

Homework 1

Gifty Osei

2026-02-10

Acknowledgement / Disclosure Statement

I discussed this homework with Jacob Antici at some point specifically on understanding some of the problems/questions, however my work is completely independent. I also used Chat-GPT to fine tune my writings and occasionally in fixing bugs in my code.

Part 1 - Part 2

Homework 1

Part 1: Simple theory. All estimands are defined under the super population framework. Please give specific reasons for each step.

1. In a randomized trial with a treatment group ($Z = 1$) and a control ($Z = 0$) group. Define ATE by $\text{ATE} = \mathbb{E}[Y(1) - Y(0)]$. Under the assumption of randomization, $Z \perp \{Y(1), Y(0)\}$, show that

$$\text{ATE} = \mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0].$$

Solⁿ

Under randomization, we know $Z \perp X$ and $Z \perp V$ and given $Z \perp \{Y(1), Y(0)\}$
So;

By SuTVA;

$$Y = ZY(1) + (1-Z)Y(0) \text{, meaning}$$

$$Y = Y(1) \text{ if } Z=1 \quad \text{or} \quad Y = Y(0) \text{ if } Z=0$$

We can write by the assumption;

$$Y(1) = Y|Z=1 ; E[Y(1)] = E[Y|Z=1]$$

so by $Z \perp Y(1) \Rightarrow$ distribution of $Y(1)$ is not affected by Z

Similarly; Consistent for $Z=0$ and $Z \perp Y(0)$

$$\therefore \text{ATE} = E[Y(1) - Y(0)]$$

$$= E[Y(1)] - E[Y(0)]$$

$$\text{ATE} = E[Y|Z=1] - E[Y|Z=0]$$

□

2. In the same randomized trial, we define ATT as $\text{ATT} = \mathbb{E}[Y(1) - Y(0)|Z = 1]$ and ATC as $\text{ATC} = \mathbb{E}[Y(1) - Y(0)|Z = 0]$, show that

$$\text{ATE} = \text{ATT} = \text{ATC}.$$

Sol:

We define ;

$$\text{ATT} = E[Y(1) - Y(0) | Z=1] ; \text{ATC} = E[Y(1) - Y(0) | Z=0]$$

Under the assumption of SUTVA & randomization ;
 $Z \perp\!\!\!\perp \{Y(0), Y(1)\}$

so

$$\text{ATT} = E[Y(1) - Y(0) | Z=1] \stackrel{\substack{\text{b/c of independence} \\ \downarrow}}{=} E[Y(1) - Y(0)] = \text{ATE}$$

by independence

from part 1

$$\text{ATC} = E[Y(1) - Y(0) | Z=0] \stackrel{\substack{\downarrow \\ \text{from part 1}}}{=} E[Y(1) - Y(0)] = \text{ATE}$$

Hence

$$\text{ATE} = \text{ATT} = \text{ATC}$$

□

3. In an observational study, assuming unconfoundedness, $\{Y(1), Y(0)\} \perp Z | X$, show that

$$\text{ATE} = \mathbb{E}_X\{\mathbb{E}(Y|X, Z=1)\} - \mathbb{E}_X\{\mathbb{E}(Y|X, Z=0)\}.$$

Now we assume weak unconfoundedness, $Y(0) \perp Z | X$, show that

$$\text{ATT} = \mathbb{E}[Y|Z=1] - \mathbb{E}_X\{\mathbb{E}(Y|X, Z=0)|Z=1\}.$$

We call the above set of simple identities the *identification conditions* (since the right hand side only involves the observed data and can be estimated).

Sol:

So recall; $\text{ATE} = E[Y(1) - Y(0)]$ and Law of iterated expectations

$$E[Y(1) - Y(0)] = E_x[E(Y(1) - Y(0)) | x]$$

b/c $E[x] = E[E(x|y)]$
if $x \perp y$

from part 1 and 2

$$E[Y(1) - Y(0) | x] = E[Y(1) | x] - E[Y(0) | x]$$

by Conditional expectation

now by Conditional unconfoundedness;

$$E[Y(1) | x] = E[Y(1) | x, Z=1]$$

b/c $\{Y(1), Y(0)\} \perp Z | x$

and

$$E[Y(0) | x] = E[Y(0) | x, Z=0]$$

by Consistency;

$$E[Y(1) | x, Z=1] = E[Y | x, Z=1]$$

$$E[Y(0) | X, Z=0] = E[Y | X, Z=0]$$

So

$$\text{ATE} = E\left(E[Y | X, Z=1] - E[Y | X, Z=0]\right)$$

under Superpopulation

$$\text{ATE} = E_x\left[E(Y | X, Z=1)\right] - E\left[E(Y | X, Z=0)\right]$$

by $0 < P(Z=z | X=x) < 1$ on the support of X .

(b) We define:

$$\text{ATT} = E[Y(1) - Y(0) | Z=1] = E[Y(1) | Z=1] - E[Y(0) | Z=1]$$

by linearity of conditional expectation;

$$E[Y(1) | Z=1] = E[Y | Z=1]$$

again:

$$E[Y(0) | Z=1] = E\left(E[Y(0) | X, Z=1] | Z=1\right)$$

by Law of iterated expectation; $Y(0) \perp Z | X$:

$$E[Y(0)|X, Z=1] = E[Y(0)|X, Z=0]$$

given X , the $E[Y(0)]$ is same for both groups; treated and Control

$$E[Y(0)|X, Z=0] = E[Y|X, Z=0]$$

when $I=0$, $Y=Y(0)$

$$E[Y(0)|Z=1] = E[E(Y|X, Z=0)|Z=1]$$

Together;

$$ATT = E[Y|Z=1] - E[E(Y|X, Z=0)|Z=1]$$

so $X|Z=1$

$$ATT = E[Y|Z=1] - E_{x|z=1}[E(Y|X, Z=0)]$$

Part 2

1. Write out the definition of the *unconfoundedness assumption* (e.g. the assumption of no unmeasured confounders). Using direct calculations, show that the unconfoundedness assumption holds.

Solⁿ:

Unconfoundedness assumption : $\{Y_i(1), Y_i(0)\} \perp\!\!\!\perp I_i | X_i$

The table gives the joint distribution :

$$P[Y_i(1) = y_1, Y_i(0) = y_0, X = x, I = z]$$

with binary variables.

So: $P[I = 1 | Y_i(1), Y_i(0), X] = P[I = 1 | X]$

For $X = 1$:

$$P[I = 1, X = 1] = \frac{1}{6} + \frac{1}{9} + \frac{1}{18} = \frac{6}{18} = \frac{1}{3}$$

$$P[X = 1] = \frac{1}{2}$$

so $P[I = 1 | X = 1] = \frac{P[I = 1, X = 1]}{P(X = 1)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$

For $X = 0$:

$$P[I = 1, X = 0] = \frac{1}{36} + \frac{1}{18} + \frac{1}{12} = \frac{6}{36} = \frac{1}{6}$$

$$P(X = 0) = \frac{1}{2}$$

$$P[Z=1, X=0] = \frac{1/6}{1/2} = \frac{1}{3}$$

Now we check $P(Z=1 | Y(1), Y(0), X)$

Case $X=1$

If $(Y(1), Y(0)) = (1, 1)$ and $X=1$ then $Z=1 : \frac{1}{6}$ and

$Z=0 : \frac{1}{12}$

so $P(Z=1 | Y(1)=1, Y(0)=1, X=1) = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{12}} = \frac{\frac{1}{6}}{\frac{1}{4}} = \frac{2}{3}$

If $[Y(1), Y(0)] = (1, 0)$, $X=1$

$Z=1 : \frac{1}{9}$ and $Z=0 : \frac{1}{18}$

$$P[Z=1 | 1, 0, X=1] = \frac{\frac{1}{9}}{\frac{1}{9} + \frac{1}{18}} = \frac{\frac{2}{18}}{\frac{2}{18} + \frac{1}{18}} = \frac{2}{3}$$

If $(Y(1), Y(0)) = (0, 0)$, $X=1$;

$Z=1 : \frac{1}{18}$, $Z=0 : \frac{1}{36}$

$$P[Z=1 | 0, 0, X=1] = \frac{\frac{1}{18}}{\frac{1}{18} + \frac{1}{36}} = \frac{\frac{2}{36}}{\frac{2}{36} + \frac{1}{36}} = \frac{2}{3}$$

Case $X=0$

. $(1,1), X=0 : Z=1 : \frac{1}{18}$

$$P[Z=1 | 1,1, X=0] = \frac{\frac{1}{36}}{\frac{1}{36} + \frac{1}{18}} = \frac{1}{3}$$

. $(1,0), X=0 : Z=1 : \frac{1}{18}, Z=0 : \frac{1}{9}$

$$P[Z=1 | 1,0, X=0] = \frac{\frac{1}{18}}{\frac{1}{18} + \frac{1}{9}} = \frac{1}{3}$$

. $(0,0), X=0 : Z=1 : \frac{1}{12}, Z=0 : \frac{1}{6}$

$$P(Z=1 | 0,0, X=0) = \frac{\frac{1}{12}}{\frac{1}{12} + \frac{1}{6}} = \frac{1}{3}$$

So $P(Z=1 | X=1) = \frac{2}{3}$ and $P(Z=1 | X=0) = \frac{1}{3}$

$\therefore \{Y(1), Y(0)\} \perp Z | X \text{ holds}$

2. Suppose you didn't have X (X was unmeasured). Using direct calculations, show that randomization of treatment (e.g. *marginal unconfoundedness*) does not hold.

Sol:

Marginal Unconfoundedness requires;

$$\{Y(1), Y(0)\} \perp Z \Leftrightarrow P[Z=1 | Y(1), Y(0)] = P[Z=1]$$

$$P[Z=1] = \frac{1}{2}$$

Now pick $[Y(1), Y(0)] = (1, 1)$

- o $P(Y(1)=1, Y(0)=1, Z=1) = \frac{1}{6} + \frac{1}{36} = \frac{6}{36} + \frac{1}{36} = \frac{7}{36}$
- o $P(Y(1)=1, Y(0)=1, Z=0) = \frac{1}{12} + \frac{1}{18} = \frac{3}{36} + \frac{2}{36} = \frac{5}{36}$

So

$$P(Z=1 | Y(1)=1, Y(0)=1) = \frac{\frac{7}{36}}{\frac{7}{36} + \frac{5}{36}} = \frac{7}{12} \neq \frac{1}{2}$$

hence marginal unconfoundedness (randomization without X) fails.

3. What additional assumption is needed to obtain the distribution of the observed data (Y, Z, X) ?

Under this assumption, deduce the distribution of the observed data. Produce a table with all 8 combinations of Y, Z and X together with their probabilities.

Solⁿ:

The needed assumption to go from potential outcome
to observed data.

$$Y(1), Y(0) \xrightarrow{\text{Consistency}} Y^{\text{obs}}$$

$$Y = Y(Z) - \text{No interference}$$

We compute $P(Y=y, Z=z, X=x)$ for all 8 combinations

Like

$$P(Y=1, Z=1, X=1) = P(Y(1)=1, X=1, Z=1) = \frac{1}{6} + \frac{1}{9} = \frac{5}{18}$$

No	X	Z	Y	$P(Y, Z, X)$
1	1	1	1	$\frac{1}{6} + \frac{1}{9} = \frac{5}{18}$
2	1	1	0	$\frac{1}{18}$
3	1	0	1	$\frac{1}{12}$
4	1	0	0	$\frac{1}{18} + \frac{1}{36} = \frac{1}{12}$
5	0	1	1	$\frac{1}{18} + \frac{1}{36} = \frac{1}{12}$
6	0	1	0	$\frac{1}{12}$
7	0	0	1	$\frac{1}{18}$
8	0	0	0	$\frac{1}{6} + \frac{1}{9} = \frac{5}{18}$

They sum to 1,,

4. Using direct calculations, show that $\mathbb{E}[Y|Z=1] \neq \mathbb{E}[Y(1)]$ and $\mathbb{E}[Y|Z=0] \neq \mathbb{E}[Y(0)]$. Give an intuitive explanation for this result.

Solⁿ:
Find $E[Y(1)]$ and $E[Y(0)]$, since outcomes are binary,

$$E[Y(1)] = P(Y(1)=1), \quad E[Y(0)] = P[Y(0)=1]$$

from table:

$Y(1)=1$ happens in 8 rows and the sum is $\frac{2}{3}$ so

$$E[Y(1)] = \frac{2}{3}$$

$Y(0)=1$ occurs only in rows with $[Y(1), Y(0)] = (1, 1)$

Summing gives $\frac{1}{3}$ so

$$E[Y(0)] = \frac{1}{3}$$

Hence:

$$E[Y|Z=1] = ?$$

$$P[Y=1, Z=1] = P(Y=1, Z=1, X=1) + P(Y=1, Z=1, X=0)$$

$$= \frac{5}{18} + \frac{1}{12}$$

$$= \frac{13}{36}$$

$$P[Z=1] = \frac{1}{2} ;$$

$$E[Y|Z=1] = P[Y=1 | Z=1] = \frac{13/36}{18/36} = \frac{13}{18}$$

$$\frac{13}{18} \neq \frac{2}{3}$$

$$- E[Y|Z=0] ?$$

$$\begin{aligned} P[Y=1, Z=0] &= P[Y=1, Z=0, X=1] + P[Y=1, Z=0, X=0] \\ &= \frac{1}{12} + \frac{1}{18} \\ &= \frac{5}{36} \end{aligned}$$

$$\begin{aligned} E[Y|Z=0] &= \frac{5/36}{18/36} \\ &= \frac{5}{18} \neq \frac{1}{3} \end{aligned}$$

Z is not randomized marginally; it depends on X
Note: $P(Z=1 | X=1) = \frac{2}{3}$ but $P(Z=1 | X=0) = \frac{1}{3}$
and also X predicts outcomes;

- $E[Y(1) | X=1] = \frac{5}{6}$ vs $E[Y(1) | X=0] = \frac{1}{2}$
 - $E[Y(0) | X=1] = \frac{1}{2}$ vs $E[Y(0) | X=0] = \frac{1}{6}$
- So the treated group has more $X=1$ (a "higher outcome group) too large and $E[Y|Z=0]$ too small relative to the true potential means.

5. Using direct calculations, show that $\mathbb{E}[Y(1)] = \mathbb{E}_X\{\mathbb{E}(Y|Z=1, X)\}$ and $\mathbb{E}[Y(0)] = \mathbb{E}_X\{\mathbb{E}(Y|Z=0, X)\}$. Further, use direct calculations to show that these are both equal to the inverse probability weighted mean, $\mathbb{E}[I(Z=z)Y/P(Z=z|X)]$.

Solⁿ: Show $E[Y(1)] = E_X\{E(Y|Z=1, X)\}$, first we find

$$- E[Y|Z=1, X=1] = P[Y=1 | Z=1, X=1] = \frac{P(Y=1, Z=1, X=1)}{P(Z=1, X=1)}$$

$$= \frac{\frac{5}{18}}{\left(\frac{5}{18} + \frac{1}{18}\right)} = \frac{\frac{5}{18}}{\frac{6}{18}} = \frac{5}{6}$$

$$- E[Y|Z=1, X=0] = \frac{\frac{1}{2}}{\frac{1}{12} + \frac{1}{12}} = \frac{\frac{1}{2}}{\frac{1}{6}} = \frac{1}{6}$$

Now average over X and $P(X=1) = P(X=0) = \frac{1}{2}$

$$f_x\{E(Y|Z=1, X)\} = \frac{1}{2} \cdot \frac{5}{6} + \frac{1}{2} \cdot \frac{1}{2} = \frac{8}{12} = \frac{2}{3} = E[Y(1)]$$

Part 2, $E[Y(0)] = f_x\{E(Y|Z=0, X)\}$

$$\cdot E[Y|Z=0, X=1] = \frac{\frac{1}{12}}{\frac{1}{12} + \frac{1}{12}} = \frac{1}{2}$$

$$\cdot E[Y|Z=0, X=0] = \frac{\frac{1}{18}}{\frac{1}{18} + \frac{5}{18}} = \frac{\frac{1}{18}}{\frac{6}{18}} = \frac{1}{6}$$

and average over X :

$$f_x\{E(Y|Z=0, X)\} = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{4} + \frac{1}{12} = \frac{1}{3} = E[Y(0)]$$

for the IPW estimand $E\left[\frac{I(Z=z)Y}{P(Z=z|X)}\right]$

for $Z=1$:

$$\begin{aligned} E\left[\frac{I(Z=1)Y}{P(Z=1|X)}\right] &= \sum_x \frac{P(Y=1, Z=1, X=x)}{P(Z=1|X)} \\ &= \frac{\frac{5}{18}}{\frac{2}{3}} + \frac{\frac{1}{12}}{\frac{1}{3}} = \frac{\frac{15}{36}}{\frac{12}{36}} + \frac{\frac{3}{12}}{\frac{12}{36}} = \frac{\frac{8}{12}}{\frac{12}{36}} = \frac{2}{3} = E[Y(1)] \end{aligned}$$

For $Z = 0$

$$E\left[\frac{I(Z=0)Y}{P(Z=0|X)}\right] = \sum_x \frac{P(Y=1, Z=0, X=x)}{P(Z=0 | X=x)}$$

$$= \frac{1/2}{1/3} + \frac{1/8}{2/3} = \left(\frac{1}{12} \cdot 3\right) + \left(\frac{1}{18} \cdot \frac{3}{2}\right) = \frac{3}{12} + \frac{1}{12} = \frac{1}{3} = E[Y_{(0)}]$$

So for this distribution:

$$E[Y_{(1)}] = E_x\{E(Y|Z=1, X)\} = E\left[\frac{I(Z=1)Y}{P(Z=1|X)}\right]$$

$$E[Y_{(0)}] = E_x\{E(Y|Z=0, X)\} = E\left[\frac{I(Z=0)Y}{P(Z=0|X)}\right]$$

6. Calculate the causal risk difference, risk ratio, odds ratio. Is there any advantage to using one or the other causal estimand for binary outcome?

Soln.

$$E[Y(1)] = P[Y(1)=1] = \frac{2}{3}, E[Y(0)] = P[Y(0)=1] = \frac{1}{3}$$

We let :

$$P_1 : P(Y(1)=1) = E[Y(1)] = \frac{2}{3}; P_0 : P(Y(0)=1) = E[Y(0)] = \frac{1}{3}$$

Causal Risk Difference (RD)

$$RD = P_1 - P_0$$

$$= \frac{2}{3} - \frac{1}{3}$$

$$= \frac{1}{3}$$

Causal Risk Ratio (RR)

$$RR = \frac{P_1}{P_0} = \frac{\frac{2}{3}}{\frac{1}{3}} = 2$$

Causal Odds Ratio (OR)

Odds under treatment :

$$Odds_{\text{treatment}} = \frac{P_1}{1-P_1} = \frac{\frac{2}{3}}{\frac{1}{3}} = 2$$

odds under Control

$$\text{odds}_0 = \frac{P_0}{1 - P_0} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$$

So

$$OR = \frac{\text{odds}_1}{\text{odds}_0} = \frac{2}{\frac{1}{2}} = 4$$

So

$$RD = \frac{1}{3}, \quad RR = 2, \quad OR = 4$$

For binary outcomes;

RD gives the absolute effect which is very interpretable for public health or decision making

RR is the relative effect (doubling risk), interpretable but can be harder to model directly but still common

OR: Quite convenient for logistic regression and case-control designs; but can exaggerate effects when outcome is common, and is non-collapsible.

7. Calculate $\mathbb{E}(Y|Z=1, X) = \mathbb{E}[Y(1)|X]$ and $\mathbb{E}(Y|Z=0, X) = \mathbb{E}[Y(0)|X]$, for each $X = 0, 1$.

Calculate the conditional risk difference, risk ratio, odds ratio for each $X = 0, 1$. Is there an interaction between Z and X and why?

Sol:

From the table we have;

Conditional means (risks)

For $X=1$:

$$\mathbb{E}[Y|Z=1, X=1] = \frac{5}{6}, \quad \mathbb{E}[Y|Z=0, X=1] = \frac{1}{2}$$

For $X=0$:

$$\mathbb{E}[Y|Z=1, X=0] = \frac{1}{2}, \quad \mathbb{E}[Y|Z=0, X=0] = \frac{1}{6}$$

Unconfoundedness holds given X ;

$$\mathbb{E}[Y|Z=1, X] = \mathbb{E}[Y(1)|X], \quad \mathbb{E}[Y|Z=0, X] = \mathbb{E}[Y(0)|X]$$

So define;

$$P_{1x} : P(Y(1)=1 | X=x), \quad P_{0x} : P(Y(0)=1 | X=x)$$

We have;

$$x=1 : P_{11} = \frac{5}{6}, \quad P_{01} = \frac{1}{2}$$

$$x=0 : P_{10} = \frac{1}{2}, \quad P_{00} = \frac{1}{6}$$

Conditional causal effects for $X=1$

RD:

$$RD(1) = P_{11} - P_{01} = \frac{5}{6} - \frac{1}{2} = \frac{2}{6} = \frac{1}{3}$$

$$RR(1) = \frac{P_{11}}{P_{01}} = \frac{\frac{5}{6}}{\frac{1}{2}} = \frac{10}{6} = \frac{5}{3}$$

OR(1)

Odds treatment at $X=1$

$$X=1; \quad \frac{P_{11}}{1-P_{11}} = \frac{\frac{5}{6}}{\frac{1}{6}} = 5$$

Odds Control at $X=1$

$$\frac{P_{01}}{1-P_{01}} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1$$

So

$$OR(1) = \frac{5}{1} = 5$$

Conditional Causal effects for $X=0$

$$RD(0) = P_{10} - P_{00} = \frac{1}{2} - \frac{1}{5} = \frac{1}{3}$$

$$RR(0) = \frac{P_{10}}{P_{00}} = \frac{\frac{1}{2}}{\frac{1}{5}} = 3$$

OR

Odds treatment at $X=0$:

$$\frac{P_{10}}{1-P_{10}} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1$$

Odds Control at $X=0$

$$\frac{P_{00}}{1-P_{00}} = \frac{\frac{1}{5}}{\frac{4}{5}} = \frac{1}{4}$$

So $OR(0) = \frac{1}{\frac{1}{4}} = 4$

Interaction between Z and X depends on the scales;

- Additive Scale (RD): $RD(1) = \frac{1}{3}$, $RD(0) = \frac{1}{3}$

Same; no interaction

- Multiplicative risk Scale (RR)

$$RR(1) = \frac{5}{3}, \quad RR(0) = 3$$

Different \rightarrow there exists an interaction

- Odds scale:

$$OR(1) = 5, \quad OR(0) = 5$$

Same \rightarrow no interaction on logit scale

hence no interaction on RD and OR scales; interaction on RR scale.

Part 3

```
## Data from table
placebo <- c(8.62, 1.48, 8.93, 9.57, 2.65, 7.3)
vitamin_A <- c(0.06, 1.72, 2.19, 7.32, 7.53, 7.62)

## Outcome vector
y <- c(placebo, vitamin_A)

n <- length(y)
# number treated
m <- 6
idx_all <- 1:n

## Observed assignment: last 6 are Vitamin_A, first 6 are Placebo
treat_obs <- 7:12

## Given test statistic
T_obs <- mean(y[treat_obs]) - mean(y[-treat_obs])
T_obs

## [1] -2.018333

# function to compute the test statistic
T_stat <- function(treat_idx, y) {
  mean(y[treat_idx]) - mean(y[-treat_idx])
}

a

## a

## Exact randomization p-value (two-tailed)

## finding the space of completely randomized assignment 6 x 924 matrix; each
## column = treated indices
treat_sets <- combn(n, m)

## Fast computation using sums: If  $S_t = \sum(y \text{ in treated})$ ,  $total = \sum(y)$ ,
##  $mean\_treat = S_t/m$   $mean\_ctrl = (total - S_t)/(n-m)$   $T = mean\_treated -$ 
##  $mean\_control$  treated sums for each assignment
S_t <- colSums(matrix(y[treat_sets], nrow = m))
S_all <- sum(y)
# 924 test statistics
T_all <- (S_t/m) - ((S_all - S_t)/(n - m))

## Two-sided exact p-value we are computing the probability that we observed
## something at least or as extreme than the t-stats if the null hypothesis is
## true.
p_exact <- mean(abs(T_all) >= abs(T_obs))
p_exact
```

```

## [1] 0.2705628

1b

## Bootstraps samples approximation with B=1000

set.seed(789)

B <- 1000

T_mc <- replicate(B, {
  tr <- sample(idx_all, m, replace = FALSE)
  T_stat(tr, y)
})

p_mc <- mean(abs(T_mc) >= abs(T_obs))
se_mc <- sqrt(p_mc * (1 - p_mc)/B)

p_mc

## [1] 0.279

se_mc

## [1] 0.01418305

## 95% CI
c(lower = p_mc - 1.96 * se_mc, upper = p_mc + 1.96 * se_mc)

##      lower      upper
## 0.2512012 0.3067988

Comment on Comparing

We draw 1000 bootstrap samples from the distribution of the statistics under the null hypothesis. The p-value from the simulation is 0.279 and that from the exact p-value (fisher's test) from 1a is 0.2705628. Since the p-value is larger than 0.05 in both cases, then we fail to reject the sharp null hypothesis of no treatment effect at a significance level of 0.05 and conclude that there there is no improvement by taking vitamin A in children.

c

## Two-sample t-test (Vitamin_A vs Placebo)

t_out_welch <- t.test(vitamin_A, placebo, alternative = "two.sided", var.equal = FALSE)

t_out_pooled <- t.test(vitamin_A, placebo, alternative = "two.sided", var.equal = TRUE)

t_out_welch$p.value

## [1] 0.3367803

```

```

t_out_pooled$p.value

## [1] 0.3367792

## output

cat("\nObserved T (Vitamin_A - Placebo):", T_obs, "\nExact Fisher p-value:", p_exact,
    "\nMC Fisher p-value (B=1000):", p_mc, " (MC SE:", se_mc, ")",
    "\nT-test p-value (Welch):", t_out_welch$p.value, "\n")

## 
## Observed T (Vitamin_A - Placebo): -2.018333
## Exact Fisher p-value: 0.2705628
## MC Fisher p-value (B=1000): 0.279 (MC SE: 0.01418305 )
## T-test p-value (Welch): 0.3367803

```

1d

For the exact fisher test; It uses the assignment mechanism by completely randomizing with 6 treated out of 12 children. It also uses the sharp null for each unit i , $Y_i(1) = Y_i(0)$ (i.e. completely impute). Under this approach, the missing Potential Outcome are imputed exactly, so every assignment implies a fully known set of observed outcomes.

For the simulation test; it is an approximation to the exact randomization distribution. The approximation is computational: instead of enumerating all 924 assignments, it samples 1000 assignments from the mechanism. Hence, it approximates the exact.

And for the t-test; It replaces the randomization distribution of the statistic with an analytic/parametric distribution. It is justified by modeling assumptions and large sample CLT. Also it is framed under super-population framework where randomization is not only from the treatment.

So (c) approximate (a) using a parametric/asymptotic distributional approximation rather than exact randomization.

2

```

## Now we create a paired data Each row is a matched pair: (placebo_i, vitA_i)
## are the two children in pair i
pairs <- data.frame(pair = 1:6, ctrl = placebo, trt = vitamin_A)

## Outcomes in each pair as a vector of length 2
y_pair <- as.matrix(pairs[, c("ctrl", "trt")])

```

2a

```

## Observed assignment: given in table within-pair differences (treated -
## control)
d_obs <- pairs$trt - pairs$ctrl

# equals diff-in-means overall under 1-1 pairing
T_obs <- mean(d_obs)
T_obs

## [1] -2.018333

```

```

## Under pairwise randomization, in each pair the sign can flip:
## treated-control difference is either +d_i or -d_i, each with prob 1/2
## independently.

sign_mat <- as.matrix(expand.grid(rep(list(c(-1, +1)), 6)))
# each row gives mean(s_i * d_i)
T_all <- as.numeric(sign_mat %*% d_obs)/6

p_exact_pair <- mean(abs(T_all) >= abs(T_obs))
p_exact_pair

```

[1] 0.375

2b

```

set.seed(985)

B <- 1000
T_mc <- replicate(B, {
  s <- sample(c(-1, +1), size = 6, replace = TRUE) # independent sign flips within each pair
  mean(s * d_obs)
})

p_mc_pair <- mean(abs(T_mc) >= abs(T_obs))
se_mc_pair <- sqrt(p_mc_pair * (1 - p_mc_pair)/B)

p_mc_pair

```

[1] 0.405

se_mc_pair

[1] 0.01552337

2c

```

# t - test by paired

t_paired <- t.test(d_obs, mu = 0, alternative = "two.sided", paired = FALSE)
t_paired$p.value

```

[1] 0.365161

2d

- (a) exact conditions on the matched-pair assignment mechanism (within each pair, 1 treated, 1 control) and uses fisher's sharp null to make all potential outcomes known under any assignment implies exact randomization distribution over 64 assignments.
- (b) Bootstrap is the same mechanism plus same sharp null, but samples 1000 of the 64 possibilities with replacement. We could sample all 64 possibilities which will mean a pure Monte Carlo approximation to (a).

- (c) Paired-test replaces the exact randomization distribution of \bar{d} with a t-distribution under assumptions like approximate normality of the within-pair differences. It is for short a parametric approximation to (a).

3

One good reason is the dataset is one realization: Any single sample can yield higher/lower p-values under different valid tests purely by chance. Design comparisons must be made in expectation or across repeated samples. The other is that paring only helps if the matching variables is prognostic for outcomes. If the selected X doesn't strongly predict Y (or matching inducing little reduction within-pair variability), pairwise randomization may not increase power and can look "worse" in a given sample. Completely randomization uses 924 assignment spaces as compared to 64 under pairwise so there is a distribution change.

4. By writing the test statistic (difference in sample means) as an explicit function of the potential outcomes and the assignment indicators, show that, under complete randomization, the statistic is unbiased for the average treatment effect. Do the same for the pairwise randomization.

Sol:

Showing that the test statistic is unbiased for ATE:

$$\text{ie } E[\hat{t}] = \text{ATE}$$

So

$$Y_i = Z_i Y_i(1) + (1 - Z_i) Y_i(0)$$

$$\begin{aligned} \hat{t} &= \frac{1}{m} \sum_{i=1}^n Z_i Y_i - \frac{1}{n-m} \sum_{i=1}^n (1 - Z_i) Y_i, \quad \begin{matrix} m \rightarrow \text{group} \\ \text{size} \\ n \rightarrow \text{total} \end{matrix} \\ &= \bar{Y}_1 - \bar{Y}_0 \end{aligned}$$

i) Under Completely randomized
Each unit i has equal chance of receiving treatment

$$\text{So } P[Z_i = 1] = \frac{m}{n}, \quad E[Z_i] = \frac{m}{n} \quad \text{and} \quad E[1 - Z_i] = \frac{n-m}{n}$$

$$\begin{aligned} E[\hat{t} | Y_i(1), Y_i(0)] &= \frac{1}{m} \sum_{i=1}^m E[Z_i] Y_i(1) - \frac{1}{n-m} \sum_{i=1}^{n-m} E[1 - Z_i] Y_i(0) \\ &= \frac{1}{m} \cdot \frac{m}{n} \sum_{i=1}^m Y_i(1) - \frac{1}{n-m} \cdot \frac{n-m}{n} \sum_{i=1}^{n-m} Y_i(0) \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{n} \sum_{i=1}^n \gamma_i(1) - \frac{1}{n} \sum_{i=1}^n \gamma_i(0) \\
 &= \frac{1}{n} \sum_{i=1}^n \left\{ \gamma_i(1) - \gamma_i(0) \right\} \\
 &= ATE
 \end{aligned}$$

hence unbiased under both finite and superpopulation framework.

ii Pairwise Randomization

$n = 2K$ with pairs $= k = 1, \dots, K$, each pair has 2 units say a, b . For each pair exactly 1 is treated

$$D_k = \begin{cases} \gamma_{ka}(1) - \gamma_{kb}(0), & a \text{ treated} \\ \gamma_{kb}(1) - \gamma_{ka}(0) & b \text{ treated} \end{cases}$$

The overall diff. in means equals the average of

D_k'

$$\frac{1}{K} = \frac{1}{K} \sum_{k=1}^K D_k$$

Within pair k , the labels of treatment status occurs with probability $\frac{1}{2}$

$$E[D_k | \gamma_{i(1)}, \gamma_{i(0)}] = \frac{1}{2} \left\{ \gamma_{ka}(1) - \gamma_{kb}(0) \right\} + \frac{1}{2} \left\{ \gamma_{kb}(1) - \gamma_{ka}(0) \right\}$$

Over pairs:

$$E[\bar{\tau}_k | \gamma_{i(1)}, \gamma_{i(0)}] = \frac{1}{2K} \sum_{k=1}^K \left[(\gamma_{ka}(1) - \gamma_{ka}(0)) + (\gamma_{kb}(1) - \gamma_{kb}(0)) \right]$$

$$= \frac{1}{n} \sum_{i=1}^n (\gamma_{i(1)} - \gamma_{i(0)}) \quad b/c \quad n=2K$$

5. Researchers are interested in obtaining some more information about the treatment effect. In particular, they now wish to test a sharp null hypothesis specifying a constant additive treatment effect:

$$H_0 : Y_i(1) = Y_i(0) + 2.5 \text{ for all } i$$

Could you Answer 1(a)-1(b) to report the P-values for this new hypothesis test?

Solⁿ:

This is Fisher's sharp null, so we can adjust outcomes to a common baseline so all units have same outcome regardless of assignment.

We define: $\tilde{Y}_i^{(0)} = Y_i - 2.5 Z_i$

Under H_0 ; $\tilde{Y}_i^{(0)} = Y_i(0)$ for each unit, so we can run the same randomization test as before on $\tilde{Y}^{(0)}$

Yes I can do 1a - 1b from part 1.

```

## Observed data combinesd
y_obs <- c(placebo, vitamin_A)
z_obs <- c(rep(0, 6), rep(1, 6))

tau0 <- 2.5

## Adjust outcomes under sharp null
y_adj <- y_obs - tau0 * z_obs

# Observed stats derived on adjusted outcomes
T_obs_adj <- mean(y_adj[z_obs == 1]) - mean(y_adj[z_obs == 0])
T_obs_adj

```

[1] -4.518333

```

# 924 assignments by the number of assignment formula
n <- length(y_adj)
m <- 6
treat_sets <- combn(n, m)
S_t <- colSums(matrix(y_adj[treat_sets], nrow = m))
S_all <- sum(y_adj)
T_all_adj <- (S_t/m) - ((S_all - S_t)/(n - m))

p_exact_const <- mean(abs(T_all_adj) >= abs(T_obs_adj))
p_exact_const

```

[1] 0.05194805

```

# bootstrap draws

set.seed(456)
B <- 1000
idx_all <- 1:n
T_mc_adj <- replicate(B, {
  tr <- sample(idx_all, m, replace = FALSE)
  mean(y_adj[tr]) - mean(y_adj[-tr])
})
p_mc_const <- mean(abs(T_mc_adj) >= abs(T_obs_adj))
p_mc_const

```

[1] 0.061