

Problem Set 4: Experimentation

1 Purpose

The purpose of this problem set is to assess your understanding of one key method of quantitative public opinion research: experimental design and analysis.

2 Your Task

1. Read the Gerber and Green (2008) chapter provided on Moodle.
2. In your own words, explain the “potential outcomes framework” of causal inference and explain how experiments a way to identify causal effects.

Solution:

There are multiple ways to think about this.

- Randomized experiments do not allow us to see individual-level causal effects unless we assume unit homogeneity. In all other cases, randomization allows us to assess average causal effects of a treatment on an outcome.
- Randomized experiments randomly sample potential outcomes in an unbiased manner, allowing for estimation of an average treatment effect.
- Randomized experiments randomly expose one of multiple potential outcomes for each individual in the sample.
- Randomized experiments eliminate selection bias into treatment assignment.
- Randomized experiments eliminate confounding.

3. A researcher wants to understand how a televised party leaders debate affected citizens’ vote intentions and considers two alternative research designs. The first design involves interviewing a representative sample of citizens, asking whether they watched the debate, and comparing vote intentions among viewers and non-viewers. The second design involves recruiting a non-representative sample of citizens into a laboratory session where one half of participants are randomly assigned to watch the debate and the other half is randomly assigned to watch a non-political program. Vote intentions are measured at the end of the laboratory session. Discuss the trade-offs involved in these designs, including what would be required to obtain an estimate of the causal effect of the debate on vote intentions in each design.

Solution:

There are numerous trade-offs here and there is no correct answer. The first design is a population-based survey with no experimental component. The second design is a laboratory experiment involving a non-representative sample.

The former design has claims to “external” validity because sample characteristics are unbiased estimates of population parameters. In attempting to infer causation, however, the design does not necessarily address selection bias: it does not induce debate viewing, so we cannot know

why people watched the debate and whether any of those factors also explain vote intentions (e.g., some third variable explains both forms of political engagement). The measurement of vote intention occurs some time after the debate, possibly suggesting durability of any influence.

The latter design has claims to “internal” validity because randomized assignment to debate viewing enables an estimation of an average causal effect of viewing the debate but that effect does not necessarily reflect the average effect in any population. The measurement of vote intentions occurs immediately. In both designs, this measurement of vote intention may not say anything about actual voting behavior.

Both designs likely present other trade-offs: feasibility, cost, etc.

4. Consider an experiment on 500 individuals in which one group is randomly assigned to read a treatment message from David Cameron support the “Remain” position in the upcoming European Union referendum and another group is assigned to a control condition that receives no information. Measures of opinions for the European Union are recorded for both groups on a 0 to 1 scale, with higher scores indicating greater favorability toward British membership in the European Union.

- (a) Assuming the treatment group mean score was 0.68 and the control group mean score was .51, what is the average treatment effect? Is this substantively large or small?

Solution:

The sample average treatment effect is simply the mean difference: $0.68 - 0.51 = 0.17$.

Deciding whether this is large or small requires some kind of benchmark for comparison. We could say this is 17% of the scale (from 0 to 1). Or, we could compare it to the standard deviation of the outcome to express a “standardized mean-difference” but that information is not provided here. Assuming the standard deviation were 0.4, then the effect size would be $\frac{0.17}{0.40} = 0.425$, and so forth.

- (b) Assuming the t-statistic for the mean-difference is 1.76, should we consider this effect to be statistically large and distinguishable from zero?

Solution:

Answering this question technically requires consulting a t-distribution. You can find one on Wikipedia. Given the very large sample size, use the ∞ row of the table. A *t*-statistic of this size is considered statistically distinguishable from zero in the case of a one-tailed test ($p < 0.05$) but not in a two-tailed test ($p < 0.10$). If we had a strong directional prediction that the treatment outcome would be higher than the control outcome, we could consider the one-tailed test appropriate but probably not otherwise.

5. The statistical power of a two-sample t-test (which is, in essence, the power of a posttest-only, two-group experimental design) is influenced by four things: the size of each experimental group, the difference-in-means (i.e., difference in mean values of the outcome in the two groups), the variance of the outcome measure, and alpha (the significance level or “Type 1” error probability).

- (a) If α (the Type 1 error probability) is 0.05, how often should we expect to find a “statistically significant” effect size when one is not present?

Solution:

This is simply 0.05 or 5% of the time.

- (b) If you increase the size of your treatment groups in an experiment while the expected effect size remains unchanged, what happens to the power of your experiment? Are you more or less likely to obtain a “false negative” result? What about “false positives”?

Solution:

The power of the test increases. Recall the definition of power:

$$Power = \phi \left(\frac{|\mu_1 - \mu_0| \sqrt{N}}{2\sigma} - \phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right) \quad (1)$$

Without worrying about all the details of the equation, note that sample size is in the numerator of the first term, so more observations means more power. This directly translates into a lower likelihood of false negatives ($1 - \beta$) where power is denoted β .

Technically, this is unrelated to the false positive rate, which is a function of the selected significance level, α , only.

- (c) Imagine we are expecting to find a small effect but we can only collect a small number of observations in our experiment, so the minimum detectable effect size in our study is larger than the effect size we would expect to observe given our theory. If our experiment reveals an effect that is statistically distinguishable from zero, what are the two possible interpretations of this result?

Solution:

- i. The effect is actually larger than we expected.
- ii. The effect in our experiment is a massive overestimate.

We cannot distinguish which alternative is correct.

6. Download the data for the Klar (2014) article we read in class during Week 4.¹ Using this data set and the code provided below (in R or Stata), reproduce the analysis reported in all of the figures of the paper. The code below creates variable representing each of the treatment “factors” in the design — ambivalence and group type (homogeneous, heterogeneous, or none). You should use regression analysis or another appropriate technique to estimate treatment means for each outcome; which are named as follows:

- energypriority
- healthpriority
- oil
- alternative
- competition
- subsidies
- energygroup
- healthgroup

R code

```
dat <- rio::import("Klar_AJPS_4.dta")

# analysis focuses only on Democrats
dems <- dat[dat[["pid"]] %in% 1:3,]
dim(dems)
# [1] 277 49

# treatment variable
```

¹A data file can be found here: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/22892>

```

table(dems[["condition_number"]])
dems[["treatment"]] <- NA_real_
dems[["treatment"]][dems[["condition_number"]] == 2] <- 1L
dems[["treatment"]][dems[["condition_number"]] == 3] <- 4L
dems[["treatment"]][dems[["condition_number"]] == 4] <- 2L
dems[["treatment"]][dems[["condition_number"]] == 5] <- 5L
dems[["treatment"]][dems[["condition_number"]] == 6] <- 3L
dems[["treatment"]][dems[["condition_number"]] == 7] <- 6L
dems[["ambivalent"]] <- NA_real_
dems[["ambivalent"]][dems[["treatment"]] %in% 1:3] <- 1L
dems[["ambivalent"]][dems[["treatment"]] %in% 4:6] <- 0L
dems[["group_type"]] <- NA_real_
dems[["group_type"]][dems[["treatment"]] %in% c(1,4)] <- 0L
dems[["group_type"]][dems[["treatment"]] %in% c(2,5)] <- 1L # homogeneous
dems[["group_type"]][dems[["treatment"]] %in% c(3,6)] <- 2L # heterogeneous

```

Stata code

```

use "Klar_AJPS_4.dta"

* analysis focuses only on Democrats
drop if pid > 3

* treatment variable
tab condition_number

gen treatment = .
replace treatment = 1 if condition_number == 2
replace treatment = 4 if condition_number == 3
replace treatment = 2 if condition_number == 4
replace treatment = 5 if condition_number == 5
replace treatment = 3 if condition_number == 6
replace treatment = 6 if condition_number == 7

gen ambivalent = .
replace ambivalent = 1 if treatment < 4
replace ambivalent = 0 if treatment > 3

gen group_type = .
* no group
replace group_type = 0 if treatment == 1 | treatment == 4
* homogeneous
replace group_type = 1 if treatment == 2 | treatment == 5
* heterogeneous
replace group_type = 2 if treatment == 3 | treatment == 6

```

Solution: Complete results are shown at the end of this document for both R and Stata.

3 Submission Instructions

Please submit your answers as a PDF document via Moodle. It should be no more than 4 pages, single-spaced, in Times New Roman font size 12, on A4 paper with standard 2.54cm margins. This problem set is self-assessed. A solution set will be provided on the course website and the activity will be discussed in class.

4 Feedback

Group feedback will be provided during class. If you would like more specific individual feedback on your work, please ask the instructor during office hours.

5 Klar (2014) Analysis

R Code (.R)

```
# Download data from:
# https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/22892

# or use the dataverse package:
if ("Klar_AJPS_4.dta" %in% dir()) {
  if (!require("ghit")) {
    install.packages("ghit")
    library("ghit")
  }
  ghit::install_github("IQSS/dataverse-client-r")
  library("dataverse")
  d <- dataverse::get_file("Klar_AJPS_4.dta", dataset = "doi:10.7910/DVN/22892")
  writeBin(as.raw(d), con = "Klar_AJPS_4.dta")
}

# load the data
if (!require("rio")) {
  install.packages("rio")
  library("rio")
}
dat <- rio::import("Klar_AJPS_4.dta")

# analysis focuses only on Democrats
dems <- dat[dat[["pid"]] %in% 1:3,]
dim(dems)
# [1] 277 49

# treatment variable
table(dems[["condition_number"]])
dems[["treatment"]] <- NA_real_
dems[["treatment"]][dems[["condition_number"]] == 2] <- 1L
dems[["treatment"]][dems[["condition_number"]] == 3] <- 4L
dems[["treatment"]][dems[["condition_number"]] == 4] <- 2L
dems[["treatment"]][dems[["condition_number"]] == 5] <- 5L
dems[["treatment"]][dems[["condition_number"]] == 6] <- 3L
dems[["treatment"]][dems[["condition_number"]] == 7] <- 6L
dems[["ambivalent"]] <- NA_real_
dems[["ambivalent"]][dems[["treatment"]] %in% 1:3] <- 1L
dems[["ambivalent"]][dems[["treatment"]] %in% 4:6] <- 0L
dems[["group_type"]] <- NA_real_
dems[["group_type"]][dems[["treatment"]] %in% c(1,4)] <- 0L
dems[["group_type"]][dems[["treatment"]] %in% c(2,5)] <- 1L # homogeneous
dems[["group_type"]][dems[["treatment"]] %in% c(3,6)] <- 2L # heterogeneous

# Figure 1
aggregate(energypriority ~ ambivalent, data = dems, FUN = mean, na.rm = TRUE)
aggregate(healthpriority ~ ambivalent, data = dems, FUN = mean, na.rm = TRUE)
## regression version
summary(lm(energypriority ~ ambivalent, data = dems))
summary(lm(healthpriority ~ ambivalent, data = dems))

# Figure 2
aggregate(oil ~ ambivalent, data = dems, FUN = mean, na.rm = TRUE)
```

```

aggregate(alternative ~ ambivalent, data = dems, FUN = mean, na.rm = TRUE)
aggregate(competition ~ ambivalent, data = dems, FUN = mean, na.rm = TRUE)
aggregate(subsidies ~ ambivalent, data = dems, FUN = mean, na.rm = TRUE)
## regression version
summary(lm(oil ~ ambivalent, data = dems))
summary(lm(alternative ~ ambivalent, data = dems))
summary(lm(competition ~ ambivalent, data = dems))
summary(lm(subsidies ~ ambivalent, data = dems))

# Figure 3
aggregate(energypriority ~ group_type, data = dems, FUN = mean, na.rm = TRUE)
aggregate(healthpriority ~ group_type, data = dems, FUN = mean, na.rm = TRUE)
## regression version
summary(lm(energypriority ~ group_type, data = dems))
summary(lm(healthpriority ~ group_type, data = dems))

# Figure 4
aggregate(oil ~ group_type, data = dems, FUN = mean, na.rm = TRUE)
aggregate(alternative ~ group_type, data = dems, FUN = mean, na.rm = TRUE)
aggregate(competition ~ group_type, data = dems, FUN = mean, na.rm = TRUE)
aggregate(subsidies ~ group_type, data = dems, FUN = mean, na.rm = TRUE)
## regression version
summary(lm(oil ~ group_type, data = dems))
summary(lm(alternative ~ group_type, data = dems))
summary(lm(competition ~ group_type, data = dems))
summary(lm(subsidies ~ group_type, data = dems))

# Figure 5
aggregate(energypriority ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4)], FUN = mean, na.rm = TRUE)
aggregate(healthpriority ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4)], FUN = mean, na.rm = TRUE)
## regression version
summary(lm(energypriority ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4)]))
summary(lm(healthpriority ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4)]))

# Figure 6
aggregate(oil ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4)], FUN = mean, na.rm = TRUE)
aggregate(competition ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4)], FUN = mean, na.rm = TRUE)
## regression version
summary(lm(oil ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4)]))
summary(lm(competition ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4)]))

# Figure 7
aggregate(energygroup ~ ambivalent, data = dems, FUN = mean, na.rm = TRUE)
aggregate(healthgroup ~ ambivalent, data = dems, FUN = mean, na.rm = TRUE)
## regression version
summary(lm(energygroup ~ ambivalent, data = dems))
summary(lm(healthgroup ~ ambivalent, data = dems))

# Figure 8
aggregate(energygroup ~ group_type, data = dems, FUN = mean, na.rm = TRUE)

```

```

aggregate(healthgroup ~ group_type, data = dems, FUN = mean, na.rm = TRUE)
## regression version
summary(lm(energygroup ~ group_type, data = dems))
summary(lm(healthgroup ~ group_type, data = dems))

# Figure 9
aggregate(energygroup ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4),], FUN = mean, na.rm = TRUE)
aggregate(healthgroup ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4),], FUN = mean, na.rm = TRUE)
## regression version
summary(lm(energygroup ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4),]))
summary(lm(healthgroup ~ ambivalent + group_type,
          data = dems[!dems[["treatment"]] %in% c(1,4),]))

```

Stata Code (.do)

* Download data from: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/22892>

```
use "Klar_AJPS_4.dta"
```

```
* analysis focuses only on Democrats
drop if pid > 3
```

```
* treatment variable
tab condition_number
```

```
gen treatment = .
replace treatment = 1 if condition_number == 2
replace treatment = 4 if condition_number == 3
replace treatment = 2 if condition_number == 4
replace treatment = 5 if condition_number == 5
replace treatment = 3 if condition_number == 6
replace treatment = 6 if condition_number == 7
```

```
gen ambivalent = .
replace ambivalent = 1 if treatment < 4
replace ambivalent = 0 if treatment > 3
```

```
gen group_type = .
* no group
replace group_type = 0 if treatment == 1 | treatment == 4
* homogeneous
replace group_type = 1 if treatment == 2 | treatment == 5
* heterogeneous
replace group_type = 2 if treatment == 3 | treatment == 6
```

```
* Figure 1
tabstat energypriority, by(ambivalent)
tabstat healthpriority, by(ambivalent)
** regression version
reg energypriority ambivalent
reg healthpriority ambivalent
```

```
* Figure 2
tabstat oil, by(ambivalent)
```



```

tabstat alternative, by(ambivalent)
tabstat competition, by(ambivalent)
tabstat subsidies, by(ambivalent)
** regression version
reg oil ambivalent
reg alternative ambivalent
reg competition ambivalent
reg subsidies ambivalent

* Figure 3
tabstat energypriority, by(group_type)
tabstat healthpriority, by(group_type)
** regression version
reg energypriority i.group_type
reg healthpriority i.group_type

* Figure 4
tabstat oil, by(group_type)
tabstat alternative, by(group_type)
tabstat competition, by(group_type)
tabstat subsidies, by(group_type)
** regression version
reg oil i.group_type
reg alternative i.group_type
reg competition i.group_type
reg subsidies i.group_type

* Figure 5
tabstat energypriority, by(treatment)
tabstat healthpriority, by(treatment)
** regression version
reg energypriority i.treatment if treatment != 1 & treatment != 4
reg healthpriority i.treatment if treatment != 1 & treatment != 4

* Figure 6
tabstat oil, by(treatment)
tabstat competition, by(treatment)
** regression version
reg oil i.treatment if treatment != 1 & treatment != 4
reg alternative i.treatment if treatment != 1 & treatment != 4

* Figure 7
tabstat energygroup, by(ambivalent)
tabstat healthgroup, by(ambivalent)
** regression version
reg energygroup i.ambivalent
reg healthgroup i.ambivalent

* Figure 8
tabstat energygroup, by(group_type)
tabstat healthgroup, by(group_type)
** regression version
reg energygroup i.group_type
reg healthgroup i.group_type

* Figure 9
tabstat energygroup, by(treatment)
tabstat healthgroup, by(treatment)

```

```
** regression version
reg energygroup i.treatment if treatment != 1 & treatment != 4
reg healthgroup i.treatment if treatment != 1 & treatment != 4
```