p8122_hw1_jsg2145

Jared Garfinkel

9/14/2020

Problem 1

Read in data

```
counterfactuals = tibble(individual = c(1:8),
                         y_0 = c(0, 1, 0, 1, 1, 0, 1, 0),
                         y_1 = c(0, 0, 1, 0, 0, 1, 0, 0))

counterfactuals
```

```
## # A tibble: 8 x 3
##   individual   y_0   y_1
##        <int> <dbl> <dbl>
## 1          1     0     0
## 2          2     1     0
## 3          3     0     1
## 4          4     1     0
## 5          5     1     0
## 6          6     0     1
## 7          7     1     0
## 8          8     0     0
```

Problem a

The treatment group, y__1, can have no causal effect on health outcome if the treatment difference is 0, a beneficial effect if the treatment difference is 1, or a harmful causal effect if the difference is -1, meaning an individual in a counterfactual was cured in the control group, y__0, but sick in the treatment group.

```
treffect = counterfactuals %>%
  mutate(difference = y_1 - y_0)

treffect
```

```
## # A tibble: 8 x 4
##   individual   y_0   y_1 difference
##        <int> <dbl> <dbl>      <dbl>
## 1          1     0     0          0
## 2          2     1     0         -1
## 3          3     0     1          1
```

```
## 4            4      1      0          -1
## 5            5      1      0          -1
## 6            6      0      1           1
## 7            7      1      0          -1
## 8            8      0      0           0
```

Problem b

The Average Caussal Effect (ACE) is -1/4 =

$$E[y_1] - E[y_0]$$

```
aceffect = treffect %>%
  summarize(sum_y_0 = sum(y_0),
            sum_y_1 = sum(y_1),
            ace = sum_y_1/8 - sum_y_0/8)

aceffect
```

```
## # A tibble: 1 x 3
##   sum_y_0 sum_y_1   ace
##     <dbl>   <dbl> <dbl>
## 1       4       2 -0.25
```

Problem b

In the treatment assignment, understanding that we cannot know the counterfactual, the association of the treatment with the outcome is -2/4 =

$$E[y_1|A = 1] - E[y_0|A = 0]$$

```
trgroup_0 = treffect %>%
  mutate(tr = c(1, 0, 1, 1, 0, 0, 0, 1)) %>%
  filter(tr == 0) %>%
  summarize(sum_y_0 = sum(y_0))

trgroup_1 = treffect %>%
  mutate(tr = c(1, 0, 1, 1, 0, 0, 0, 1)) %>%
  filter(tr == 1) %>%
  summarize(sum_y_1 = sum(y_1))

trgroups = tibble(trgroup_0, trgroup_1) %>%
  mutate(effect = (sum_y_1 - sum_y_0)/4)

trgroups
```

```
## # A tibble: 1 x 3
##   sum_y_0 sum_y_1 effect
##     <dbl>   <dbl>  <dbl>
## 1       3       1   -0.5
```

Problem d

Create a partition and label the treatment groups.

```
set.seed(719)
group_0 = counterfactuals %>%
  sample_frac(size = .5, replace = FALSE) %>%
  mutate("group" = 0)

group_1 = counterfactuals %>%
  anti_join(group_0) %>%
  mutate("group" = 1)

groups = bind_rows(group_0, group_1)

groups
```

```
## # A tibble: 8 x 4
##   individual   y_0   y_1 group
##        <int> <dbl> <dbl> <dbl>
## 1          4     1     0     0
## 2          5     1     0     0
## 3          7     1     0     0
## 4          2     1     0     0
## 5          1     0     0     1
## 6          3     0     1     1
## 7          6     0     1     1
## 8          8     0     0     1
```

Find the observed treatment effects.

```
treatment_0 = groups %>%
  filter(group == 0) %>%
  summarize(sum_y_0 = sum(y_0))

treatment_1 = groups %>%
  filter(group == 1) %>%
  summarize(sum_y_1 = sum(y_1))

treatments = tibble(treatment_0, treatment_1) %>%
  mutate(effect = (sum_y_1 - sum_y_0)/4)

treatments
```

```
## # A tibble: 1 x 3
##   sum_y_0 sum_y_1 effect
##     <dbl>   <dbl>  <dbl>
## 1       4       2   -0.5
```

The observed effect of treatment is -0.5 =

$$E[y_1|A=1] - E[y_0|A=0]$$

Problem 2

During a check-up, a physician finds that his patient's blood pressure levels are too low. He prescribes medication at a high dose and asks her to be re-tested in a month. At the second test, the patient's blood pressure levels are now too high, so the physician switches her to a low dose of medication and again asks her to be re-tested in a month. At the third test, the patient's blood pressure levels are perfect, and so the doctor decides that she should stay at the low dose indefinitely, with no further testing.

Problem a

The units are the patient at different times.

Problem b

The treatment is a dose of medicine. It can be HIGH or LOW.

Problem c

The potential outcomes are blood pressures. They can take values of cured (1) or not cured (0)

Problem d

The difference in treatments for the patient was 1 when they took a low dose compared to when they took a high dose. The doctor calculated that the patient was cured on the low dose.

So, the average causal effect (ACE) is 1 =

$$E[y_{low}] - E[y_{high}]$$

```
patient0 = tibble(ptid = c(0, 0), y_high = c(0, 0), y_low = c(1, 1))

patient0
```

```
## # A tibble: 2 x 3
##    ptid y_high y_low
##   <dbl>  <dbl> <dbl>
## 1     0      0     1
## 2     0      0     1
```

```
aceffect = patient0 %>%
  summarize(sum_y_high = sum(y_high),
            sum_y_low = sum(y_low),
            ace = sum_y_low/2 - sum_y_high/2)

aceffect
```

```
## # A tibble: 1 x 3
##   sum_y_high sum_y_low   ace
##        <dbl>     <dbl> <dbl>
## 1          0         2     1
```

Problem e

SUTVA is not plausible because the patient at the first visit was used to influence the next dosage. We should assume that the treatment assignment was not based on the outcome of the high dose so that SUTVA is plausible.

Problem f

SUTVA must be plausible because the counterfactual cannot be observed in real life. The difference between potential outcomes is only possible when the assumptions of SUTVA hold.

Problem g

The assignment is individualistic because there is only one participant.

It's not probabilistic because the doctor will change the treatment group to put the patient on the treatment that cures them.

The assignment mechanism is not unconfounded because the doctor will assign the patient to a treatment group based on their reaction to the dosage.

This assignment mechanism is known but not controlled.

Problem h

Assign the patient a lower dose and increase it until the blood pressure is normal in a planned way.