Intro to R Markdown lecture

Professor Richard Levine

1/11/2018

## 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:

### Code chunk example

We will flip a fair coin and compute the number of tosses until a head. We will have R perform this experiment repeatedly (1000 times here) and report the average and standard deviation of the number of tosses until a head over these repeated experiments, as well as minimum and maximum number of tosses.

# Flipping a fair coin: compute number of tosses until a head  
simnum = 1000 # number of experiments  
prob = 0.5 # probability of tossing a head  
heads = 0 # initialize heads vector storage  
  
# for-loop and while-loop version  
numtoss<-0  
for(j in 1:simnum){  
 numtoss[j]=0; head=0  
 while(head==0){  
 u=runif(1)  
 head=(u<prob) # toss coin  
 numtoss[j]=numtoss[j]+1  
 }  
}  
c(mean(numtoss), sd(numtoss), min(numtoss), max(numtoss))

## [1] 2.007000 1.370704 1.000000 10.000000

We may type text to appear in our Word report:

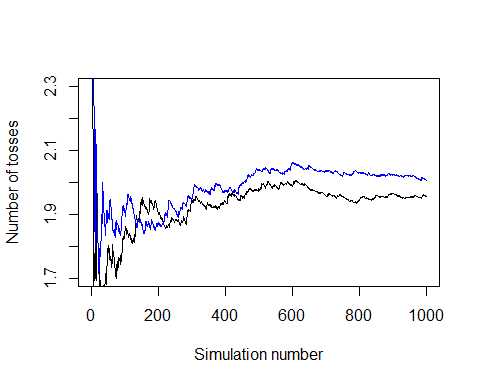
The true values

* mean is 1/p = 2;
* variance is (1-p)/p^2 = 2;
* sd is sqrt(2) = 1.414

As we can see, the simulation of 1000 coin flips arrives close to these true values! Not surprisingly, the minimum value is 1: toss a head on the first coin flip. But we can potentially string together a long sequence of tails, the maximum number of tosses exceeding 10.

### Including Plots

You can also embed plots. For example, let us present a plot the convergence of the average number of tosses from two independent simulation experiments.



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

Again we can type text to interpret the graphic:

The above figure presents a running average of number of tosses for the two experiments (black and blue curves respectively). The figure shows that the simulation experiments vary, especially for the first 200 tosses, due to different sequences in heads and tails tossed. But they both converge to the true average number of tosses until a head of 2 tosses. In fact, the simulation experiments arrive close to this value even after about 300 experiments.