INFO 201 Lab 5 Demo Solutions

Yubing Tian

2/2/2020

In this demo we will use a dataset on infidelity, known as the Fair's Affairs dataset. The `Affairs` dataset is available as part of the `AER` package in **R**. This data comes from a survey conducted by *Psychology Today* in 1969.

See Greene, W.H. (2003). Econometric Analysis, 5th edition. Upper Saddle River, NJ: Prentice Hall. and Fair, R.C. (1978). A Theory of Extramarital Affairs. Journal of Political Economy, 86, 45–61. for more information.

The dataset contains various self-reported characteristics of 601 participants, including how often the respondent engaged in extramarital sexual intercourse during the past year, as well as their gender, age, year married, whether they had children, their religiousness (on a 5-point scale, from 1 = anti to 5 = very), education, occupation (Hollingshead 7-point classification with reverse numbering), and a numeric self-rating of their marriage (from 1 = very unhappy to 5 = very happy).

```r
#install.packages("AER")
library(AER)
data("Affairs")
```

```r
library(dplyr)
# Removing `occupation` because the Hollingshead 7-point classification scheme is outdated
# also reverse numbering makes it a little tedious to interpret
affairs_clean <- Affairs %>%
  select(-occupation)
```

```r
# Inspect dataset observations and variables
dim(affairs_clean)
```

```
## [1] 601   8
```

```r
# Inspect first few rows
head(affairs_clean)
```

```
##    affairs gender age yearsmarried children religiousness education rating
## 4        0   male  37        10.00       no             3        18      4
## 5        0 female  27         4.00       no             4        14      4
## 11       0 female  32        15.00      yes             1        12      4
## 16       0   male  57        15.00      yes             5        18      5
## 23       0   male  22         0.75       no             2        17      3
## 29       0 female  32         1.50       no             2        17      5
```

```
# Inspect last few rows
tail(affairs_clean)
```

```
##      affairs gender age yearsmarried children religiousness education rating
## 1935       7   male  47         15.0      yes             3        16      2
## 1938       1   male  22          1.5      yes             1        12      5
## 1941       7 female  32         10.0      yes             2        18      4
## 1954       2   male  32         10.0      yes             2        17      5
## 1959       2   male  22          7.0      yes             3        18      2
## 9010       1 female  32         15.0      yes             3        14      5
```

```
# By using the summary function we can begin to know our data.
summary(affairs_clean)
```

```
##     affairs          gender         age        yearsmarried      children
##  Min.   : 0.000   female:315   Min.   :17.50   Min.   : 0.125   no :171
##  1st Qu.: 0.000   male  :286   1st Qu.:27.00   1st Qu.: 4.000   yes:430
##  Median : 0.000                Median :32.00   Median : 7.000
##  Mean   : 1.456                Mean   :32.49   Mean   : 8.178
##  3rd Qu.: 0.000                3rd Qu.:37.00   3rd Qu.:15.000
##  Max.   :12.000                Max.   :57.00   Max.   :15.000
##  religiousness     education         rating
##  Min.   :1.000   Min.   : 9.00   Min.   :1.000
##  1st Qu.:2.000   1st Qu.:14.00   1st Qu.:3.000
##  Median :3.000   Median :16.00   Median :4.000
##  Mean   :3.116   Mean   :16.17   Mean   :3.932
##  3rd Qu.:4.000   3rd Qu.:18.00   3rd Qu.:5.000
##  Max.   :5.000   Max.   :20.00   Max.   :5.000
```

Let's write down our interpretation of our descriptive statistics:

- **Affairs** From this sample participants engaged in approximately 1.5 extramarital sexual intercourse in the last year, the minimum number of affairs was 0 and the maximum was 12. *Note* One thing that remains ambiguous is for instance whether 12 was with the same individual or multiple individuals note that the survey only asked for instances.

- **Gender** The proportion of females to males in this sample is quite balanced however at 52.4% there are slightly more females represented than males.

- **Age** The average age of survey participants was ~33 years old. The youngest participant was 17.5 years old whilst the oldest was 57.

- **Years Married** On average participants were married for 8 years and 2 months. The participants who were married the longest had been married for 15 years. I was quite surprised to see that for some survey participants they had been married for 0.125 years (i.e. for ~ 1.5 months). I suppose infidelity could occurr at any point during a marriage and there is also the possibility that these participants are in consenting non-monogamous relationships.

- **Children** The majority of survey participants - 71.5% - had children.

- **Religiousness** On average participants are non-religious, however there are participants that rate themselves as anti-religious to very religious.

- **Education** On average participants have ~ 16 years of education. If we count the first 13 as part of K-12 then this means that on average participants have some college education or completed degree programs that are less than 4 years. Participants reported between 9 to 20 years of education. This range translates to no high school education all the way to postgraduate studies.

- **Rating** On average participants rate themselves happy (4) in terms of their marriage which is quite surprising to me. I would have hypothesized for below average marriage happiness ratings.