Quantitative Exercise

Marine Collery and Benjamin Seregi

10/29/2017

Introduction

We are looking at the Monty Hall problem.

Dataset

The dataset is generated by a C program that simulates the multiple cases.

Setting the directory to the data directory.

```
getwd()
```

```
## [1] "/Users/Marine/Documents/1_KTH/4_Courses/3_RMSW/Assignments/Quantitative"
```

```
setwd("/Users/Marine/Documents/1_KTH/4_Courses/3_RMSW/Assignments/Quantitative")

getwd()
```

```
## [1] "/Users/Marine/Documents/1_KTH/4_Courses/3_RMSW/Assignments/Quantitative"
```

Read in data frame

```
mdf <- read.csv("data.csv",header=T,sep=',',stringsAsFactors = F)

# check what variables are there
names(mdf)
```

```
## [1] "door"     "selected" "S"        "NS"
```

```
summary(mdf)
```

```
##      door             selected              S
##  Length:2           Length:2           Length:2
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##       NS
##  Length:2
##  Class :character
##  Mode  :character
```

```
#factorize the data

#mdf$doorf=as.factor(mdf$door)
#names(mdf)[3]='label'
mdf$Sf=as.factor(mdf$S)
mdf$NSf=as.factor(mdf$NS)
summary(mdf)
```

```
##      door             selected              S
##  Length:2           Length:2           Length:2
##  Class :character   Class :character   Class :character
```

```
##  Mode  :character   Mode  :character   Mode  :character
##       NS            Sf    NSf
##  Length:2           f:1   f:1
##  Class :character   w:1   w:1
##  Mode  :character
```

```
#Count all the wins in 2 sets : Switching or Not
countWin <- function(df) {
    # df is a dataframe - want to get first character of "phases" label and insert into label2 column

    l1 <- length(df$NS)
    sumWinS <- 0
    sumWinNS <- 0

    for (i in 1:l1) {

        if (df$S=='w' & df$NS=='f') {
          sumWinS <- sumWinS+1
          df$sumWinSwitched[i] <- sumWinS

          #Not changed:
          df$sumWinNotSwitched[i] <- sumWinNS
        }
        else if (df$S=='f' & df$NS=='w'){
          sumWinNS <- sumWinNS+1
          df$sumWinNotSwitched[i] <- sumWinNS

          #Not changed:
          df$sumWinSwitched[i] <- sumWinS
        }
    }
    return(df)

}

mdf <- countWin(mdf)
```

```
## Warning in if (df$S == "w" & df$NS == "f") {: the condition has length > 1
## and only the first element will be used

## Warning in if (df$S == "w" & df$NS == "f") {: the condition has length > 1
## and only the first element will be used
```

```
summary(mdf)
```

```
##      door              selected              S
##  Length:2           Length:2           Length:2
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##       NS            Sf    NSf   sumWinSwitched sumWinNotSwitched
##  Length:2           f:1   f:1   Min.   :1.00   Min.   :0
##  Class :character   w:1   w:1   1st Qu.:1.25   1st Qu.:0
##  Mode  :character               Median :1.50   Median :0
```

```
##                                 Mean   :1.50   Mean   :0
##                                 3rd Qu.:1.75   3rd Qu.:0
##                                 Max.   :2.00   Max.   :0
```

Variables

Independant Variables

Dependant Variables

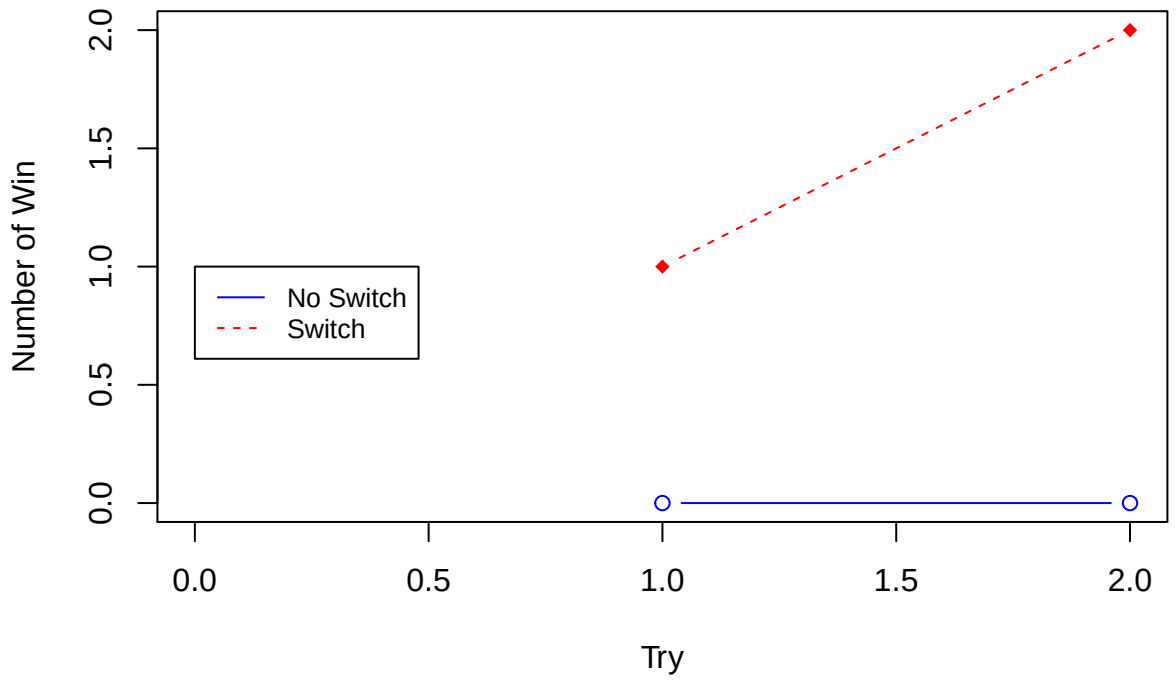
Analysis

```
plot(mdf$sumWinNotSwitched, xlim=c(0.0,length(mdf$NS)), ylim=c(0.0,length(mdf$NS)), xlab='Try', ylab='Nu


#par(new=T)
#plot(mdf$sumWinSwitched, xlim=c(0.0,length(mdf$NS)), ylim=c(0.0,length(mdf$NS)), xlab='', ylab='', col=

lines(mdf$sumWinSwitched, pch=18, col="red", type="b", lty=2)

legend(0, length(mdf$NS)-1, legend=c("No Switch", "Switch"),
       col=c("blue", "red"), lty=1:2, cex=0.8)
```




```
#par(new=F)
```

R Markdown

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.

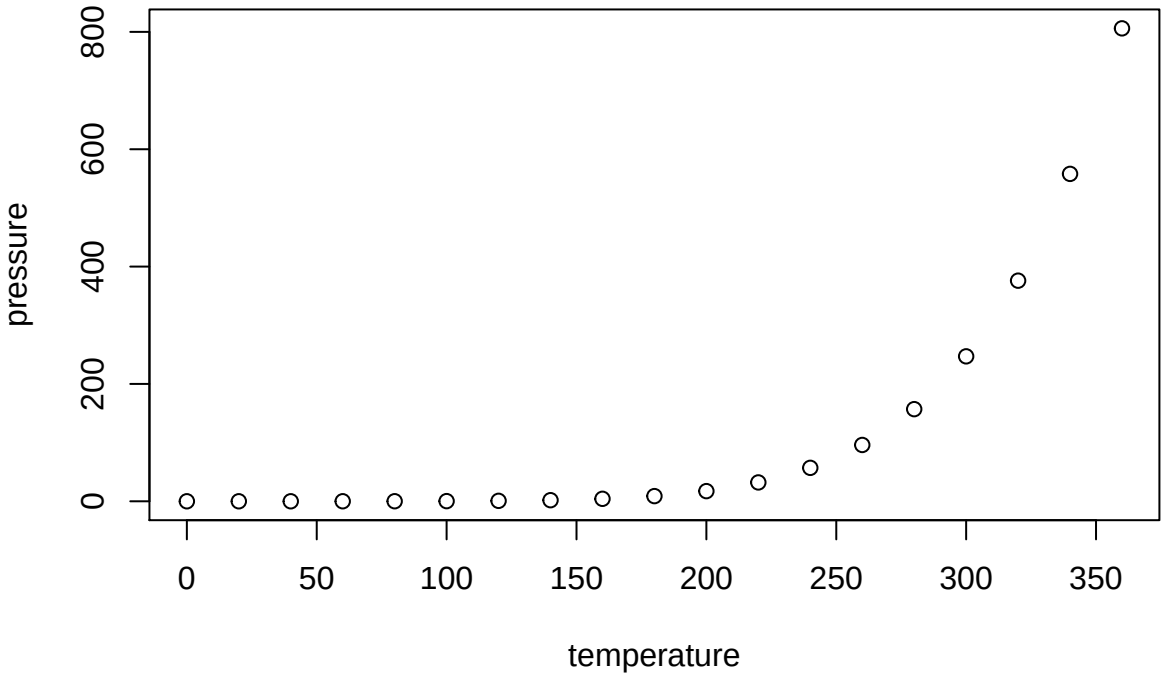
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.