POL211 TA Session 1

Gento Kato

October 11, 2018

Intro to R

Where to Find

- **R**: Go to https://cran.r-project.org/. Select the appropriate version for your OS to download.
- **RStudio**: The user-friendly interface to use R. Go to https://www.rstudio.com/products/RStudio and download the free desktop version.

Set up RStudio

- Tools -> Global Options -> Appearance to change the appearance (e.g., dark background and light color texts).
- Tools -> Global Options -> Pane Layout to change the layout.

Good Custom to Have

- Create the new folder to save codes, datasets, etc... This is the clever way to manage your project.
- Always use the script file to type and save your codes. Make sure to add .R extention. Avoid typing codes directly to console. This custom will help you make your results reproducible.

Basic Rules to Remember.

1. Codes are case sensitive.
2. Shortcut commands to execute the current script line: Ctrl+Return in Windows; Command+Return in Mac.
3. Save script, avoid overwriting the datasets. (Reproducibility!)
4. Leave memo in the script to make codes sense for you (Texts entered after # sign in each row are recognized as memo).
5. Reference help ???les when you can: ? before the command will open help ???le.

```
?rm # This opens help file for the command "rm
```

Set working directory

The working directory is the folder to store and save all files relevant to the analysis.

```
setwd("C:/GoogleDrive/Lectures/2018_10to12_UCD/POL211 TA/POL211_TA_resource")
# If using WINDOWS, replace all backslashes in the file path with slash.
getwd()
```

```
## [1] "C:/GoogleDrive/Lectures/2018_10to12_UCD/POL211 TA/POL211_TA_resource"
```

```
# Easier Way (only in R Studio)
setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
getwd()
```

```
## [1] "C:/GoogleDrive/Lectures/2018_10to12_UCD/POL211 TA/POL211_TA_resource"
```

Clear workspace

```
rm(list = ls()) # Remove all objects from workspace.
```

Basic R functions

Calculation

```
(3+5/78)^3*7 # is 201.3761
```

```
## [1] 201.3761
```

```
sqrt(2^2) ## is 2
```

```
## [1] 2
```

```
exp(log(4)) # is 4
```

```
## [1] 4
```

True or False Statements

```
1 < 2
```

```
## [1] TRUE
```

```
1 == 2
```

```
## [1] FALSE
```

```
1 != 2
```

```
## [1] TRUE
```

```
2 == 4-2
```

```
## [1] TRUE
```

```
exp(log(6)) == 6
```

```
## [1] TRUE
```

```
"A" == "A" # Can be applied to texts too.
```

```
## [1] TRUE
```

```
exp(log(6)) == 6 & 1 == 2 # TRUE if both are TRUE, FALSE if at least one is FALSE
```

```
## [1] FALSE
```

```r
exp(log(6)) == 6 | 1 == 2 # TRUE if either one of the statement is TRUE, FALSE if both are FALSE
```

```
## [1] TRUE
```

Manipulation of Objects

- R recognizes & manages almost anything as object. Object can be numbers, texts, vector (set of numbers/texts), data frame (dataset), ...
- All objects in the current workspace appears in Environment tab in RStudio.
- If you create the new object with the same name the old object, the new object overwrites the old object. (Be careful!)

```r
test <- 3
TEST <- 5
Test <- sqrt(49)
tesT <- 3^2

test # is 3
```

```
## [1] 3
```

```r
TEST # is 5
```

```
## [1] 5
```

```r
Test # is 7
```

```
## [1] 7
```

```r
tesT # is 9
```

```
## [1] 9
```

```r
class(test) # The object class is named as "numeric"
```

```
## [1] "numeric"
```

```r
text <- "UC Davis"; text
```

```
## [1] "UC Davis"
```

```r
class(text) # The object class is "character"
```

```
## [1] "character"
```

```r
vecN <- c(1,2,3,4,5) # The vector with numbers 1,2,3,4,5.
vecT <- c("A","B","C","A","A") # The vector with characters
vecN; vecT
```

```
## [1] 1 2 3 4 5
```

```
## [1] "A" "B" "C" "A" "A"
```

```r
vecT[3] # Show the third element of vecT
```

```
## [1] "C"
```

```r
vecT[c(1,3)] # Show the first and third element of vecT
```

```
## [1] "A" "C"
```

```r
vecN53 <- vecN[c(5,3)] # vecN53 as the fifth and third elements of vecN
vecN53
```

```
## [1] 5 3
```

Replication of Week 1 Code

```r
#Probability Space
samplespace <- c(1,-1) #outcome
p <- c(1/2, 1/2)

#sample without replacement
sample(x = samplespace,
       size = 1, replace = FALSE,
       prob = p)
```

```
## [1] 1
```

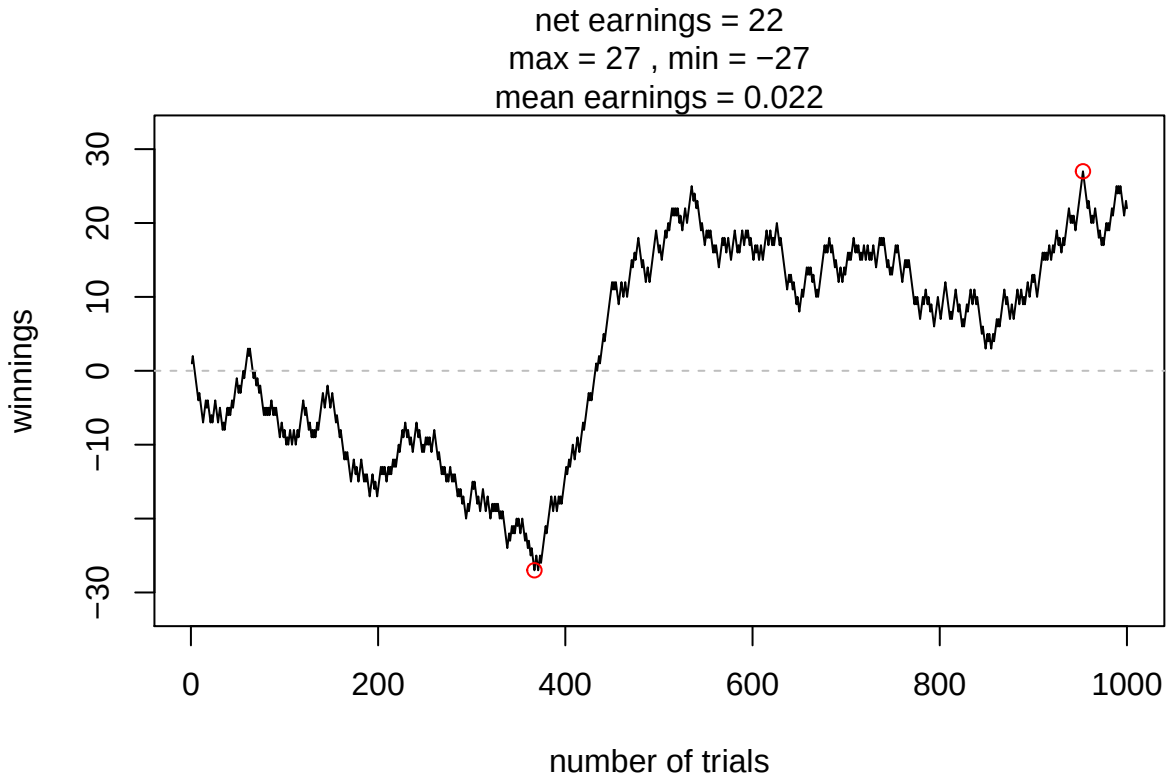
```r
#sample with replacement
set.seed(34567) # For the replicable result
sample(x = samplespace,
       size = 10, replace = TRUE, prob = p)
```

```
##  [1]  1  1 -1 -1 -1 -1 -1 -1  1 -1
```

```r
#law of large numbers
set.seed(34567) # For the replicable result
n <- 1000 #change n to see LLN in action
#trials <- sample(x = samplespace, size = n, replace = TRUE, prob = p); trials
trials <- replicate(n, sample(samplespace, 1, prob = p))
mean(trials)
```

```
## [1] 0.022
```

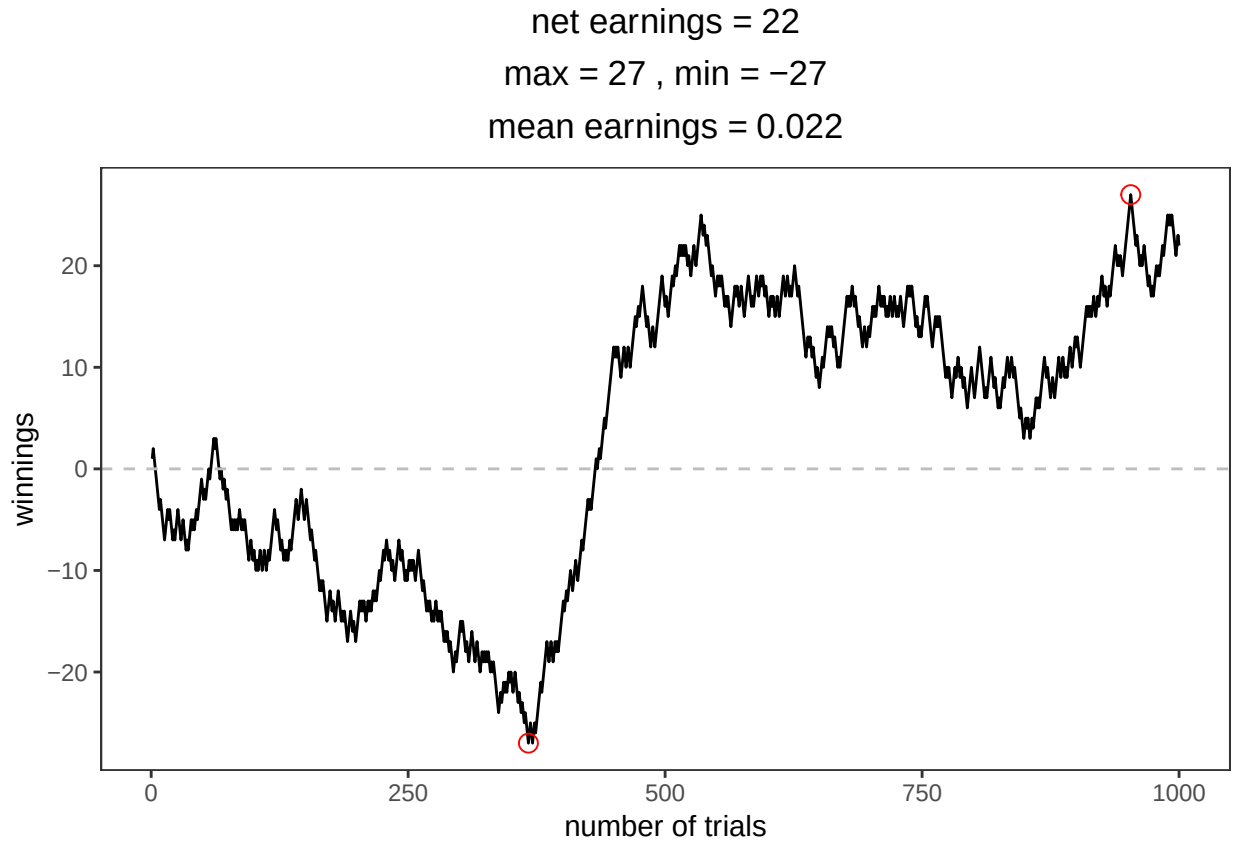
```r
#mean wins says nothing about overall earnings of player 1
plot(x = c(1:n), y = cumsum(trials), type = "l",
     xlab = "number of trials", ylab = "winnings",
     ylim = c(min(cumsum(trials)) - 5, max(cumsum(trials)) + 5))
abline(h = 0, lty = "dashed", col = "grey")
points(which.max(cumsum(trials)), max(cumsum(trials)), col = "red")
points(which.min(cumsum(trials)), min(cumsum(trials)), col = "red")
mtext(paste("mean earnings =", mean(trials), sep = " "), line = 0)
mtext(paste("max =", max(cumsum(trials)), ", min =", min(cumsum(trials)), sep = " "),
      line = 1)
mtext(paste("net earnings =", sum(trials), sep = " "), line = 2)
```




```
## ggplot replication of the above plot
#install.packages("ggplot2")
library(ggplot2)

d <- data.frame(x=c(1:n),y=cumsum(trials)) # Need Dataset Object

ggplot(d,aes(x=x,y=y)) +
  geom_line() +
  geom_hline(aes(yintercept=0),linetype=2, col="grey") +
  annotate("point",colour="red",size=3,shape=1,
           x=which.max(cumsum(trials)),
           y=max(cumsum(trials))) +
  annotate("point",colour="red",size=3,shape=1,
           x=which.min(cumsum(trials)),
           y=min(cumsum(trials))) +
  xlab("number of trials") + ylab("winnings") +
  ggtitle(bquote(atop(textstyle("net earnings =" ~ .(sum(trials))),
              atop(textstyle("max =" ~ .(max(cumsum(trials))) ~
                             ", min =" ~ .(min(cumsum(trials)))),
              atop(scriptscriptstyle(""),
                   textstyle("mean earnings =" ~ .(mean(trials))))))) +
  theme_bw() + theme(panel.grid = element_blank(),
                     plot.title = element_text(hjust=0.5))
```




```
## include
```

Examples of sum, abs, colSums

```
# Summing Up all values
sum(trials)
```

```
## [1] 22
```

```
sum(trials)/n
```

```
## [1] 0.022
```

```
mean(trials) # Similar to mean
```

```
## [1] 0.022
```

```
# Absolute values
abs(trials)
```

```
##    [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##   [35] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##   [69] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [103] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [137] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [171] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [205] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [239] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
##  [273] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [307] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [341] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [375] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [409] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [443] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [477] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [511] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [545] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [579] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [613] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [647] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [681] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [715] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [749] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [783] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [817] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [851] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [885] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [919] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [953] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [987] 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
min(abs(trials)) # is larger than 0
```

```
## [1] 1
```

```
# Column Sums
trials10 <- matrix(trials,nrow=100,ncol=10)
 # Just Make trials as a 100 coin toss for 10 times.
colSums(trials10)
```

```
##  [1] -8 -8  0  2 30  0 -2 -6  4 10
```

```
sum(trials10[,1]) == colSums(trials10)[1] # 1st element is 1st column sum
```

```
## [1] TRUE
```

```
sum(trials10[,10]) == colSums(trials10)[10] # 10th element is 10th column sum
```

```
## [1] TRUE
```