PSY 221A Homework 3

Cristopher Garduno Luna

10/17/2017

Disclaimer: The methods used in this *C1*, *C2* and *C3* are not necessarily the most efficient for this particualr assignment, but they were written while keeping generalization in mind such that they can easily be adapted to other tasks.

Chapter 3

C1

Find the mode, median, and mean for each of the quantitative variables in Ihno's data set.

Quantitative variables: Num_cups, Phobia, Prevmath, Mathquiz, Statquiz, Exp_sqz, Hr_base, Hr_pre, Hr_post, Anx_post, Anx_base, Anx_pre, Anx_post

```
# Load data
library(haven)
```

```
## Warning: package 'haven' was built under R version 3.2.5
```

```
filelocation = "~/Desktop/UCSB/fall2017/psych221a/hw/data_hw1.sav"
dataset      = as.data.frame(read_sav(filelocation))

# quant_var is a vector with column names
# dataTable is empty dataframe for final data
quant_var = names(dataset)[which(names(dataset) == "Num_cups"):length(names(dataset))]
dataTable = data.frame(matrix(ncol = length(quant_var), nrow = 3))

# Create mode function
mymode = function(values) {
   uniq_val = unique(values)
   uniq_val[which.max(tabulate(match(values, uniq_val)))][[1]][[1]]
}

# Loop through each column of quantitative variables
for (i in 1:length(quant_var)) {
  # Create vector for current variable
  curr = c()

  # Append mode to index 1, median to index 2, and mean to index 3
  curr[1] = mymode(dataset[quant_var[i]])
  curr[2] = as.numeric(sapply(dataset[quant_var[i]], median, na.rm = TRUE))
  curr[3] = as.numeric(sapply(dataset[quant_var[i]], mean, na.rm = TRUE))

  # Append vector to datatable
  dataTable[i] = curr
}

names(dataTable) = quant_var
```

```
row.names(dataTable) = c("Mode", "Median", "Mean")
dataTable
```

```
##        Num_cups Phobia Prevmath Mathquiz Statquiz Exp_sqz Hr_base Hr_pre
## Mode       0.00   1.00     3.00 43.00000     6.00    7.00   71.00  68.00
## Median     0.00   3.00     1.00 30.00000     7.00    7.00   72.00  74.00
## Mean       0.68   3.31     1.38 29.07059     6.86    6.83   72.27  73.85
##        Hr_post Anx_base Anx_pre Anx_post
## Mode      65.0    17.00   22.00     20.0
## Median    73.0    18.00   19.00     19.0
## Mean      72.8    18.43   19.58     19.4
```

C2

Find the mode for the undergraduate major variable.

```
majorl     = c("Psychology", "Premed", "Biology", "Sociology", "Economics")
major_fac  = factor(dataset$Major,    level = c(1:5), majorl)
major_mode = mymode(major_fac)
major_mode
```

```
## [1] Psychology
## Levels: Psychology Premed Biology Sociology Economics
```

C3

Find the range, semi-interquartile range, unbiased variance, and unbiased standard deviation for each of the quantitative variables in Ihno's data set.

```
# Subset quantitative variables and initialize data table
data_quant = subset(dataset, select = quant_var)
dataTable2 = data.frame(matrix(ncol = length(quant_var), nrow = 5))
names(dataTable2) = names(data_quant)
row.names(dataTable2) = c('Range Start', 'Range End', 'Semi-Interquartile Range', 'Unbiased Variance',
                          'Unbiased Standard Deviation')

for (i in 1:length(names(data_quant))) {
  # Initialize empty vector and create current data object
  curr   = c()
  curdat = data_quant[names(data_quant)[i]]

  # Find range, semi-interquartile range, unbiased variance, unbiased standard deviation
  curr[1] = range(curdat, na.rm = TRUE)[1]
  curr[2] = range(curdat, na.rm = TRUE)[2]
  curr[3] = IQR(as.numeric(unlist(curdat)), na.rm = TRUE, type = 6)/2
  curr[4] = var(curdat, na.rm = TRUE)
  curr[5] = sd(as.numeric(unlist(curdat)), na.rm = TRUE)

  dataTable2[i] = curr
}

# Display Data
dataTable2
```

```
##                               Num_cups    Phobia Prevmath  Mathquiz
## Range Start                 0.0000000  0.000000 0.000000  9.000000
## Range End                   3.0000000 10.000000 6.000000 49.000000
## Semi-Interquartile Range    0.5000000  1.500000 0.500000  7.000000
## Unbiased Variance           0.7450505  5.973636 1.571313 89.875910
## Unbiased Standard Deviation 0.8631631  2.444102 1.253520  9.480291
##                               Statquiz   Exp_sqz   Hr_base    Hr_pre
## Range Start                   1.000000  1.000000 64.000000 62.000000
## Range End                    10.000000 11.000000 80.000000 87.000000
## Semi-Interquartile Range      1.000000  1.000000  2.000000  4.000000
## Unbiased Variance             2.889293  4.465758 10.340505 26.330808
## Unbiased Standard Deviation   1.699792  2.113234  3.215666  5.131355
##                                Hr_post Anx_base   Anx_pre  Anx_post
## Range Start                 64.000000 10.00000  8.000000  9.000000
## Range End                   86.000000 39.00000 39.000000 40.000000
## Semi-Interquartile Range     3.500000  2.00000  5.500000  3.000000
## Unbiased Variance           22.464646 18.75263 41.377374 22.747475
## Unbiased Standard Deviation  4.739688  4.33043  6.432525  4.769431
```