D., Maria-Isabel Covas, D.Pharm., Ph.D., Dolores Corella, D.Pharm., Ph.D., Fernando Arós, M.D., Ph.D., Enrique Gómez-Gracia, M.D., Ph.D., Valentina Ruiz-Gutiérrez, Ph.D., Miquel Fiol, M.D., Ph.D., José Lapetra, M.D., Ph.D., Rosa Maria Lamuela-Raventos, D.Pharm., Ph.D., Lluís Serra-Majem, M.D., Ph.D., et al., for the PREDIMED Study Investigators*

Article   Figures/Media   Metrics

April 4, 2013
N Engl J Med 2013; 368:1279-1290
DOI: 10.1056/NEJMoa1200303

Article on PubMed

Interventions and causality

Ideal: Intervention + Multiverse → Causality

Examples:

- Medical treatment (e.g. kidney stone treatment)
- Social outcomes (e.g. university admissions)
- Business outcomes (e.g. click-through rate, hit rate)

In-practice:

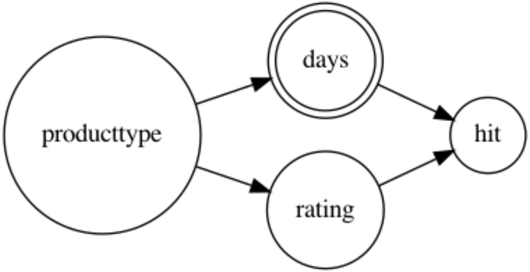
- Correlation: approximate multiverse by comparing intervention at t to result at $t - 1$
- Random population: approximate multiverse by splitting sample well
- A / B testing: random populations A / B + intervention in one

Formalizing interventions: the intuition of "do" for hit-rate

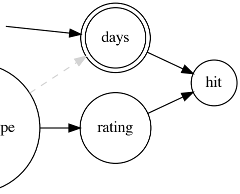
For business application, quantity of interest is effect of intervention / counterfactual

Not $P(\text{hit} = 1 | \text{days} = d)$ but $P(\text{hit} = 1 | \text{do}(\text{days} = d))$

$G =$



$G' = G_{\underline{\text{days}}} =$



Causality vs correlation mean different business decisions

Compute relative average treatment effect for different values of days:

$$\text{relative-ate}_G = \frac{P_G(\text{hit} = 1|\text{days} = d) - P_G(\text{hit} = 1|\text{days} = d+1)}{P_G(\text{hit} = 1|\text{days} = d)}$$

$$\text{relative-ate}_{G'} = \frac{P_G(\text{hit} = 1|\text{do}(\text{days} = d)) - P_G(\text{hit} = 1|\text{do}(\text{days} = d+1))}{P_G(\text{hit} = 1|\text{do}(\text{days} = d))}$$

$$= \frac{P_{G'}(\text{hit} = 1|\text{days} = d) - P_{G'}(\text{hit} = 1|\text{days} = d+1)}{P_{G'}(\text{hit} = 1|\text{days} = d)}$$

from-d	to-d	ate-given	ate-do
0	1	0.170153	0.297187
1	2	0.252329	0.395158
2	3	0.473538	0.102707

Reality check and wrap-up

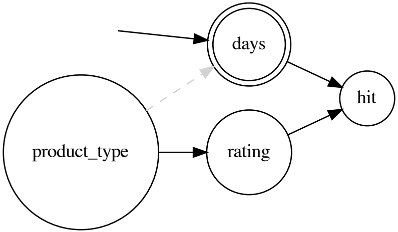
- The do-calculus models interventions better than correlation / conditionals, but what about model misspecification?
- Causal reasoning mitigates risk of outsourcing thinking to correlations

Appendices

For more context and code samples, see the risk-ai-workshop repo and slides.

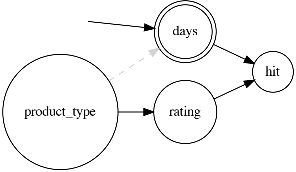
Formalizing interventions: the intuition of "do"

First, find quantities unchanged between G and $G' = G_{\underline{\text{days}}}$



$$P_{G'}(\text{producttype} = p, \text{rating} = r)$$
$$= P_G(\text{producttype} = p, \text{rating} = r) \quad (1)$$
$$P_{G'}(\text{hit} = 1|\text{producttype} = p, \text{rating} = r)$$
$$= P_G(\text{hit} = 1|\text{producttype} = p, \text{rating} = r) \quad (2)$$

Formalizing interventions: the intuition of "do"



$$
\begin{aligned}
&P(\text{hit} = 1 | \text{do}(\text{days}) = d) \\
&\quad = P_{G'}(\text{hit} = 1 | \text{days} = d), \text{ by definition} \\
&\quad = \sum_{p,r} P_{G'}(\text{hit} = 1 | \text{days} = d, \text{producttype} = p, \text{rating} = r) \\
&\qquad\qquad P_{G'}(\text{producttype} = p, \text{rating} = r | \text{days} = d), \text{ by total probability} \\
&\quad = \sum_{p,r} P_{G'}(\text{hit} = 1 | \text{days} = d, \text{producttype} = p, \text{rating} = r) \\
&\qquad\qquad P_{G'}(\text{producttype} = p, \text{rating} = r), \text{ by substitution} \\
&\quad = \sum_{p,r} P_{G}(\text{hit} = 1 | \text{days} = d, \text{producttype} = p, \text{rating} = r) \\
&\qquad\qquad P_{G}(\text{producttype} = p, \text{rating} = r), \text{ our } \textit{adjustment} \text{ formula}
\end{aligned}
$$

References: Judea Pearl et. al, Causal Inference in Statistics, Christopher Prohm, Causality and Function Approximation

Judea Pearl's Rules of Causality

Let X, Y, Z and W be arbitrary disjoint sets of nodes in a DAG G. Let $G_{\underline{X}}$ be the graph obtained by removing all arrows pointing into (nodes of) X. Denote by $G_{\overline{X}}$ the graph obtained by removing all arrows pointing out of X. If, e.g. we remove arrows pointing out of X and into Z, we the resulting graph is denoted by $G_{\underline{X}\overline{Z}}$

Rule 1: Insertion / deletion of observations

$$P(y|\mathrm{do}(x), z, w) = P(y|\mathrm{do}(x), w) \text{ if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}$$

Rule 2: Action / observation exchange

$$P(y|\mathrm{do}(x), \mathrm{do}(z), w) = P(y|\mathrm{do}(x), z, w) \text{ if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}$$

Rule 3: Insertion / deletion of actions

$$P(y|\mathrm{do}(x), \mathrm{do}(z), w) = P(y|\mathrm{do}(x), w) \text{ if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\overline{Z(W)}}},$$

where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\underline{X}}$.

Special cases of the causal rules

By judicious setting of sets of nodes to be empty, we obtain some useful corollaries of the causal rules.

Rule 1': Insertion / deletion of observations, with $W = \emptyset$

$$P(y|\mathrm{do}(x), z) = P(y|\mathrm{do}(x)) \text{ if } (Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}}}$$

Rule 2': Action / observation exchange, with $X = \emptyset$

$$P(y|\mathrm{do}(z), w) = P(y|z, w) \text{ if } (Y \perp\!\!\!\perp Z|W)_{G_{\underline{Z}}}$$

Rule 3': Insertion / deletion of actions, with $X, W = \emptyset$

$$P(y|\mathrm{do}(z)) = P(y) \text{ if } (Y \perp\!\!\!\perp Z)_{G_{\overline{Z}}}$$

Special cases of the causal rules

By judicious setting of sets of nodes to be empty, we obtain some useful corollaries of the causal rules.

Rule 1': Insertion / deletion of observations, with $W = \emptyset$

$$P(y|\mathrm{do}(x), z) = P(y|\mathrm{do}(x)) \text{ if } (Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}}}$$

Rule 2': Action / observation exchange, with $X = \emptyset$

$$P(y|\mathrm{do}(z), w) = P(y|z, w) \text{ if } (Y \perp\!\!\!\perp Z|W)_{G_{\underline{Z}}}$$

Rule 3': Insertion / deletion of actions, with $X, W = \emptyset$

$$P(y|\mathrm{do}(z)) = P(y) \text{ if } (Y \perp\!\!\!\perp Z)_{G_{\overline{Z}}}$$

$\Longrightarrow$ d-separation + causal rules = *adjustment formulas*: do queries as normal queries.

Causality vs correlation mean different business decisions

Quantity of interest: *average treatment effect* or *ATE*

$P(\text{hit} = 1 | \text{days} = d)$

	hit
days	
0	0.532706
1	0.442064
2	0.330519
3	0.174006

$P(\text{hit} = 1 | \text{do}(\text{days} = d))$

	prob
days	
0	0.565343
1	0.397330
2	0.240322
3	0.215639