CKME 132 Summer 2016 - Assignment #2

Tamer Abdou

June 30, 2016

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

Use RStudio for this assignment. Edit the file `assignment-2.Rmd` and insert your R code where wherever you see the string "INSERT YOUR ANSWER HERE"

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document.

When your are done with your answers and before submitting, save the file with the following naming convention :your **Lastname_firstname**

Submit **both** the rmd and the pdf output(or word or html) files, failing to submit **both** will be subject to mark deduction.

This assignment makes use of data that is provided by the `ISwR` package.

```
library(ISwR)
```

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

Sample Question and Solution

Use `seq()` to create the vector $(1, 2, 3, \ldots, 10)$.

```
seq(1,10)
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10
```

Question 1 - 35%

In this question you will explore the relation between blood pressure and obesity. The data frame `bp.obese` measures blood pressure, obsesity, and sex on a random sample of adults.

a) Print the head of `bp.obese` to familiarize yourself with the data.

```
#Insert your answer here
```

b) Print a summary of `bp.obese`.

```
#Insert your answer here
```

Note that `sex` is a numeric vector but it should be a factor.

c) Convert `sex` from numeric to a factor type with two levels in bp.obese.

```
#Insert your answer here
```

d) Use `str` to check the structure after the conversion

```
#Insert your answer here
```

e) Use `sapply`, `any`, and `is.na` to check if there are any `NA` values in each column of `bp.obese`.

```
#Insert your answer here
```

f) Assign the obesity and blood pressure measurements to numeric vectors `obesity` and `bp`. Generate a scatter plot of `bf` on the vertical axis versus `obesity` on the horizontal axis.

```
#Insert your answer here
```

Note that blood pressure appears to be positively correlated with obesity.

g) Compute the Pearson correlation r and covariance of `obesity` and `bp`.

```
#Insert your answer here
```

h) determine the type and strength of the relationship. (note when you inserst your answer make sure to use a # at the begining)

```
#Insert your answer here
```

Question 2 - 35%

Consider the probability distribution associated with rolling 3 fair dice. We can label the faces of a single die using the numbers from 1 to 6. We can therefore label the simple events in this distribution by triples of numbers from 1 to 6. Let `d1`, `d2`, and `d3` represent the labels on each of the dice.

a) Set `d1` to the sequence (1,2,...,6) repeated 36 times.

```
#Insert your answer here
```

b) Set `d2` to the sequence consisting of 6 repetitions of the sequence in which each of the numbers (1,2,...,6) is repeated 6 times.

```
#Insert your answer here
```

c) Set `d3` to the sequence in which each of the numbers (1,2,...,6) is repeated 36 times.

```
#Insert your answer here
```

d) Create a new data frame `three.dice` from `d1`, `d2`, and `d3` and print it. Visually confirm that there are $6 \times 6 \times 6 = 216$ rows and each row contains a unique combination of dice labels.

```
#Insert your answer here
```

e) Since the dice are fair and independent, each simple event has the same probability, namely $\frac{1}{216}$. Add the column `P` to the data frame with this value.

```
#Insert your answer here
```

f) Add a new column `sum` equal to the sum of the dice labels. Add another new column `mean` equal to the average of the dice labels.

```
#Insert your answer here
```

g) Plot a probability histogram of `three.dice$sum`.

```
#Insert your answer here
```

h) Compute the probability that the sum of the dice is greater then 12 and less than 18.

HINT: Use `subset()` to select the events and sum `P`.

```
#Insert your answer here
```

i) Compute the probability that the sum is even.

```
#Insert your answer here
```

j) Compute the probability that the mean is exactly 4.

```
#Insert your answer here
```

Question 3 - 30%

a)You have two groups of distinctly different items, 10 in the first group and 8 in the second. If you select one item from each group, how many different pairs can you form?

```
#use the mn rule
#Insert your answer here
```

b)Evaluate the following permutation P_3^5

```
#Insert your answer here
```

c)In how many ways can you select five people from a group of eight if the order of selection is important?

```
#Since order is important, you use permutations
#Insert your answer here
```

d)In how many ways can you select two people from a group of 20 if the order of selection is not important?

```
#Since order is unimportant, you use combinations
#Insert your answer here
```

END of Assignment #2.