Causal inference: for statistics, social, and biomedical sciences

Chapter 25: Model-based analysis in instrumental variable settings: randomized experiments with two-sided noncompliance

Kunwoong Kim

2022.8.30.

Contents

25.1 Introduction

- Handle with average treatment effects on randomized experiments with two-sided noncompliance.
- Develop a model-based alternative to the moment-based analyses (Chapter 23 and 24).

25.2 The McDonald-Hiu-Tierney influenza vaccination data

For the i^{th} patient, we have the following notations.

- Z_i : indicator of whether the physician of i^{th} patient received a letter encouraging vaccination **(random)**
- W_i : indicator of receiving a flu shot
- Y_i : binary response indicating the hospitalization for flu-related illnesses
- X_i : a set of pre-treatment variables
- G_i : indicator of four compliance groups **(bijective to (W_i, Z_i))**

$$G_i = \begin{cases} \text{nt} & \text{if } W_i(0) = 0,\ W_i(1) = 0, \\ \text{co} & \text{if } W_i(0) = 0,\ W_i(1) = 1, \\ \text{df} & \text{if } W_i(0) = 1,\ W_i(1) = 0, \\ \text{at} & \text{if } W_i(0) = 1,\ W_i(1) = 1. \end{cases}$$

25.2 The McDonald-Hiu-Tierney influenza vaccination data

- 1,931 female patients.
- Table 25.1: averages by treatment and assignment group for outcomes and covariates.
- Table 25.2: the number of individuals in each of the eight subsamples defined by Z_i, W_i, Y_i $(2^3 = 8)$ with means of three X_is: `age`, `copd`, `heart`.
- Note: the design of this experiment involved randomization physicians rather than patients; for physicians with multiple patients, outcomes of those patients would be correlated.

25.2 The McDonald-Hiu-Tierney influenza vaccination data

Table 25.1. ***Summary Statistics for Women by Assigned Treatment, Received Treatment: Covariates and Outcome for Influenza Vaccination Data***

			Means			Means		
			No Letter	Letter	t-Stat	No Flu Shot	Flu Shot	t-Stat
	Mean	STD	$Z_i = 0$	$Z_i^{obs} = 1$	dif	$W_i^{obs} = 0$	$W_i^{obs} = 1$	dif
letter (Z_i)	0.53	(0.50)	0	1	–	0.49	0.63	[7.8]
flu shot (W_i^{obs})	0.24	(0.43)	0.18	0.29	[7.7]	0	1	–
hosp (Y_i^{obs})	0.08	(0.27)	0.09	0.06	[−3.2]	0.08	0.07	[−0.4]
age	65.4	(12.8)	65.2	65.6	[1.1]	64.9	67.1	[4.9]
copd	0.20	(0.40)	0.21	0.20	[−1.3]	0.20	0.23	[2.4]
heart	0.56	(0.50)	0.56	0.57	[0.6]	0.55	0.60	[2.4]

Table 25.2. ***Summary Statistics for Women by Assigned Treatment, Received Treatment and Outcome, and Possible Latent Compliance Status for Influenza Vaccination Data***

Type under Monotonicity and Exclusion Restr.	Assign. (Letter) Z_i	Receipt of Flu Shot W_i^{obs}	Hosp. Y_i^{obs}	# of Units 1,931	Means age	copd	heart
Complier or nevertaker	0	0	0	685	64.7	0.18	0.524
Complier or nevertaker	0	0	1	64	62.9	0.33	0.77
Alwaystaker	0	1	0	148	67.8	0.28	0.60
Alwaystaker	0	1	1	20	68.9	0.30	0.70
Nevertaker	1	0	0	672	65.4	0.19	0.55
Nevertaker	1	0	1	51	62.0	0.29	0.69
Complier or alwaystaker	1	1	0	277	66.6	0.20	0.57
Complier or alwaystaker	1	1	1	14	67.3	0.21	0.79

25.3 Covariates

Assumptions

1. $Z_i \perp (W_i(0), W_i(1), Y_i(0,0), Y_i(0,1), Y_i(1,0), Y_i(1,1)) | X_i$
2. $Y_i(0, W_i(0)) = Y_i(1, W_i(1))$ for all nevertakers and alwaystakers
3. $Z_i \perp Y_i(Z_i, W_i(Z_i)) | X_i, G_i = nt$
4. $Z_i \perp Y_i(Z_i, W_i(Z_i)) | X_i, G_i = at$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

Notations

- Let $\mathbf{W}(0), \mathbf{W}(1)$ the N-vectors of secondary potential outcomes with i^{th} element equal to $W_i(0), W_i(1)$, indicating the primary treatment received under assignment to $Z_i = 0$ and $Z_i = 1$ respectively, and let $\mathbf{W} = (\mathbf{W}(0), \mathbf{W}(1))$.
- We are interested in the **local average treatment effect** for compliers,

$$\tau_{late} = \frac{1}{N_c} \sum_{i:G_i=co} (Y_i(1) - Y_i(0)) \quad (1)$$

- Let missing and observed values for the treatment received in similar fashion:

$$W_i^{mis} = W_i(1 - Z_i), W_i^{obs} = W_i(Z_i).$$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

Check the textbook for details and chapter 8 for model-based approach without compliance.

- We cannot directly specify the posterior predictive distribution of the missing data

$$f(\mathbf{Y}^{mis}, \mathbf{W}^{mis} | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})$$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

- To derive the posterior of missing outcomes, we assume probabilistic models on variables.
- Let θ roughly denotes all parameters used in the models that are used in this model-based approach.
- We begin with setting two probabilistic models:
  - $f(\mathbf{Y}|\mathbf{G}, \mathbf{X}; \theta) = \prod_{i=1}^{N} f(Y_i(0), Y_i(1)|G_i, X_i; \theta)$, and
  - $f(\mathbf{G}|\mathbf{X}; \theta) = \prod_{i=1}^{N} f(G_i|X_i, \theta)$.

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

$$f(\mathbf{Y}|\mathbf{G},\mathbf{X};\theta) = \prod_{i=1}^{N} f(Y_i(0), Y_i(1)|G_i, X_i;\theta)$$

- Compliers

$$Y_i(0)|G_i = co, X_i;\theta \sim \mathcal{N}(X_i\beta_{co,c}, \sigma^2_{co,c})$$

$$Y_i(1)|G_i = co, X_i;\theta \sim \mathcal{N}(X_i\beta_{co,t}, \sigma^2_{co,t})$$

- Nevertakers, Alwaytakers

$$Y_i(0)|G_i = nt, X_i;\theta \sim \mathcal{N}(X_i\beta_{nt}, \sigma^2_{nt})$$

$$Y_i(1)|G_i = at, X_i;\theta \sim \mathcal{N}(X_i\beta_{at}, \sigma^2_{at})$$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

$$f(\mathbf{G}|\mathbf{X};\theta) = \prod_{i=1}^{N} f(G_i|X_i,\theta)$$

- Multinomial logit model

$$\mathbb{P}(G_i = co|X_i,\theta) = \frac{1}{1+\exp(X_i\gamma_{nt})+\exp(X_i\gamma_{at})}$$

$$\mathbb{P}(G_i = nt|X_i,\theta) = \frac{\exp(X_i\gamma_{nt})}{1+\exp(X_i\gamma_{nt})+\exp(X_i\gamma_{at})}$$

$$\mathbb{P}(G_i = at|X_i,\theta) = \frac{\exp(X_i\gamma_{at})}{1+\exp(X_i\gamma_{nt})+\exp(X_i\gamma_{at})}$$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

Derivation of $f(\mathbf{Y}^{mis}, \mathbf{W}^{mis} | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z}, \theta)$

- We can now compute

$$f(\mathbf{Y}(0), \mathbf{Y}(1), \mathbf{W}(0), \mathbf{W}(1) | \mathbf{X}, \theta)$$

1. $f(\mathbf{Y}, \mathbf{G} | \mathbf{X}, \theta) = f(\mathbf{Y} | \mathbf{G}, \mathbf{X}; \theta) f(\mathbf{G} | \mathbf{X}; \theta)$ (remember $\mathbf{G}$ is one-to-one function of $\mathbf{W}(0), \mathbf{W}(1)$)
2. $(\mathbf{Y}(0), \mathbf{Y}(1), \mathbf{W}(0), \mathbf{W}(1))$ is one-to-one of $(\mathbf{Y}^{mis}, \mathbf{W}^{mis}, \mathbf{Y}^{obs}, \mathbf{W}^{obs})$
3. $f(\mathbf{Y}(0), \mathbf{Y}(1), \mathbf{W}(0), \mathbf{W}(1) | \mathbf{X}, \theta) = f(\mathbf{Y}(0), \mathbf{Y}(1), \mathbf{W}(0), \mathbf{W}(1) | \mathbf{X}, \mathbf{Z}, \theta)$ due to the unconfoundness assumption.

- Thus we can derive $f(\mathbf{Y}^{mis}, \mathbf{W}^{mis}, \mathbf{Y}^{obs}, \mathbf{W}^{obs} | \mathbf{X}, \mathbf{Z}, \theta)$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

Then we infer the conditional distribution as

$$f(\mathbf{Y}^{mis}, \mathbf{W}^{mis} | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z}, \theta)$$
$$= \frac{f(\mathbf{Y}^{mis}, \mathbf{Y}^{obs}, \mathbf{W}^{mis}, \mathbf{W}^{obs} | \mathbf{X}, \mathbf{Z}, \theta)}{\int \int f(\mathbf{Y}^{mis}, \mathbf{Y}^{obs}, \mathbf{W}^{mis}, \mathbf{W}^{obs} | \mathbf{X}, \mathbf{Z}, \theta) \, d\mathbf{Y}^{mis} \, d\mathbf{W}^{mis}}.$$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

- Derivation of the posterior distribution $p(\theta|\mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})$

$$\mathcal{L}(\theta|\mathbf{Y}^{pbs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z}) = f(\mathbf{Y}^{obs}, \mathbf{W}^{obs}|\mathbf{X}, \mathbf{Z}, \theta)$$
$$= \int\int f(\mathbf{Y}^{mis}, \mathbf{Y}^{obs}, \mathbf{W}^{mis}, \mathbf{W}^{obs})\, d\mathbf{Y}^{mis}\, d\mathbf{W}^{mis}.$$

- We multiply this likelihood function of θ by the prior distribution for θ, $p(\theta)$ as

$$p(\theta|\mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z}) = \frac{p(\theta)\cdot f(\mathbf{Y}^{obs}, \mathbf{W}^{obs}|\mathbf{X}, \mathbf{Z}, \theta)}{f(\mathbf{Y}^{obs}, \mathbf{W}^{obs}|\mathbf{X}, \mathbf{Z})}$$
$$= \frac{p(\theta)\cdot f(\mathbf{Y}^{obs}, \mathbf{W}^{obs}|\mathbf{X}, \mathbf{Z}, \theta)}{\int p(\theta)\cdot f(\mathbf{Y}^{obs}, \mathbf{W}^{obs}|\mathbf{X}, \mathbf{Z}, \theta)\, d\theta}. \tag{2}$$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

Derivation of the posterior distribution of missing potential outcomes $f(\mathbf{Y}^{mis}, \mathbf{W}^{mis} | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})$

- We combine
  - conditional posterior distribution of the missing potential outcomes given the parameter θ : $f(\mathbf{Y}^{mis}, \mathbf{W}^{mis} | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z}, \theta)$
  - posterior distribution of θ : $p(\theta | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})$
- so that we obtain

$$f(\mathbf{Y}^{mis}, \mathbf{W}^{mis} | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})$$
$$\int_\theta f(\mathbf{Y}^{mis}, \mathbf{W}^{mis} | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z}, \theta) \cdot p(\theta | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})\, d\theta. \tag{3}$$

25.4 Analyses for Randomized Experiments with Two-Sided Noncompliance

Derivation of the posterior distribution of estimands

- Infer the posterior distribution of τ given the observed data $(\mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})$ using
  - $f(\mathbf{Y}^{mis}, \mathbf{W}^{mis} | \mathbf{Y}^{obs}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})$ and
  - the fact that any estimand is a function of $(\mathbf{Y}(0), \mathbf{Y}(1), \mathbf{W}(0), \mathbf{W}(1), \mathbf{X}, \mathbf{Z})$ can be rewritten as a function of $(\mathbf{Y}^{mis}, \mathbf{Y}^{obs}, \mathbf{W}^{mis}, \mathbf{W}^{obs}, \mathbf{X}, \mathbf{Z})$.

25.5 Simulation methods for obtaining draws from the posterior distribution of the estimand given the data

- (1) The conditional joint distribution of $\mathbf{Y}$ is

$$f(\mathbf{Y}|\mathbf{G},\mathbf{X};\theta)=\prod_{i=1}^{N} f(Y_i(0)|G_i,X_i,\theta)\cdot f(Y_i(1)|G_i,X_i,\theta). \tag{4}$$

- (2) The compliance type probability is

$$f(\mathbf{G}|\mathbf{X};\gamma)=\prod_{i=1}^{N} p(G_i|X_i,\gamma) \tag{5}$$

where γ is a subvector of θ.

25.5 Simulation methods for obtaining draws from the posterior distribution of the estimand given the data

(3) Likelihood of θ

- There are four possible patterns of missing and observed data: $(Z_i, W_i^{obs}) = (0,0), (0,1), (1,0), (1,1)$.
- Then, we have four different likelihood functions ($\mathcal{L}_{(z,w),i}$ is the likelihood of i^{th} individual of $Z_i = z$ and $W_i = w$):

$$\begin{aligned}
\mathcal{L}_{(0,1),i} &= P(G_i = at | X_i, Z_i, \gamma) \cdot f(Y_i(1) | G_i = at, X_i, Z_i, \beta_{at}) \\
\mathcal{L}_{(1,0),i} &= P(G_i = nt | X_i, Z_i, \gamma) \cdot f(Y_i(1) | G_i = nt, X_i, Z_i, \beta_{nt}) \\
\mathcal{L}_{(0,0),i} &= P(G_i = nt | X_i, Z_i, \gamma) \cdot f(Y_i(1) | G_i = nt, X_i, Z_i, \beta_{nt}) \\
&+ P(G_i = co | X_i, Z_i, \gamma) \cdot f(Y_i(0) | G_i = co, X_i, Z_i, \beta_{co,c}) \\
\mathcal{L}_{(1,1),i} &= P(G_i = at | X_i, Z_i, \gamma) \cdot f(Y_i(1) | G_i = at, X_i, Z_i, \beta_{at}) \\
&+ P(G_i = co | X_i, Z_i, \gamma) \cdot f(Y_i(1) | G_i = co, X_i, Z_i, \beta_{co,t})
\end{aligned} \tag{6}$$

25.5 Simulation methods for obtaining draws from the posterior distribution of the estimand given the data

- Summing up, we have the overall likelihood function as

$$\mathcal{L}_{obs}(\theta|\mathbf{Z}^{obs}, \mathbf{W}^{obs}, \mathbf{Y}^{obs}, \mathbf{X}^{obs}) = \prod_{i \in \mathcal{S}(0,1)} \mathcal{L}_{(0,1),i} \prod_{i \in \mathcal{S}(1,0)} \mathcal{L}_{(1,0),i} \prod_{i \in \mathcal{S}(0,0)} \mathcal{L}_{(0,0),i} \prod_{i \in \mathcal{S}(1,1)} \mathcal{L}_{(1,1),i} \tag{7}$$

where $\mathcal{S}_{(z,w)}$ is the subset of indices of observed individuals having $Z = z$ and $W = w$.

25.5 Simulation methods for obtaining draws from the posterior distribution of the estimand given the data

Before obtaining completed likelihood, we compute the probability of compliance as

$$\Pr(G_i = \text{co}|Y_i^{\text{obs}} = y, W_i^{\text{obs}} = 0, Z_i = 0, X_i = x, \theta) \tag{25.8}$$

$$= \frac{p(\text{co}|x,\beta) \cdot f(y|x,\beta_{\text{co,c}})}{p(\text{co}|x,\beta) \cdot f(y|x,\beta_{\text{co,c}}) + p(\text{nt}|x,\beta) \cdot f(y|x,\beta_{\text{nt}})}.$$

Similarly,

$$\Pr(G_i = \text{co}|Y_i^{\text{obs}} = y, W_i^{\text{obs}} = 1, Z_i = 1, X_i = x, \theta) \tag{25.9}$$

$$= \frac{p(\text{co}|x,\beta) \cdot f(y|x,\beta_{\text{co,t}})}{p(\text{co}|x,\beta) \cdot f(y|x,\beta_{\text{co,t}}) + p(\text{at}|x,\beta) \cdot f(y|x,\beta_{\text{at}})}.$$

25.5 Simulation methods for obtaining draws from the posterior distribution of the estimand given the data

- We have the complete-data likelihood function as

$$
\begin{aligned}
&\mathcal{L}_{\text{comp}}(\theta|\mathbf{G},\mathbf{Z},\mathbf{W}^{\text{obs}},\mathbf{Y}^{\text{obs}},\mathbf{X})\\
&= \prod_{i:G_i=\text{nt}} f(Y_i(0)|G_i=\text{nt},X_i,Z_i,\beta_{\text{nt}})\\
&\quad\times \prod_{i:G_i=\text{at}} f(Y_i(1)|G_i=\text{at},X_i,Z_i,\beta_{\text{at}}) \prod_{i:G_i=\text{co},Z_i=0} f(Y_i(0)|G_i=\text{co},X_i,Z_i,\beta_{\text{co,c}})\\
&\quad\times \prod_{i:G_i=\text{co},Z_i=1} f(Y_i(1)|G_i=\text{at},X_i,Z_i,\beta_{\text{co,t}}) \prod_{i:G_i=\text{co}} p(G_i=\text{co}|X_i,Z_i,\beta)\\
&\quad\times \prod_{i:G_i=\text{at}} p(G_i=\text{at}|X_i,Z_i,\beta) \prod_{i:G_i=\text{nt}} p(G_i=\text{nt}|X_i,Z_i,\beta).
\end{aligned}
$$

25.5 Simulation methods for obtaining draws from the posterior distribution of the estimand given the data

- Using these distributions, we can now obtain the posterior distribution of outcomes.
- How to implement?
  1. Initialize θ, simulate the compliance type with (25.8) and (25.9).
  2. Given compliance type $\mathbf{G}$ and θ, draw from the posterior of the missing potential outcomes for compliers, i.e., impute $Y_i(1)$ if $Z_i = 0$ and $Y_i(0)$ if $Z_i = 1$.
  3. Compute $\tau_{late} = \frac{1}{N_{co}} \sum_{i:G_i=co}(Y_i(1) - Y_i(0))$ where $N_{co} = \sum_{i=1}^{N} I_{G_i=co}$ is the number of compliers.
  4. Update θ given $\mathbf{Y}^{obs}, \mathbf{G}, \mathbf{Z}, \mathbf{X}$ (maximize the complete likelihood).
  5. Return to Step 2 until break.

25.6 Models for the Influenza vaccination data

(1) Model for $\mathbf{Y}|\mathbf{G}, \mathbf{X}; \theta$

$$\Pr(Y_i(0) = y|X_i = x, G_i = \text{co}, \theta) = \frac{\exp(y \cdot x\beta_{\text{co,c}})}{1 + \exp(x\beta_{\text{co,c}})},$$

$$\Pr(Y_i(1) = y|X_i = x, G_i = \text{co}, \theta) = \frac{\exp(y \cdot x\beta_{\text{co,t}})}{1 + \exp(x\beta_{\text{co,t}})},$$

for nevertakers,

$$\Pr(Y_i(0) = y|X_i = x, G_i = \text{nt}, \theta) = \frac{\exp(y \cdot x\beta_{\text{nt}})}{1 + \exp(x\beta_{\text{nt}})},$$

and, finally, for alwaystakers,

$$\Pr(Y_i(1) = y|X_i = x, G_i = \text{at}, \theta) = \frac{\exp(y \cdot x\beta_{\text{at}})}{1 + \exp(x\beta_{\text{at}})}.$$

25.6 Models for the Influenza vaccination data

(2) Model for $\mathbf{G}|\mathbf{X};\gamma$

$$\Pr(G_i = g|X_i = x) = \begin{cases} \dfrac{1}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})}, & \text{if } g = \text{co} \\ \dfrac{\exp(x\gamma_{\text{nt}})}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})}, & \text{if } g = \text{nt} \\ \dfrac{\exp(x\gamma_{\text{at}})}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})}, & \text{if } g = \text{at}, \end{cases}$$

25.6 Models for the Influenza vaccination data

(3) Likelihood of θ

$$
\begin{aligned}
&\mathcal{L}_{\text{obs}}(\theta|\mathbf{Z}^{\text{obs}},\mathbf{W}^{\text{obs}},\mathbf{Y}^{\text{obs}},\mathbf{X}^{\text{obs}}) \\
&= \prod_{i\in\mathcal{S}(0,0)} \left[\frac{\exp(X_i\gamma_{\text{nt}})}{1+\exp(X_i\gamma_{\text{at}})+\exp(X_i\gamma_{\text{nt}})} \cdot \frac{\exp(Y_i^{\text{obs}}\cdot X_i\beta_{\text{nt}})}{1+\exp(X_i\beta_{\text{nt}})} \right. \\
&\quad \left. + \frac{1}{1+\exp(X_i\gamma_{\text{at}})+\exp(X_i\gamma_{\text{nt}})} \cdot \frac{\exp(Y_i^{\text{obs}}\cdot X_i\beta_{\text{co0}})}{1+\exp(X_i\beta_{\text{co},0})} \right] \\
&\quad \times \prod_{i\in\mathcal{S}(0,1)} \frac{\exp(X_i\gamma_{\text{at}})}{1+\exp(X_i\gamma_{\text{at}})+\exp(X_i\gamma_{n})} \cdot \frac{\exp(Y_i^{\text{obs}}\cdot X_i\beta_{\text{at}})}{1+\exp(X_i\beta_{\text{at}})} \\
&\quad \times \prod_{i\in\mathcal{S}(1,0)} \frac{\exp(X_i\gamma_{\text{nt}})}{1+\exp(X_i\gamma_{\text{at}})+\exp(X_i\gamma_{\text{nt}})} \cdot \frac{\exp(Y_i^{\text{obs}}\cdot X_i\beta_{\text{nt}})}{1+\exp(X_i\beta_{\text{nt}})} \\
&\quad \times \prod_{i\in\mathcal{S}(1,1)} \left[\frac{\exp(X_i\gamma_{\text{at}})}{1+\exp(X_i\gamma_{\text{at}})+\exp(X_i\gamma_{\text{nt}})} \cdot \frac{\exp(Y_i^{\text{obs}}\cdot X_i\beta_{\text{at}})}{1+\exp(X_i\beta_{\text{at}})} \right. \\
&\quad \left. + \frac{1}{1+\exp(X_i\gamma_{\text{at}})+\exp(X_i\gamma_{\text{nt}})} \cdot \frac{\exp(Y_i^{\text{obs}}\cdot X_i\beta_{\text{co},1})}{1+\exp(X_i\beta_{\text{co1}})} \right].
\end{aligned}
$$

25.6 Models for the Influenza vaccination data

(4) Compliance probability

$$\Pr(G_i = \text{co} | Y_i^{\text{obs}} = y, W_i^{\text{obs}} = 0, Z_i = 0, X_i = x, \theta) \tag{25.10}$$

$$= \frac{p(\text{co}|x,\beta) \cdot f(y|x,\beta_{\text{co,c}})}{p(\text{co}|x,\beta) \cdot f(y|x,\beta_{\text{co,c}}) + p(\text{nt}|x,\beta) \cdot f(y|x,\beta_{\text{nt}})}$$

$$= \left(\frac{\frac{1}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})} \cdot \frac{\exp(x\beta_{\text{co,c}})}{1+\exp(x\beta_{\text{co,c}})}}{\frac{1}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})} \cdot \frac{\exp(x\beta_{\text{co,c}})}{1+\exp(x\beta_{\text{co,c}})} + \frac{\exp(\gamma_{\text{nt}})}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})} \cdot \frac{\exp(x\beta_{\text{nt}})}{1+\exp(x\beta_{\text{nt}})}} \right)^{y}$$

$$\times \left(\frac{\frac{1}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})} \cdot \frac{1}{1+\exp(x\beta_{\text{co,c}})}}{\frac{1}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})} \cdot \frac{1}{1+\exp(x\beta_{\text{co,c}})} + \frac{\exp(x\gamma_{\text{nt}})}{1+\exp(x\gamma_{\text{at}})+\exp(x\gamma_{\text{nt}})} \cdot \frac{1}{1+\exp(x\beta_{\text{nt}})}} \right)^{1-y}.$$

Similarly deriving for units with $Z_i = 1$ and $W_i^{obs} = 1$.

(4) Results (Estimates of LATE)

	MLE (flat prior)	Post Mode	Quantiles of Posterior Distribution with Model for Potential Outcomes given Covariates								
			No Cov			Parallel			Unrestricted		
			$q^{.025}$	**med**	$q^{.975}$	$q^{.025}$	**med**	$q^{.975}$	$q^{.025}$	**med**	$q^{.975}$
τ_{late}	−0.11	−0.10	−0.32	**−0.15**	−0.02	−0.32	**−0.14**	−0.01	−0.48	**−0.16**	0.13
ITT_{W}	0.12	0.13	0.09	**0.12**	0.15	−0.03	**−0.02**	−0.00	0.05	**0.10**	0.13
ITT_{Y}	−0.01	−0.01	−0.04	**−0.02**	−0.00	0.09	**0.12**	0.15	−0.04	**−0.02**	0.01
$\mathbb{E}[Y_i(0)\|G_i = \text{co}]$	0.11	0.15	0.06	**0.19**	0.35	0.69	**0.71**	0.72	0.01	**0.22**	0.53
$\mathbb{E}[Y_i(1)\|G_i = \text{co}]$	0.00	0.06	0.00	**0.03**	0.09	0.16	**0.18**	0.19	0.00	**0.04**	0.31
$\mathbb{E}[Y_i(0)\|G_i = \text{nt}]$	0.08	0.08	0.06	**0.07**	0.08	0.06	**0.18**	0.35	0.06	**0.07**	0.08
$\mathbb{E}[Y_i(1)\|G_i = \text{at}]$	0.11	0.11	0.08	**0.09**	0.10	0.00	**0.04**	0.09	0.06	**0.09**	0.10

25.6 Models for the Influenza vaccination data

(4) Results (Marginal posterior distribution of τ_{late})

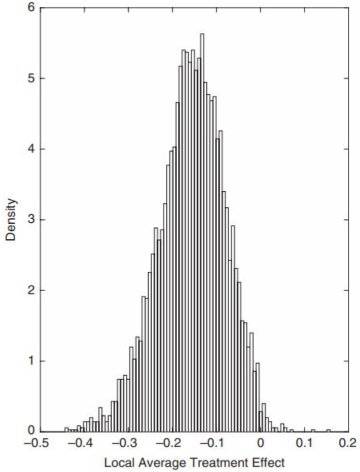

Figure 25.1. Histogram-based estimate of the distribution of the LATE, influenza vaccination data

25.6 Models for the Influenza vaccination data

What is the advantage of using model-based approach rather than IV estimate?

- When we use this data and perform the moment-based, we obtain $\mathbf{E}(Y_i(1)|G_i = co) = -0.077 < 0$ which should be larger than or equal to zero. That is, the IV estimate does not impose this restriction.
- Model-based approach naturally restrict.