Foundations of Design Based Methods

Cyrus Samii

September 01, 2020

Properties of SI and SIR

Use simulation to complete the exercises below.

Let's suppose we have a population with N units and we want to sample n of them.

```
N <- 400
n <- 20
```

Exercise 1. *Show the distributions of samples of size* $0, ..., N$ *and of the* π_i *values under SI and SIR.*

Here are our population outcomes, $(y_1, ..., y_N)$:

```
y <- rpois(N, 5)
```

Let's suppose that we are using SI to sample n values $(Y_1, ..., Y_n)$ from this population.

Exercise 2. *Use simulation to compute* $E[Y_i]$*,* $Var[Y_i]$*, and* $Cov[Y_i, Y_j]$*. Compare to the theoretically derived quantities in the slides.*

Exercise 3. *Use simulation to compute* $E[\bar{Y}]$ *and* $Var[\bar{Y}]$*. Compare this to the theoretically derived quantities in the slides.*

Exercise 4. *Show how the distribution of* $\bar{Y}$ *varies for* $n = 20, 40, 60$*. Compare the variance for these distributions to the expression* $s^2_{\bar{Y}}$ *from the text.*

Properties of Complete Random Assignment

Let's translate the results from above to the situation of a randomized experiment with complete random assignment.

Here are the potential outcomes and average treatment effect for our population:

```
y0 <- rpois(N, 5)
y1 <- y0 + 1 + rpois(N, 2)
mean(y1-y0)
```

```
## [1] 3
```

Suppose we assign treatment using complete random assignment. Let $\widehat{\beta} := \bar{Y}_1 - \bar{Y}_0$. Suppose we want to assign n units to treatment and $N - n$ to control:

```r
n <- 20
```

Exercise 5. *Use simulation to compute* $E[\widehat{\beta}]$ *and* $Var[\widehat{\beta}]$*. Comment on unbiasedness and how the variance compares to the Neyman variance,* $V_{\widehat{\beta}}$*.*

You can use the `estimatr` package to estimate the "HC2" standard error, which is equivalent to the square root of the Neyman variance.

Exercise 6. *For* $n = 15, 30, 60$*, perform the following using simulation:*

- *Show the distributions of* $\widehat{\beta}/\sqrt{\widehat{V}_{\widehat{\beta}}}$*,*
- *Show the CDF of* p *values from* t *tests of the null hypothesis* $H_0 : \widehat{\beta} = 0$*,*
- *For tests with* 95% *confidence, what is the power of these tests of the null?*

Exercise 7. *Using the MDE expression from the slides, estimate the* n *needed to have 80% power to detect the treatment effect for this population with 95% confidence. Construct a simulation to validate your calculation.*