Bayesian causal inference: an introduction

Peng Ding

UC Berkeley, Statistics

April 25, 2024

More advanced topics are in:

Li, Ding, and Mealli (2023) Bayesian causal inference: a critical review.

Philosophical Transactions of the Royal Society, Series A

What is causal inference and why causal inference?

- Causal inference: evaluate effect of treatment on outcome
  - economic return of college education: labor economics
  - effect of drug: drug approval of Food and Drug Administration (FDA)
  - gene and environment effect on health
- Causal inference and decision
  - what if we change something?
  - causal inference is about hypothetical intervention
- David Hume: causal inference is impossible based on experience
  - we can only learn "constant conjunctions" from experience
  - "constant conjunctions" are not causal
  - statistics offers a constructive perspective for causal inference

Causal inference using potential outcomes

- Units $i = 1, \ldots, n$
- Treatment or intervention: Z_i
  - $Z_i = 1$ for treatment; $Z_i = 0$ for control
  - binary Z is interesting enough; can extend to general Z
- Outcome of interest Y_i
- Potential outcomes (Neyman–Rubin model, Rubin Causal Model):

$$Y_i(1), Y_i(0)$$

- Causal effects = comparisons of potential outcomes

Hidden assumptions in the notation of potential outcomes

- Consistency: there is only one version of the treatment
  - Z college or not; college does not matter, major does not matter
  - Z covid vaccine or not; Moderna or Pfizer does not matter
  - strong assumption; can be relaxed with more detailed data
- No interference: $Y_i(z)$ do not depend on other units' treatments
  - violation in infectious diseases
  - violation in network experiments: experiments at Meta or LinkedIn
  - can be relaxed by isolating units
  - studying interference can be of scientific interest
- Consistency + No interference = SUTVA
  - Stable Unit Treatment Value Assumption (SUTVA)
  - $Y_i(1), Y_i(0)$'s are well defined

"Science Table"—Rubin's terminology

unit	under treatment	under control
$i = 1$	$Y_1(1)$	$Y_1(0)$
$i = 2$	$Y_2(1)$	$Y_2(0)$
$\vdots$		
$i = n$	$Y_n(1)$	$Y_n(0)$

- If we knew all potential outcomes, causal inference problem is solved
  - average treatment effect $\tau = n^{-1} \sum_{i=1}^{n} \{ Y_i(1) - Y_i(0) \}$
  - median treatment effect $\text{median}\{ Y_i(1) \}_{i=1}^{n} - \text{median}\{ Y_i(0) \}_{i=1}^{n}$

- Fundamental problem of causal inference: for each unit, only one potential outcome is observed, with the other missing

Causal inference can be viewed as a missing data problem

unit	treatment status	under treatment	under control
$i = 1$	$Z_1 = 1$	$Y_1(1)$	?
$i = 2$	$Z_2 = 0$	?	$Y_2(0)$
$\vdots$			
$i = n$	$Z_n = 1$	$Y_n(1)$	?

- Half of the potential outcomes are missing
- Key for causal inference: infer or impute missing potential outcomes

The central role of the treatment assignment mechanism

- The mechanism for ? matters a lot for causal inference

- Examples from Ding (2024 book), with true average effect $= -0.5$

```
> n   = 500
> Y0 = rnorm(n)
> tau = - 0.5 + Y0
> Y1 = Y0 + tau
```

A perfect doctor assigns the treatment to the patient if s/he knows that the individual causal effect is non-negative. This results in a positive difference in means of the observed outcomes:

```
> Z = (tau >= 0)
> Y = Z*Y1 + (1 - Z)*Y0
> mean(Y[Z==1]) - mean(Y[Z==0])
[1] 2.166509
```

A clueless doctor does not know any information about the individual causal effects and assigns the treatment to patients by flipping a fair coin. This results in a difference in means of the observed outcomes close to the true average causal effect:

```
> Z = rbinom(n, 1, 0.5)
> Y = Z*Y1 + (1 - Z)*Y0
> mean(Y[Z==1]) - mean(Y[Z==0])
[1] -0.552064
```

Treatment assignment mechanisms

- The examples on the last page are two extreme cases
  - treatment assignment depends on missing potential outcomes
  - treatment assignment does not depend on potential outcomes
- Intermediate case 1: treatment assignment depends on observed covariates but not potential outcomes – unconfounded studies
- Intermediate case 2: treatment assignment does not depend on potential outcomes but the final treatment received depends on unmeasured covariates – noncompliance and instrumental variable

Completely randomized experiment

- $Z_1, \ldots, Z_n$ IID Bernoulli
  - Treated group: $Z_i = 1$, observe $Y_i(1)$, missing $Y_i(0)$
  - Control group: $Z_i = 0$, missing $Y_i(1)$, observe $Y_i(0)$

- Infer the population/sample average treatment effects
  - PATE:

$$\tau^{\mathsf{P}} = \mu_1 - \mu_0,$$

    where μ_1 and μ_0 are means of $Y_i(1)$ and $Y_i(0)$
  - SATE:

$$\tau^{\mathsf{S}} = n^{-1} \sum_{i=1}^{n} \{Y_i(1) - Y_i(0)\}$$

Completely randomized experiment: PATE

- Infer μ_1: standard Bayesian problem under $Y_i \mid Z_i = 1 \sim N(\mu_1, \sigma_1^2)$
  - Normal model: obtain the posterior of μ_1
- Infer μ_0: standard Bayesian problem under $Y_i \mid Z_i = 0 \sim N(\mu_0, \sigma_0^2)$
  - Normal model: obtain the posterior of μ_0
- Obtain the posterior of

$$\tau^{\mathsf{P}} = \mu_1 - \mu_0$$

Completely randomized experiment: SATE

- High-level idea: impute all missing potential outcomes
  - based on their posterior distribution given the data
- Subtle issue: impute $Y_i(1)$ given $Y_i(0)$; impute $Y_i(0)$ given $Y_i(1)$
  - it involves the joint distribution of $\{Y_i(1), Y_i(0)\}$
- A bivariate Normal potential outcome model:

$$\begin{pmatrix} Y_i(1) \\ Y_i(0) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_0 \\ \rho\sigma_1\sigma_0 & \sigma_0^2 \end{pmatrix} \right)$$

  - marginal distributions $Y_i(1) \sim N(\mu_1, \sigma_1^2)$ and $Y_i(0) \sim N(\mu_0, \sigma_0^2)$

Completely randomized experiment: SATE, more details

- How do we obtain the posterior distributions of parameters?
  - μ_1, σ_1^2 and μ_0, σ_0^2: identical to classic Bayesian Normal model
  - ρ: no information from the data, not identifiable
- We encounter a fundamental problem about ρ
  - if we impose a prior on ρ, the posterior is identical
  - we can vary ρ within a region as sensitivity analysis
- How do we impute the missing potential outcomes?
  - for treated unit, we sample the control potential outcome:
    $Y_i(0) \mid Y_i(1) = Y_i \sim N(\mu_0 + \rho\sigma_0/\sigma_1 \cdot (Y_i - \mu_1), \sigma_0^2(1-\rho^2))$
  - for control unit, we sample the treated potential outcome:
    $Y_i(1) \mid Y_i(0) = Y_i \sim N(\mu_1 + \rho\sigma_1/\sigma_0 \cdot (Y_i - \mu_0), \sigma_1^2(1-\rho^2))$

Causal inference with observational studies

- Treatment Z_i not randomized
- Observed pretreatment covariates X_i
- Factorize the joint density

$$\begin{aligned} & \Pr\{Y_i(0), Y_i(1), Z_i, X_i\} \\ = \; & \Pr\{Z_i \mid Y_i(0), Y_i(1), X_i\} \cdot \Pr\{Y_i(0), Y_i(1) \mid X_i\} \cdot \Pr(X_i) \end{aligned}$$

- Key assumption of ignorability $\Pr\{Z_i \mid Y_i(0), Y_i(1), X_i\} = \Pr\{Z_i \mid X_i\}$
  - why "ignorability"? – this assumption allows us to ignore modeling Z
  - Bayesian causal inference reduces to outcome modeling and prediction

Causal inference with observational studies: an example

- Bivariate Normal linear model:

$$\begin{pmatrix} Y_i(1) \\ Y_i(0) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \beta_1' X_i \\ \beta_0' X_i \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_0 \\ \rho\sigma_1\sigma_0 & \sigma_0^2 \end{pmatrix} \right)$$

  - marginally, $Y_i(z) \mid X_i, \beta_z, \sigma_z^2 \sim \mathcal{N}(\beta_z' X_i, \sigma_z^2)$ for $z = 0, 1$.

- Focus on SATE

- Key: impute missing potential outcomes

Causal inference with observational studies: an example

- $\beta_1, \beta_0, \sigma_1^2, \sigma_0^2$: standard Bayesian posterior
- ρ: no information in the data—sensitivity analysis
- Impute missing potential outcomes
  - for treated unit, we sample the control potential outcome:
    $Y_i(0) \mid Y_i(1) = Y_i \sim N(\beta_0' X_i + \rho\sigma_0/\sigma_1 \cdot (Y_i - \beta_1' X_i), \sigma_0^2(1-\rho^2))$
  - for control unit, we sample the treated potential outcome:
    $Y_i(1) \mid Y_i(0) = Y_i \sim N(\beta_1' X_i + \rho\sigma_1/\sigma_0 \cdot (Y_i - \beta_0' X_i), \sigma_1^2(1-\rho^2))$
- Side note: the method can also be used in randomized experiment with covariates, which is the Bayesian version of analysis of covariance

Problem of outcome modeling and role of propensity score

- What is the outcome model is wrong?
- Are the results robust to model misspecification?
- Ignoring the treatment mechanism may not be a good idea
  - strange to analyze experiments and observational studies in same way
- Rosenbaum and Rubin (1983) propose the propensity score:

$$e(X_i) = \Pr(Z_i = 1 \mid X_i)$$

- Key property of the propensity score:

$$Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0)\} \mid X_i \Longrightarrow Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0)\} \mid e(X_i)$$

Use propensity score in Bayesian causal inference

- Key property of the propensity score: given the propensity score, the treatment is essentially randomized
- We first fit a model for $e(X_i)$: e.g. logistic regression $Z_i \mid X_i$
- We discretize the estimated propensity score
  - e.g. 5 strata
  - make sure covariates are balanced across treatment and control groups
- Analyze the data stratified on the discretized propensity score
  - can further include X_i in outcome modeling
  - average over all strata to obtain SATE

Related papers

- Ding, P. (2024) A first course in causal inference. Chapman & Hall.
- Li, F., Ding, P. and Mealli, F. (2023) Bayesian causal inference: a critical review. *Philosophical Transactions of the Royal Society A*
- Ding, P. and Guo, T. (2023) Posterior Predictive Propensity Scores and p-Values. *Observational Studies*
- Ding, P. and Li, F. (2018) Causal inference: a missing data perspective. *Statistical Science*