Lecture 9
Non-compliance in randomized experiments, instrumental variables
Part I

Outline

- Non-compliance in randomized experiment
  - Principal stratification
  - The monotonicity and exclusion restriction assumptions
  - CATE estimand and the moment-based estimator
  - Uncertainty quantification of the moment-based estimator
- Textbook Chapters: Imbens and Rubin Chapters 23.1-23.7, 23.9 & 24.1-24.5, Peng Chapter 21.1-21.2

Ideal randomized experiment

- We have for now only considered an ideal randomized experiment
  - No loss to follow-up
  - Full adherence to the assigned treatment over the duration of the study
    ex. most severely ill individuals in the control group tend to seek a heart outside of the study.
  - No measurement errors
    ex. The PCR tests of COVID-19 may introduce false signals (depending on virus loading) when evaluating the causal effect of vaccine
  - A single version of treatment: different dosage of a drug
  - Double-blind assignment
    in real life, both patients and doctors are aware of the received treatment

The Sommer-Zeger vitamin A supplement data

- Sommer and Zeger study the effect of vitamin A supplements on infant mortality in Indonesia
- The vitamin supplements were randomly assigned to villages, but some of the individuals in villages assigned to the treatment group failed to receive them.
- None of the individuals assigned to the control group received the supplements

- $N = 23{,}682$ infants
- Outcome: binary variable indicating survival of an infant

- $W_i^{\text{obs}} \in \{0,1\}$ whether the infant receives the vitamin supplement or not
- $Z_i \in \{0,1\}$ whether the infant is assigned to the treatment group or not

- We ignore the fact that treatment assignment is at the village level (clustered randomized experiment) and consider the experiment as from a completely randomized experiment

The Sommer-Zeger vitamin A supplement data

- In principle, 8 different possible values of the triple $(Z_i, W_i^{\text{obs}}, Y_i^{\text{obs}})$
- Non-compliance: $Z_i \neq W_i^{\text{obs}}$

Assignment Z_i	Vitamin Supplements W_i^{obs}	Survival Y_i^{obs}	Number of Units ($N = 23{,}682$)
0	0	0	74
0	0	1	11,514
1	0	0	34
1	0	1	2385
1	1	0	12
1	1	1	9663

Three types of traditional analyses

Method	Estimate	Calculation	Row Comparison
ITT	0.0026	$= \frac{2385 + 9663}{12 + 9663 + 34 + 2385} - \frac{11514}{74 + 11514}$	3, 4, 5, & 6 vs. 1 & 2
As-treated	0.0065	$= \frac{9663}{12 + 9663} - \frac{11514 + 2385}{74 + 11514 + 34 + 2385}$	5 & 6 vs. 1, 2, 3, & 4
Per-protocol	0.0052	$= \frac{9663}{12 + 9663} - \frac{11514}{74 + 11514}$	5 & 6 vs. 1 & 2

Assignment Z_i	Vitamin Supplements W_i^{obs}	Survival Y_i^{obs}	Number of Units ($N = 23{,}682$)
0	0	0	74
0	0	1	11,514
1	0	0	34
1	0	1	2385
1	1	0	12
1	1	1	9663

- **Intention-to-Treat (ITT) analysis:**
  control assigned v.s. treatment assigned
- **As-treated analysis:**
  control received v.s. treatment received
- **Per-protocol analysis:**
  control received within control assigned v.s. treatment received within treatment assigned

Non-compliance in randomized experiments

- In practice, randomized experiments are often not ideal

- Often, for ethical and logistical reasons, we cannot force all experimental units to follow the randomized treatment assignment
  - some in the treatment group refuse to take the treatment
  - some in the control group manage to receive the treatment

- Intention-to-Treat (ITT) analysis: causal effect of treatment assignment
  - ITT effect can be estimated without bias
  - ITT analysis does not yield the treatment effect

- As-treated analysis
  - comparison of the treated and untreated subjects (based on treatment received)
  - no benefit of randomization, can suffer from selection bias

- Can we provide a better estimate?

Setup of the framework

- Treatment assignment (randomized encouragement): $Z_i \in \{0,1\}$

- Potential treatment variables: $(W_i(0), W_i(1))$
  - $W_i(z) = 1$: would receive the treatment if $Z_i = z$
  - $W_i(z) = 0$: would not receive the treatment if $Z_i = z$
- Observed treatment received: $W_i^{\text{obs}} = W_i(Z_i)$

- In the non-compliance setting, there are two "treatment": assignment to treatment and receipt of treatment
- Potential outcomes: $Y_i(z, w)$ potential outcome if unit is assigned to z and receive w

- Observed outcome: $Y_i^{\text{obs}} = Y_i\big(Z_i, W_i(Z_i)\big)$
- We can also write the potential outcomes as $Y_i(z) = Y_i\big(z, W_i(z)\big)$

Underlying assumptions

- No interference assumption for $W_i(z)$ and $Y_i(z, w)$

- Randomization of the treatment assignment

$$(Y_i(0,0), Y_i(0,1), Y_i(1,0), Y_i(1,1), W_i(0), W_i(1)) \perp Z_i$$

- We don't have

$$\left(Y_i(0,0), Y_i(0,1), Y_i(1,0), Y_i(1,1)\right) \perp W_i^{\text{obs}}$$

  or

$$\left(Y_i(0,0), Y_i(0,1), Y_i(1,0), Y_i(1,1)\right) \perp W_i^{\text{obs}} | Z_i$$

  We don't know why some units comply and some units don't

- Compliance can not be controlled by randomized experiment

Intention-to-treat (ITT) effects

- ITT effect on the receipt of treatment level

$$\mathrm{ITT}_{\mathrm{W},i} = W_i(1) - W_i(0) \qquad \mathrm{ITT}_{\mathrm{W}} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{ITT}_{\mathrm{W},i} = \frac{1}{N}\sum_{i=1}^{N}\big(W_i(1) - W_i(0)\big)$$

- ITT effect on the outcome of primary interest

$$\mathrm{ITT}_{\mathrm{Y},i} = Y_i(1, W_i(1)) - Y_i(0, W_i(0)).$$

$$\mathrm{ITT}_{\mathrm{Y}} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{ITT}_{\mathrm{Y},i} = \frac{1}{N}\sum_{i=1}^{N}\big(Y_i(1, W_i(1)) - Y_i(0, W_i(0))\big)$$

Statistical analysis of ITT effects

- Statistical analyses of these effects follow exactly the same procedures as before

$$\widehat{\mathrm{ITT}}_{\mathrm{W}} = \overline{W}_1^{\mathrm{obs}} - \overline{W}_0^{\mathrm{obs}} \qquad \hat{\mathbb{V}}(\widehat{\mathrm{ITT}}_{\mathrm{W}}) = \frac{s_{W,0}^2}{N_0} + \frac{s_{W,1}^2}{N_1}$$

$$s_{W,z}^2 = \sum_{i:W_i^{\mathrm{obs}}=z} \frac{\left(W_i^{\mathrm{obs}} - \overline{W}_z^{\mathrm{obs}}\right)^2}{N_z - 1} = \frac{N_z}{N_z - 1} \overline{W}_z^{\mathrm{obs}}(1 - \overline{W}_z^{\mathrm{obs}})$$

$$\widehat{\mathrm{ITT}}_{\mathrm{Y}} = \overline{Y}_1^{\mathrm{obs}} - \overline{Y}_0^{\mathrm{obs}} \qquad \hat{\mathbb{V}}(\widehat{\mathrm{ITT}}_{\mathrm{Y}}) = \frac{s_{Y,1}^2}{N_1} + \frac{s_{Y,0}^2}{N_0}$$

- We can also use regression analyses
- Drawback is that it estimates 'programmatic effectiveness’ instead of 'biologic efficacy’

Principal stratification

- Stratify individuals based on their compliance status
- Four principal strata
  - Compliers (co) $\left(W_i(0), W_i(1)\right) = (0,1)$
  - Non-compliers (nc) $\begin{cases} \text{Always} - \text{takers (at)} \ \left(W_i(0), W_i(1)\right) = (1,1) \\ \text{never} - \text{takers (nt)} \ \left(W_i(0), W_i(1)\right) = (0,0) \\ \text{Defiers (df)} \ \left(W_i(0), W_i(1)\right) = (1,0) \end{cases}$
- Principal stratification depends on latent states (potential outcomes) of units!!

		$W_i(1)$	
		0	1
$W_i(0)$	0	nt	co
	1	df	at

Principal stratification

- Can not decide which principal strata each unit belong to simply based on the observed data
  - **one-sided compliance**: control group can never receive the treatment, but treatment group may not follow the assignment

		Assignment Z_i	
		0	1
Receipt of treatment W_i^{obs}	0	nt/co	nt
	1	–	co

  - In general

		Z_i	
		0	1
W_i^{obs}	0	nt/co	nt/df
	1	at/df	at/co

ITT effect decomposition

- Denote the proportion of individuals that fall into each strata as $\pi_c, \pi_a, \pi_n, \pi_d$
  - For one-sided compliance data, $\pi_a = \pi_d = 0$
- Define the average ITT effect for each strata
  - For the treatment received $\mathrm{ITT}_{W,c}, \mathrm{ITT}_{W,a}, \mathrm{ITT}_{W,n}, \mathrm{ITT}_{W,d}$

$$\mathrm{ITT}_{W,c} = 1, \mathrm{ITT}_{W,a} = 0, \mathrm{ITT}_{W,n} = 0, \mathrm{ITT}_{W,d} = -1$$

  - For the primary outcome $\mathrm{ITT}_c, \mathrm{ITT}_a, \mathrm{ITT}_n, \mathrm{ITT}_d$

- For the ITT effect on treatment received

$$\mathrm{ITT}_W = \sum_{i=1}^{N} \mathrm{ITT}_{W,i} = \pi_c \mathrm{ITT}_{W,c} + \pi_a \mathrm{ITT}_{W,a} + \pi_n \mathrm{ITT}_{W,n} + \pi_d \mathrm{ITT}_{W,d} = \pi_c - \pi_d$$

- For the ITT effect on primary outcome

$$\mathrm{ITT}_Y = \sum_{i=1}^{N} \mathrm{ITT}_{Y,i} = \pi_c \mathrm{ITT}_c + \pi_a \mathrm{ITT}_a + \pi_n \mathrm{ITT}_n + \pi_d \mathrm{ITT}_d$$

Instrumental variables (IV)

Assumptions for Z_i being a valid IV:

- Randomization: $Z_i \in \{0,1\}$ are randomized
- Monotonicity: no defiers $\pi_d = 0$ or $W_i(0) \le W_i(1)$ for all i
- Exclusion restriction: instrument affects the outcome only through treatment

$$Y_i(1, w) = Y_i(0, w)$$

  - For always takers

$$\mathrm{ITT}_{Y,i} = Y_i\big(1, W_i(1)\big) - Y_i\big(0, W_i(0)\big) = Y_i(1,1) - Y_i(0,1) = 0$$

    so $\mathrm{ITT}_a = 0$
  - For never takers

$$\mathrm{ITT}_{Y,i} = Y_i\big(1, W_i(1)\big) - Y_i\big(0, W_i(0)\big) = Y_i(1,0) - Y_i(0,0) = 0$$

    so $\mathrm{ITT}_n = 0$
  - For compliers

$$\mathrm{ITT}_{Y,i} = Y_i\big(1, W_i(1)\big) - Y_i\big(0, W_i(0)\big) = Y_i(1,1) - Y_i(0,0)$$

    ITT_c is the average ``biological efficacy'' of the treatment on compliers

- Relevance: $\pi_c > 0$

Instrumental variables

Assumptions of Z_i being a valid IV :

- Randomization: $Z_i \in \{0,1\}$ are randomized
- Monotonicity: no defiers $\pi_d = 0$ or $W_i(0) \leq W_i(1)$ for all i
- Exclusion restriction: instrument affects the outcome only through treatment

$$Y_i(1, w) = Y_i(0, w)$$

- Relevance: $\pi_c > 0$

- Then $\mathrm{ITT}_W = \pi_c$ and $\mathrm{ITT}_Y = \pi_c \mathrm{ITT}_c + \pi_a \mathrm{ITT}_a + \pi_n \mathrm{ITT}_n + \pi_d \mathrm{ITT}_d = \pi_c \mathrm{ITT}_c$

- IV estimand: ITT_c Complier average treatment effect (CATE)

$$\mathrm{CATE} = \mathrm{ITT}_c = \frac{\mathrm{ITT}_Y}{\mathrm{ITT}_W}$$

- We can identify ITT_Y and ITT_W, so ITT_c is also identifiable
- CATE $\neq$ ATE unless ATE for noncompliers equals CATE

The monotonicity assumption

- Monotonicity: no defiers $\pi_d = 0$ or $W_i(0) \leq W_i(1)$ for all i

- Defiers are individuals who never follow treatment assignment no matter what treatment assignment is

- For one-sided compliance data, monotonicity is always satisfied

- Check the monotonicity assumption in general:
  - $\mathrm{ITT}_W = \pi_c - \pi_d > 0$ if $\pi_d = 0$, so if we can reject the null that $\mathrm{ITT}_W \geq 0$, then monotonicity assumption must fail
  - Otherwise, the monotonicity assumption is not testable

  - Need to decide whether the monotonicity assumption is reasonable or not based on domain knowledge

The exclusion restriction assumption

- Exclusion restriction: instrument affects the outcome only through treatment
$$Y_i(1, w) = Y_i(0, w)$$

- Double-blinding in experiments guarantees exclusion restriction

- The assumption in general is not testable, and need subject-matter knowledge to judge
- The subject-matter knowledge needed is often more subtle than that required to evaluate SUTVA

Moment-based IV estimator

- Causal estimand assuming a super population

$$\text{CATE} = \frac{\text{ITT}_Y}{\text{ITT}_W} = \frac{\mathbb{E}(Y_i(1) - Y_i(0))}{\mathbb{E}(W_i(1) - W_i(0))}$$

- Method-of-moment estimator:

$$\hat{\tau}^{iv} = \frac{\widehat{\text{ITT}}_Y}{\widehat{\text{ITT}}_W}$$

Simplification under one-sided compliance:

- As $W_i(0) \equiv 0$, we have

$$\widehat{\text{ITT}}_W = \bar{W}_1^{\text{obs}} - \bar{W}_0^{\text{obs}} = \bar{W}_1^{\text{obs}}$$

proportions of units who follow the assignment in the treated group

Result in Sommer-Zeger Vitamin Supplement data

ITT Estimates:

- $N_1 = 12 + 9663 + 34 + 2385 = 12094,\ N_0 = 74 + 11514 = 11588$
- $\widehat{\mathrm{ITT}}_W = \overline{W}_1^{\mathrm{obs}} = \frac{12+9663}{N_1} = 0.8$
- $\widehat{\mathrm{ITT}}_Y = \frac{2385+9663}{N_1} - \frac{11514}{N_0} = 0.0026$

CATE estimate:

- $\hat{\tau}^{iv} = \frac{0.0026}{0.8} = 0.0032$

ITT	0.0026
As-treated	0.0065
Per-protocol	0.0052

- ITT estimate is biased down
- The as-protocol or as-treated estimates are possibly biased up

Assignment Z_i	Vitamin Supplements W_i^{obs}	Survival Y_i^{obs}	Number of Units ($N = 23{,}682$)
0	0	0	74
0	0	1	11,514
1	0	0	34
1	0	1	2385
1	1	0	12
1	1	1	9663

Uncertainty of the CATE estimator

- Method-of-moment estimator: $\hat{\tau}^{iv} = \frac{\widehat{\mathrm{ITT}}_Y}{\widehat{\mathrm{ITT}}_W}$

- How to estimate the variance of $\hat{\tau}^{iv}$?
  - Estimation of $\widehat{\mathrm{ITT}}_Y$ and $\widehat{\mathrm{ITT}}_W$ are correlated because they use the same dataset

- When the number of units N is large
  - $\widehat{\mathrm{ITT}}_Y$ and $\widehat{\mathrm{ITT}}_W$ are close to the true values ITT_Y and ITT_W

$$\widehat{\mathrm{ITT}}_Y = \mathrm{ITT}_Y + O\left(\frac{1}{\sqrt{N}}\right), \qquad \widehat{\mathrm{ITT}}_W = \mathrm{ITT}_W + O\left(\frac{1}{\sqrt{N}}\right)$$

  - Perform Taylor expansion of $\hat{\tau}^{iv}$ at ITT_Y and ITT_W:

$$\frac{\widehat{\mathrm{ITT}}_Y}{\widehat{\mathrm{ITT}}_W} = \frac{\mathrm{ITT}_Y}{\mathrm{ITT}_W} + \frac{1}{\mathrm{ITT}_W}\left(\widehat{\mathrm{ITT}}_Y - \mathrm{ITT}_Y\right) - \frac{\mathrm{ITT}_Y}{\mathrm{ITT}_W^2}\left(\widehat{\mathrm{ITT}}_W - \mathrm{ITT}_W\right) + O\left(\frac{1}{N}\right)$$

  - Then

$$\mathbb{V}(\hat{\tau}^{iv}) \approx \frac{1}{\mathrm{ITT}_W^4}\left\{\mathrm{ITT}_W^2 \mathbb{V}\left(\widehat{\mathrm{ITT}}_Y\right) + \mathrm{ITT}_Y^2 \mathbb{V}\left(\widehat{\mathrm{ITT}}_W\right) - 2\mathrm{ITT}_Y \mathrm{ITT}_W \mathrm{Cov}\left(\widehat{\mathrm{ITT}}_W, \widehat{\mathrm{ITT}}_Y\right)\right\}$$

Uncertainty of the CATE estimator

- Another equivalent way to get the formula of $\mathbb{V}(\hat{\tau}^{iv})$ (see Section 21.2.2 of Peng's book)

- When N is large, $\widehat{\mathrm{ITT}}_W = \mathrm{ITT}_W + O\left(\frac{1}{\sqrt{N}}\right)$ thus (Slutsky's theorem):

$$\hat{\tau}^{iv} - \mathrm{ITT}_c = \frac{\widehat{\mathrm{ITT}}_Y - \mathrm{ITT}_c \widehat{\mathrm{ITT}}_W}{\widehat{\mathrm{ITT}}_W} \approx \frac{\widehat{\mathrm{ITT}}_Y - \mathrm{ITT}_c \widehat{\mathrm{ITT}}_W}{\mathrm{ITT}_W}$$

- Then as $\mathrm{ITT}_c = \frac{\mathrm{ITT}_Y}{\mathrm{ITT}_W}$

$$\mathbb{V}\left(\hat{\tau}^{iv} - \mathrm{ITT}_c\right) \approx \frac{\mathbb{V}\left(\widehat{\mathrm{ITT}}_Y - \mathrm{ITT}_c \widehat{\mathrm{ITT}}_W\right)}{\mathrm{ITT}_W^2} = \mathbb{V}\left(\widehat{\mathrm{ITT}}_Y\right) + \mathrm{ITT}_c^2 \mathbb{V}\left(\widehat{\mathrm{ITT}}_W\right) - 2\mathrm{ITT}_c \mathrm{Cov}\left(\widehat{\mathrm{ITT}}_W, \widehat{\mathrm{ITT}}_Y\right)$$

$$= \frac{1}{\mathrm{ITT}_W^4}\left\{\mathrm{ITT}_W^2 \mathbb{V}\left(\widehat{\mathrm{ITT}}_Y\right) + \mathrm{ITT}_Y^2 \mathbb{V}\left(\widehat{\mathrm{ITT}}_W\right) - 2\mathrm{ITT}_Y \mathrm{ITT}_W \mathrm{Cov}\left(\widehat{\mathrm{ITT}}_W, \widehat{\mathrm{ITT}}_Y\right)\right\}$$

- Same formula as before

Estimate the covariance

- Plug-in estimator of $\mathbb{V}(\hat{\tau}^{iv})$:

$$\widehat{\mathbb{V}}\left(\hat{\tau}^{iv}\right) \approx \frac{1}{\widehat{\mathrm{ITT}}_W^4}\left\{\widehat{\mathrm{ITT}}_W^2 \widehat{\mathbb{V}}\left(\widehat{\mathrm{ITT}}_Y\right) + \widehat{\mathrm{ITT}}_Y^2 \widehat{\mathbb{V}}\left(\widehat{\mathrm{ITT}}_W\right) - 2\widehat{\mathrm{ITT}}_Y \widehat{\mathrm{ITT}}_W \widehat{\mathrm{Cov}}\left(\widehat{\mathrm{ITT}}_W, \widehat{\mathrm{ITT}}_Y\right)\right\}$$

- The covariance between $\widehat{\mathrm{ITT}}_Y$ and $\widehat{\mathrm{ITT}}_W$:

$$\mathrm{Cov}\left(\widehat{\mathrm{ITT}}_W, \widehat{\mathrm{ITT}}_Y\right) = \mathrm{Cov}\left(\bar{W}_1^{\mathrm{obs}} - \bar{W}_0^{\mathrm{obs}}, \bar{Y}_1^{\mathrm{obs}} - \bar{Y}_0^{\mathrm{obs}}\right)$$

- We have

$$\bar{W}_1^{\mathrm{obs}} - \bar{W}_0^{\mathrm{obs}} = \frac{1}{N_1}\sum_{i=1}^{N} Z_i W_i(1) - \frac{1}{N_0}\sum_{i=1}^{N}(1 - Z_i) W_i(0)$$

$$\bar{Y}_1^{\mathrm{obs}} - \bar{Y}_0^{\mathrm{obs}} = \frac{1}{N_1}\sum_{i=1}^{N} Z_i Y_i(1) - \frac{1}{N_0}\sum_{i=1}^{N}(1 - Z_i) Y_i(0)$$

- Completely randomized experiment:

$$Z_i \perp \left(W_i(0), W_i(1), Y_i(1), Y_i(0)\right)$$

- It can be shown that (condition on Z_i first)

$$\mathrm{Cov}\left(\bar{W}_1^{\mathrm{obs}} - \bar{W}_0^{\mathrm{obs}}, \bar{Y}_1^{\mathrm{obs}} - \bar{Y}_0^{\mathrm{obs}}\right) = \frac{\mathrm{Cov}\left(Y_i(1), W_i(1)\right)}{N_1} + \frac{\mathrm{Cov}\left(Y_i(0), W_i(0)\right)}{N_0}$$

Estimate the covariance

- To estimate the covariance $\text{Cov}\big(Y_i(z), W_i(z)\big)$ for $z = 0,1$:

$$\widehat{\text{Cov}}\big(Y_i(z), W_i(z)\big) = \frac{1}{N_z - 1}\sum\nolimits_{i:Z_i=z}\big(W_i^{\text{obs}} - \overline{W}_z^{\text{obs}}\big)\big(Y_i^{\text{obs}} - \overline{Y}_z^{\text{obs}}\big)$$

- So, the plug-in estimator is

$$\widehat{\text{Cov}}\big(\widehat{\text{ITT}}_W, \widehat{\text{ITT}}_Y\big) = \sum\nolimits_{z=0}^{1} \frac{\sum_{i:Z_i=z}\big(W_i^{\text{obs}} - \overline{W}_z^{\text{obs}}\big)\big(Y_i^{\text{obs}} - \overline{Y}_z^{\text{obs}}\big)}{N_z(N_z - 1)}$$

- 95% confidence interval of CATE: $\left[\hat{\tau}^{iv} - 1.96\sqrt{\widehat{\mathbb{V}}(\hat{\tau}^{iv})}, \hat{\tau}^{iv} + 1.96\sqrt{\widehat{\mathbb{V}}(\hat{\tau}^{iv})}\right]$

- Under one-sided compliance
  - $\widehat{\mathbb{V}}\big(\widehat{\text{ITT}}_W\big) = \frac{s_{W,1}^2}{N_1} = \frac{\overline{W}_1^{\text{obs}}(1-\overline{W}_1^{\text{obs}})}{N_1 - 1}$ as $s_{W,0}^2 = 0$

  - $\widehat{\text{Cov}}\big(\widehat{\text{ITT}}_W, \widehat{\text{ITT}}_Y\big) = \frac{\sum_{i:Z_i=1}\left(W_i^{\text{obs}} - \overline{W}_1^{\text{obs}}\right)\left(Y_i^{\text{obs}} - \overline{Y}_1^{\text{obs}}\right)}{N_1(N_1 - 1)}$

Result in Sommer-Zeger Vitamin Supplement data

ITT Estimates:

- $N_1 = 12094$, $N_0 = 11588$
- $\widehat{\text{ITT}}_W = \bar{W}_1^{\text{obs}} = \frac{12+9663}{N_1} = 0.8$, $\widehat{\mathbb{V}}\left(\widehat{\text{ITT}}_W\right) = \frac{\bar{W}_1^{\text{obs}}(1-\bar{W}_1^{\text{obs}})}{N_1-1} = \frac{0.2*0.8}{12093} = 0.0036^2$
- $\widehat{\text{ITT}}_Y = \frac{2385+9663}{N_1} - \frac{11514}{N_0} = 0.0026$, $\widehat{\mathbb{V}}\left(\widehat{\text{ITT}}_Y\right) = \sum_{z=0}^{1} \frac{\bar{Y}_z^{\text{obs}}(1-\bar{Y}_z^{\text{obs}})}{N_z-1} = 0.0009^2$
- 95% CI of ITT_Y: $(0.0008, 0.0044)$

CATE estimate:

- $\hat{\tau}^{iv} = \frac{0.0026}{0.8} = 0.0032$

- $\widehat{\text{Cov}}\left(\widehat{\text{ITT}}_W, \widehat{\text{ITT}}_Y\right) = -0.0000017$ (correlation -0.05)
- $\widehat{\mathbb{V}}\left(\hat{\tau}^{iv}\right) = 0.0012^2$

- 95% CI of CATE: $(0.0010, 0.0055)$

Assignment Z_i	Vitamin Supplements W_i^{obs}	Survival Y_i^{obs}	Number of Units ($N = 23{,}682$)
0	0	0	74
0	0	1	11,514
1	0	0	34
1	0	1	2385
1	1	0	12
1	1	1	9663