

# STA 640 — Causal Inference

## Chapter 1. Background and Potential Outcome Framework

Fan Li

Department of Statistical Science  
Duke University

# Causality in Ancient Greek Philosophy

*I would rather discover one **causal law** than be King of Persia.*  
— Democritus

*We have knowledge of a thing only when we have grasped **its cause**.*  
— Aristotle, *Posterior Analytics*

*We do not have knowledge of a thing until we have grasped **its why**, that is to say, **its cause**.*  
— Aristotle, *Physics*

# Questions on Causation

- ▶ Relevant questions about causation:
  - ▶ the philosophical meaningfulness of the notion of causation
  - ▶ deducing the causes of a given effect
  - ▶ understanding the details of causal mechanism
- ▶ In this class we focus on measuring the effects of causes – a place where statistics, which is concerned with measurement, has most contributions to make.

# Association vs. Causation

- ▶ The research questions that motivate most studies in statistics-based sciences are causal in nature.
- ▶ Standard statistical analysis is to infer *associations* among variables, based on which may do some *prediction*
- ▶ Causal analysis is one step further: it is about *counterfactual prediction*, predict what would have happened to the same units/subjects had they were exposed to a different (counterfactual) condition
- ▶ In most cases, *association does not imply causation*

# Causal Inference

- ▶ *How to make the leap from association to causation?*
- ▶ Key: *causal assumptions* - structural and/or modeling
- ▶ Causal inference is about
  1. build a framework and define causal effects under general scenarios
  2. specify assumptions under which one can declare/identify causation from association
  3. assess the sensitivity to the causal assumptions and find ways to mitigate

# Notations

- ▶ Treatment (e.g. intervention, exposure)  $Z$ : for illustration we will mostly focus on binary treatments
- ▶ Outcome (e.g. disease status)  $Y$
- ▶ Observed covariates or confounders  $X$
- ▶ Unobserved covariates or confounders  $U$
- ▶ Examples of question of interest
  - ▶ Causal effect of exposure on disease
  - ▶ Comparative effectiveness research: whether one drug or medical procedure is better than the other
  - ▶ Program evaluation in economics and policy

# Three Type of Biases

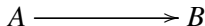
(Hernan and Robins, 2020)

- ▶ Types of biases
  - ▶ Confounding bias
  - ▶ Selection (sample) bias
  - ▶ Measurement bias
- ▶ All three have strong connections to *missing data*
- ▶ Will mostly address *confounding bias* and (if time allows) *selection bias* in this course
  - ▶ germane to *observational studies*, but also in *randomized trials*

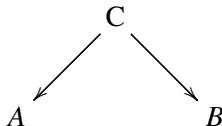
# Confounding

- ▶ Confounding (or common cause) is the main complication/hurdle between association and causation
- ▶ Two Directed Acyclic Graphs

(i) Causal relationship:



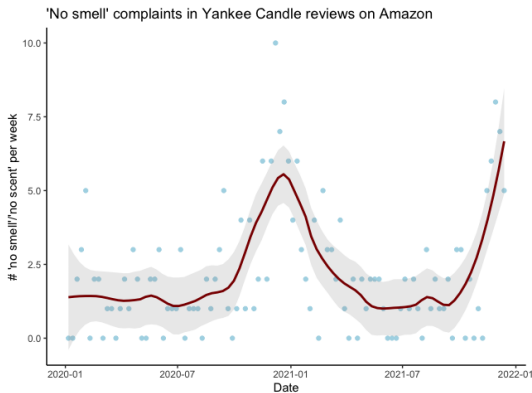
(ii) Confounding:





# Examples of Confounding

- ▶ Ice cream consumption and number of people drown.  
Confounder: temperature
- ▶ Complaints of “no smell” of Yankee Candle in Amazon reviews vs. time.



# Another COVID-19 example

## COVID Vaccine Hesitancy and Risk of a Traffic Crash

DECEMBER 02, 2021 | 1:30PM - 3:00PM

---

**Speaker:** Donald Redelmeier, Professor of Medicine at the University of Toronto; Canada Research Chair in Medical Decision Sciences; Director of Clinical Epidemiology at Sunnybrook Health Sciences Centre; Senior Scientist at the Institute for Clinical Evaluative Studies in Ontario; Staff physician in the Division of General Internal Medicine at Sunnybrook Hospital

### Abstract:

COVID vaccine hesitancy is a reflection of judgment, reasoning, and other psychological influences that may also contribute to traffic safety. We tested whether COVID vaccine hesitancy was associated with an increased risk of a serious traffic crash.

A total of 11,270,763 adults were identified, of whom 16% had not received a COVID vaccine and 84% had received a COVID vaccine. Those who had not received the vaccine accounted for a disproportionate number of crashes, equivalent to a significant increased traffic risk. The association between a lack of COVID vaccine and increased traffic risks extended to diverse patient subgroups, persisted after adjusting for measured baseline differences, applied across a spectrum of crash severity, and was similar to the relative risk associated with a diagnosis of sleep apnea.

We suggest that COVID vaccine hesitancy is associated with a significant increased risk of a serious traffic crash. An awareness of this counter-intuitive finding might contribute to more public support for the COVID vaccine.

# Simpson's paradox: Kidney Stone Treatment

(Charig et al., BMJ, 1986)

- ▶ An extreme example of confounding is Simpson's paradox: confounder reverses the sign of the correlation between treatment and outcome
- ▶ Compare the success rates of two treatments for kidney stones
- ▶ Treatment A: open surgery; treatment B: small puncture

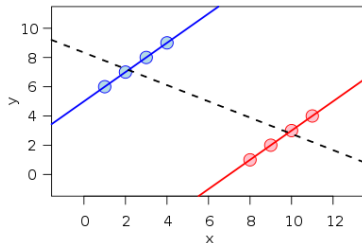
	Treatment A	Treatment B
Small stones	<b>93%</b> (81/87)	87% (234/270)
Large stones	<b>73%</b> (192/263)	69% (55/80)
Both	78% (273/350)	<b>83%</b> (289/350)

- ▶ What is the confounder here? Severity of the case

# Simpson's paradox or Yule-Simpson effect

(K Pearson et al., 1899; Yule, 1903; Simpson, 1951)

- ▶ Simpson's paradox: a trend appears in different groups of data but disappears or reverses when these groups are combined



- ▶ Mathematically, it is about conditioning
- ▶ A special case of the ecological fallacy
- ▶ It may or may not be due to confounding
- ▶ Another well-known example is the Berkeley admission gender bias (Bickel et al., Science, 1976)

# A Classic Example—Smoking and Lung Cancer

Doll and Hill (1950 BMJ)



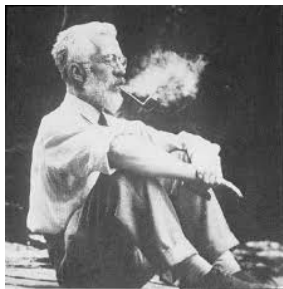
**Figure:** Sir Austin  
Bradford Hill  
(1897–1991)

- ▶ Smoking-cancer association
- ▶ Case-control study of lung cancer
- ▶ Risk ratio  $\approx$  odds ratio, is roughly 9 even after adjusting for observed covariates:

$$RR_{ZY}^{\text{obs}} = \frac{\Pr(Y = 1 \mid Z = 1)}{\Pr(Y = 1 \mid Z = 0)} \approx 9$$

- ▶ Does smoking cause lung cancer?
- ▶ Box (2013) stopped smoking after seeing Doll and Hill (1950)

# A Classic Example—Smoking and Lung Cancer



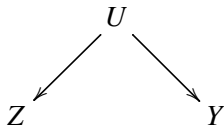
**Figure:** Sir Ronald  
Aylmer Fisher  
(1890–1962)

- ▶ Association does not imply causation
- ▶ “Common cause” (Reichenbach 1956, Fisher 1957 BMJ)
- ▶ Fisher (1957 BMJ):  
*cigarette-smoking and lung cancer, though not mutually causative, are both influenced by a **common cause**, in this case the **individual genotype**.*

# Cornfield Inequality for Smoking and Lung Cancer

Cornfield et al. (1959 JNCI)

Common cause hypothesis



- ▶ Smoking  $Z$
- ▶ Lung cancer  $Y$
- ▶ Genetic factor  $U$

- ▶ Assume Fisher is right
- ▶ The smoking-gene association must satisfy:

$$RR_{ZU} \geq RR_{ZY} \approx 9$$

- ▶ Such a genetic confounder is too strong to be realistic.
- ▶ Association must be due to causal.
- ▶ We will revisit this example later.

# Frameworks for Causal Inference

- ▶ The purpose is to construct a model or a framework that is complex enough to allow us to formalize basic intuitions concerning cause and effect
- ▶ Two commonly used frameworks
  - ▶ The *potential outcome framework*, also known as the counterfactual framework, or the Neyman-Rubin Causal Model (Neyman, 1923; Rubin, 1974; Imbens and Rubin, 2015; Hernan and Robins, 2020)
  - ▶ The *causal diagram* framework (Pearl, 2009)
  - ▶ Mathematically the two frameworks are connected (Richardson and Robins, 2013), but each has different established goals, tools and applicable areas
  - ▶ This class focuses on the potential outcome framework, and will occasionally draw directed acyclic graphs (DAGs) for simple illustration



# Potential Outcome Framework: Basic Setup

- ▶ Unit: The person, place, or thing upon which a treatment will operate, at a particular time
- ▶ *Target Population*: a well-defined **population** of units whose outcomes are going to be compared
- ▶ Data: a random sample of  $N$  units from a target population
- ▶ Treatment: an intervention, the effects of which (on some outcomes of the units) the investigator wishes to assess relative to no intervention (i.e., the control)
- ▶ For simplicity, consider a treatment with two levels:  $z = 0, 1$
- ▶ For each unit  $i$ , we observe the (binary) treatment status  $Z_i (= 0, 1)$ , a vector of  $p$  covariates  $X_i = (X_{i1}, \dots, X_{ip})$ , and an outcome  $Y_i^{\text{obs}}$  (or simply denoted as  $Y_i$  later)

# Potential Outcome Framework: Basic Setup

- ▶ **Potential outcomes**: the values of a unit's outcome (hypothetically) under treatment or control
- ▶ Each unit  $i$  has two potential outcomes:  $Y_i(0), Y_i(1)$ . Note: Such notation implicitly relies on SUTVA ( more later)
- ▶ **Causal effect**: comparison between the potential outcomes under treatment and under control for *the same unit or a common set of units*. For example:
  - ▶ **Individual causal effect (ITE)**:  $Y_i(1) - Y_i(0)$
  - ▶ **Average treatment effect (ATE)**:  $\tau = \mathbb{E}[Y_i(1) - Y_i(0)]$ , ITE averaged over a target population

# The Fundamental Problem of Causal Inference

Holland, 1986, JASA

- ▶ The fundamental problem of causal inference: We can observe at most one of the potential outcomes for each unit, the other(s) are missing/counterfactual
- ▶ Causal inference under the potential outcome framework is essentially a missing data problem
- ▶ To identify causal effects from observed data, under any mathematical framework, one must make assumptions (structural or/and stochastic)

# Multiple Units

- ▶ We must rely on multiple units exposed to different treatments to make causal inferences
  - ▶ observe the same physical object subject to different treatment levels at different points in time
  - ▶ observe different physical units at the same time
- ▶ By itself, however, the presence of multiple units does not solve the problem of causal inference

You take: I take:	Asp Asp	Not Not	Asp Not	Not Asp
Unit				
1=you	$Y_1([A, A]) = 0$	$Y_1([N, N]) = 100$	$Y_1([A, N]) = 50$	$Y_1([N, A]) = 75$
2=me	$Y_2([A, A]) = 0$	$Y_2([N, N]) = 100$	$Y_2([A, N]) = 100$	$Y_2([N, A]) = 0$

# SUTVA

Assumption 1: **Stable Unit Treatment Value Assumption (SUTVA)**  
(Rubin, 1980):

- ▶ SUTVA includes two assumptions: (1) no interference, (2) no different versions of a treatment (also known as **consistency**, Cole and Frangakis, 2009)
- ▶ Seems trivial, actually very strong assumptions
- ▶ Under SUTVA, each unit has only two potential outcomes  $Y_i(1)$ ,  $Y_i(0)$ , and
  - ▶ If  $Z_i = 1$  then  $Y_i = Y_i(1)$
  - ▶ If  $Z_i = 0$  then  $Y_i = Y_i(0)$
  - ▶ Equivalently:  $Y_i = Z_i Y_i(1) + (1 - Z_i) Y_i(0)$
- ▶ **SUTVA connects the intervention we see ( $Z$ ), with the causal intervention of interest ( $z$ )**

# Interference

- ▶ **Interference:** the potential outcome  $Y_i(z)$  for an individual  $i$  depends on what treatment other people receive
- ▶ Examples: vaccination, advertising, infectious diseases, social networks, agricultural experiments
- ▶ There are lots of possible  $Y_i(z)$ , *depending on what happens to other people*
- ▶ When in the presence of interference, other assumptions required for causal inference (e.g., Rosenbaum 2007; Hudgens and Hollaran 2008) – a hot topic lately

# Multiple Versions of Treatment

- ▶ **Multiple Versions of Treatment:** Sometimes action  $Z$  doesn't have a clear meaning, as it has many versions
- ▶ Examples
  - ▶ Dose
  - ▶ Surgeon
  - ▶ Treatment history
  - ▶ Standard of care (as a reference)
- ▶ There are lots of possible  $Y_i(Z)$ , *depending on what version gets selected*

# Causal Estimands

- ▶ **Causal estimands:** functions of the potential outcomes

- ▶ **Individual treatment effect (ITE):**

$$\tau_i = Y_i(1) - Y_i(0)$$

- ▶ **Conditional average treatment effect (CATE):**

$$\tau(x) = \mathbb{E}[Y_i(1) - Y_i(0)|X = x]$$

Note: ITE and CATE are different estimands, but often conflated in the literature

- ▶ **Average treatment effect (ATE):**

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}_x[\tau(x)].$$

- ▶ **Average treatment effect for the treated units (ATT):**

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)|Z_i = 1].$$

- ▶ **Average treatment effect for the control units (ATC):**

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)|Z_i = 0].$$



# Causal Estimands: Ratio

- ▶ Individual treatment effect in ratio:  $r_i = Y_i(1)/Y_i(0)$
- ▶ Population treatment effect in ratio:  $r = \mathbb{E}[Y_i(1)]/\mathbb{E}[Y_i(0)]$
- ▶ Relative causal effect (lift):  $l = \mathbb{E}[Y_i(1) - Y_i(0)]/\mathbb{E}[Y_i(0)]$
- ▶ Continuous outcomes are common in theoretical literature and applications social sciences
- ▶ Binary outcomes are more common in medical and health studies
  - ▶ Note: ratio is **not** an average of individual causal effect unless the effect is homogeneous

# Causal Estimands: Binary outcomes

- ▶ For binary outcomes, we have specific estimands and terminology
- ▶ Let  $\mu_1 = \Pr(Y(1) = 1)$  and  $\mu_0 = \Pr(Y(0) = 1)$ , often known as causal risks in medical research.
- ▶ Three estimands:
  - ▶ Causal risk difference (equivalent to ATE):  $\tau_{RD} = \mu_1 - \mu_0$
  - ▶ Causal risk ratio:  $\tau_{RR} = \mu_1 / \mu_0$
  - ▶ Causal odds ratio:  $\tau_{OR} = \frac{\mu_1 / (1 - \mu_1)}{\mu_0 / (1 - \mu_0)}$ .
- ▶ These estimands are **different** from their observed counterpart (denote  $m_1 = \Pr(Y^{obs} = 1 | Z = 1)$  (  $m_0 = \Pr(Y^{obs} = 1 | Z = 0)$  )
  - ▶ Observed risk difference:  $m_1 - m_0$
  - ▶ Observed risk ratio:  $m_1 / m_0$
  - ▶ Observed odds ratio:  $\frac{m_1 / (1 - m_1)}{m_0 / (1 - m_0)}$ .
- ▶ Take-home question: why these two are different?

# Potential Outcomes

- ▶ Causal effects defined by **potential outcomes**, not **model parameters**
- ▶ Why bother to introduce the **hypothetical** potential outcomes?
  - ▶ separation of the intervention and quantities of interest
  - ▶ randomness comes solely from the treatment
  - ▶ **analysis of experiments is driven by design**
- ▶ Obviously these estimands are not identifiable without further assumptions
- ▶ So, what assumptions do we need?

# Perfect Doctor

<u>Potential Outcomes</u>			<u>Observed Data</u>			
$Y(0)$	$Y(1)$		$Z$	$Y(0)$	$Y(1)$	$Y^{obs}$
13	14		1	?	14	14
6	0		0	6	?	6
4	1		0	4	?	4
5	2		0	5	?	5
6	3		0	6	?	6
6	1		0	6	?	6
8	10		1	?	10	10
8	9		1	?	9	9
True averages	7	5	Observed averages	5.4	11	

## Perfect Doctor: Comments

- ▶ The simple difference-in-means estimator does not return a valid estimate of the true causal effect. Why?
- ▶ Key: the **assignment mechanism** (Rubin, 1978) – the probabilistic rule that decides which unit gets assigned to which treatment

$$p(Z_i = 1 | X_i, Y_i(0), Y_i(1))$$

- ▶ Is the assignment mechanism in the perfect doctor example random? (in fact, what does random assignment mean?)
- ▶ The assignment depends on both  $Y_i(0)$  and  $Y_i(1)$  for each unit (of course, in reality we will unlikely to have any such perfect doctor)
- ▶ The key identifying assumptions in causal inference are on the **assignment mechanism**

# Unconfounded Assignment

- ▶ *Unconfounded Assignment*: an assignment is unconfounded if the assignment mechanism does not depend on the potential outcomes (and only on the pre-treatment covariates):

$$p(Z_i = 1|X_i, Y_i(0), Y_i(1)) = p(Z_i = 1|X_i)$$

- ▶  $p(Z_i = 1|X_i)$  is also known as the propensity score (Rosenbaum and Rubin, 1983)
- ▶ Is the assignment mechanism unconfounded
  - ▶ in randomized experiments?
  - ▶ in the perfect doctor example?
  - ▶ in the smoking and lung cancer study?

# Unconfounded Assignment

- ▶ *Unconfounded Assignment*: an assignment is unconfounded if the assignment mechanism does not depend on the potential outcomes (and only on the pre-treatment covariates):

$$p(Z_i = 1|X_i, Y_i(0), Y_i(1)) = p(Z_i = 1|X_i)$$

- ▶  $p(Z_i = 1|X_i)$  is also known as the propensity score (Rosenbaum and Rubin, 1983)
- ▶ Is the assignment mechanism unconfounded
  - ▶ in randomized experiments? Yes, and it is **known** and **controlled** by investigators
  - ▶ in the perfect doctor example? No, because it depends on both  $Y_i(1)$  and  $Y_i(0)$
  - ▶ in the smoking and lung cancer study? We do not know, as in most observational studies, the assignment mechanism is usually **unknown** and **uncontrolled**. To make causal inference, we have to make (often strong) assumptions about the assignment mechanism

# Ignorable Assignment

- ▶ Unconfoundedness (or unconfounded assignment) is often also called **ignorability** (or ignorable assignment) in the literature
- ▶ The terminology corresponds to that of missing data mechanism in the missing data literature (Rubin, 1976).
- ▶ The meaning of “ignorability” becomes more apparent when we talk about Bayesian inference of causal effect: means that the assignment mechanism drops out from the data likelihood in estimating the causal effects
- ▶ In the context of sequential treatment, ignorability is more complex, allowing conditioning on observed intermediate outcomes  $p(Z_i = 1 | X_i, Y_i^{obs})$  (more in Section 6)



# Potential Outcome Framework: Comments

- ▶ No causation without manipulation - “cause” must be (hypothetically) manipulatable, e.g., intervention, action, treatment
- ▶ Gender, time and age are not well defined “causes” under the PO framework
- ▶ Three integral components (Rubin, 1978):
  - ▶ potential outcomes corresponding to the various levels of a treatment or manipulation (no causation without manipulation)
  - ▶ assignment mechanisms
  - ▶ a (Bayesian) model for the science (the potential outcomes and covariates)
- ▶ Rooted in the statistical work on randomized experiments by Fisher (1918, 1925) and Neyman (1923), as extended by Rubin (1974, 1978) and subsequently by many others to apply to nonrandomized studies and other forms of inference.

# Outline of the Class

The class material is (mostly) organized by the classification of assignment mechanisms

- ▶ **Randomized Experiments:** the assignment mechanism is known, controlled, and random (stronger than unconfounded)
- ▶ **Observational studies:** the assignment mechanism is unknown and uncontrolled, but often assumed to be unconfounded conditional on observed covariates or unobserved quantities
  - ▶ Cross-sectional data: treatment at one time point
  - ▶ Longitudinal data: treatment at multiple time points (sequentially ignorable)
  - ▶ Panel data: treatment at one time point, comparative case studies (often different assumptions than unconfoundedness)
- ▶ **Natural/quasi-experiments:** usually involve some type of unconfoundedness assumptions (can based on unobserved quantities)
  - ▶ instrumental variables
  - ▶ principal stratification (latent ignorable)
  - ▶ regression discontinuity (local randomized)

# Methods and Modes of Inference

- ▶ Two overarching methods
  - ▶ Imputation: impute the missing potential outcomes (model-based or matching-based)
  - ▶ Weighting: weight (often function of the propensity scores) the observed data to represent a target population
- ▶ Three modes of inference
  - ▶ Frequentist: imputation, weighting, motivated by consistency, asymptotic normality, (semiparametric) efficiency, etc.
  - ▶ Bayesian: modeling and imputing missing potential outcomes based on their posterior distributions
  - ▶ Fisherian randomization: combine randomization tests with Bayesian methods, unique to randomized experiments

# Causal Inference vs. Missing Data

- ▶ Under PO framework, causal inference is a missing data problem
- ▶ A broad parallel between the classification of assignment mechanisms in causal inference and the classification of missing data mechanisms (Ding and Li, 2108, Stat Sci)
  - ▶ completely randomized experiments  $\iff$  missing completely at random (MCAR)
  - ▶ observational studies with unmeasured confounding  $\iff$  missing not at random (MNAR)
  - ▶ ignorable assignment mechanisms MAR  $\iff$  missing at random (MAR)
- ▶ But ignorable assignment mechanism can be generalized to a richer class of assignment mechanisms (sequential ignorable, locally ignorable, etc)

## A Few Comments

- ▶ Causal effects, in their definition, do not relate to probability distribution for subjects “who got different treatments”, or to coefficients of models
- ▶ We do not assume necessarily that the P.O.s are fixed. In fact we could consider an inherent status as random, but that is not important for the scale we discuss here
- ▶ We allow that the potential outcomes depend on other factors that relate to that person at the time and place when the treatment is given, and through the follow-up time where the P.O.s would be observed
- ▶ Not a before-after comparison

# On “effects of cause” not “cause of effects”

A Conversation with DB Rubin (p 446), Stat. Sci. 2014

**FL:** *In the RCM, cause/intervention should always be defined before you start the analysis. In other words, the RCM is a framework to investigate the “effects of a cause,” but not the “causes of an effect.” Some criticize this as a major limitation. Do you regard this as a limitation? Do you think it is ever possible to draw inference on the causes of effects from data, or is it, per se, an interesting question worth further investigation?*

**Rubin:** *I regard “the cause” of an event topic as more of a cocktail conversation topic than a scientific inquiry, because it leads to an essentially infinite regress. Someone says, “He died of lung cancer because he smoked three packs a day”; then someone else counters, “Oh no, he died of lung cancer because both of his parents smoked three packs a day and, therefore, there was no hope of his doing anything other than smoking three packs a day”; then another one says, “No, no, his parents smoked because his grandparents smoked - they lived in North Carolina where, back then, everyone smoked three packs a day, so the cause is where the grandparents lived,” and so on. How far back should you go? You can’t talk sensibly about the cause of an event; you can talk about “but for that cause (and there can be many ‘but for’s), what would have happened?” All these questions can be addressed hypothetically. But the cause? The notion is meaningless to me.*

# References

- Bickel, P. J., Hammel, E. A., O'Connell, J. W. (1975). Sex bias in graduate admissions: Data from Berkeley. *Science*, 187(4175), 398-404.
- Cole, S. R., Frangakis, C. E. (2009). The consistency statement in causal inference: a definition or an assumption?. *Epidemiology*, 20(1), 3-5.
- Cornfield, J., Haenszel, W., et al. (1959). Smoking and lung cancer: recent evidence and a discussion of some questions. *Journal of the National Cancer institute*, 22(1), 173-203.
- Ding, P, Li, F.(2018). Causal inference: a missing data perspective. *Statistical Science*. 33(2), 214-237.
- Doll, R., Hill, A. B. (1954). The mortality of doctors in relation to their smoking habits. *British medical journal*, 1(4877), 1451.
- Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Edinburgh, UK: Oliver and Boyd.
- Fisher, R. A. (1935). *The Design of Experiments* (1st ed.), Edinburgh, UK: Oliver and Boyd.
- Fisher, R. A. (1957). Dangers of cigarette-smoking. *British Medical Journal*, 2(5039), 297.
- Hernan MA, Robins JM. *Causal Inference*. (2020). Boca Raton: Chapman & Hall/CRC.
- Holland, P. (1986). Statistics and causal inference (with discussion). *J. Amer. Statist. Assoc.* 81, 945-970
- Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge Univ. Press, New York.

# References

Li, F, Mealli, F. (2014). A conversation with Donald B. Rubin. *Statistical Science*. 29(3), 439-457.

Neyman J. (1990). On the application of probability theory to agricultural experiments: Essay on Principles, Section 9. Masters Thesis. Portions translated into English by D. Dabrowska and T. Speed (1990). *Statistical Science*. 1993(5): 465-472

Pearl, J. (2009). *Causality*. Cambridge University Press.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* 66 688-701.

Rubin, D. B. (1976). Inference and missing data. *Biometrika* 63 581-592.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics*. 6 34-58.

Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. *Journal of the American Statistical Association*, 75(371), 591-593.

Simpson, Edward H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society, Series B*. 13: 238-241.