Problem Set 3

Tessie Dong, Derek Li, Andi Liu

Jan 15th, 2024

Problem 1: Combine the NSW data in `nswre74_control.csv` and `nswre74_treated.csv` and complete Table 1. Note that variables `1-10` are *predetermined*, i.e., capture characteristics determined at or before treatment assignment; some of these variables are background characteristics (e.g., `edu`), others capture a subject's pre-RCT labor market experience (e.g., `u75`). `re78` is the observed outcome variable. `treat` is the indicator of treatment status.

```
# load and combine the data sets into one dataframe
df1 <- utils::read.csv(file = "nswre74_control.csv")
df2 <- utils::read.csv(file = "nswre74_treated.csv")
df <- rbind(df1, df2)
```

```
# count units in each sample
dplyr::tally(dplyr::group_by(df, treat))
```

```
## # A tibble: 2 x 2
##   treat     n
##   <int> <int>
## 1     0   260
## 2     1   185
```

```
# generate mean summary statistics for each variable and treatment
dplyr::summarise_all((dplyr::group_by(df, treat)), list(mean))
```

```
## # A tibble: 2 x 12
##   treat   age   edu black   hisp married nodegree  re74  re75  re78   u74   u75
##   <int> <dbl> <dbl> <dbl>  <dbl>   <dbl>    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1     0  25.1  10.1 0.827 0.108    0.154    0.835 2107. 1267. 4555. 0.75  0.685
## 2     1  25.8  10.3 0.843 0.0595   0.189    0.708 2096. 1532. 6349. 0.708 0.6
```

A completed version of Table 1 is provided on the following page.

Variable Counter	Variable Name	Variable Definition	Sample Average	
			Treated	Control
1	`age`	Age in years	25.8	25.1
2	`edu`	Education in years	10.3	10.1
3	`nodegree`	1 if education < 12		0.835
4	`black`	1 if Black		0.827
5	`hisp`	1 if Hispanic		0.108
6	`married`	1 if married		
7	`u74`	1 if unemployed in '74		
8	`u75`	1 if unemployed in '75		
9	`re74`	Real earnings in '74 (in '82 $)		
10	`re75`	Real earnings in '75 (in '82 $)		
11	`re78`	Real earnings in '78 (in '82 $)		
12	`treat`	1 if received offer of training	1	0
Sample Size			185	260

Table 1: Descriptive statistics for the NSW data by group.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla euismod, nisl nec ultricies ultricies, nunc nisl ultricies nunc, vitae ultricies nisl nisl nec nisl. Nulla facilisi. Sed euismod, nisl nec ultricies ultricies, nunc nisl ultricies nunc, vitae ultricies nisl nisl nec nisl. Nulla facilisi. Sed euismod, nisl nec ultricies ultricies, nunc nisl ultricies nunc, vitae ultricies nisl nisl nec nisl. Nulla facilisi. Sed euismod, nisl nec ultricies ultricies, nunc nisl ultricies nunc, vitae ultricies nisl nisl nec nisl. Nulla facilisi. Sed euismod, nisl nec ultricies ultricies, nunc nisl ultricies nunc, vitae ultricies nisl nisl nec nisl. Nulla facilisi. Sed euismod, nisl nec ultricies ultricies, nunc nisl ultricies nunc, vitae ultricies nisl nisl nec nisl. Nulla facilisi. Sed euismod, nisl nec ultricies ultricies, nunc nisl ultricies nunc, vitae ultricies nisl nisl nec nisl. Nulla facilisi. Sed euismod, nisl nec ultricies ultricies, nunc nisl ultricies nunc, vitae ultricies nisl nisl nec nisl. Nulla facilisi.