STA 6866 Monte Carlo Statistical Methods
Spring 2014

Homework #1
Assigned Wednesday January 8, 2014
Due Wed January 22, 2014 (this is a lot of time, so you have a chance to get familiar with R)

1 Imagine that you are living in a universe where the value of π is unknown. Run a Monte Carlo experiment to estimate it, as follows. Consider the disk centered at $(0, 0)$, and with radius 1. The square of side 2, also centered at the origin, completely encloses this disk. If you pick a point at random inside the square (i.e. from the uniform distribution over the square), the chance that this point is inside the disk is $\pi/4$. Use this fact to write a program that gives a point estimate, together with a confidence interval, for π.

Solution To estimate π a Monte Carlo experiment is run in which we choose points uniformly at random from inside a square of side 2. If the point lies in the circle of radius 1 inscribed in the square, we record a "success" and if it falls outside, we record a failure. In this experiment, the probability of success is $\pi/4$. So repeating this experiment a large number of times enables us to estimate $\pi/4$ and hence π itself. Following is R code to carry this out.

```
set.seed(1)
ss <- 100000
x <- runif(n=ss, min=-1, max=1)
y <- runif(n=ss, min=-1, max=1)
z <- x^2 + y^2
b <- as.numeric(z < 1)

# Point estimate
hatpi <- 4*sum(b)/ss
hatpi # gives point estimate of 3.13648

# Confidence interval
prob <- hatpi/4
se <- sqrt(prob*(1-prob)/ss)
ci <- 4*c(prob - qnorm(.975)*se, prob + qnorm(.975)*se)
ci # gives 95% confidence interval of (3.12628, 3.14668)

# Alternative: use existing function to calculate the CI
# Look at help file for binom.test
4*binom.test(prob*ss, ss, conf.level=.95)$conf.int
   # gives 95% confidence interval of (3.126227, 3.146667)
```

2 Five runners (called "A" through "E") line up for the start of the 100 meter dash. The distribution of runners' times are all normal, with the same mean of exactly 12 seconds, but the SD's differ. The SD's for runners A, B, C, D, and E are .1, .19, .2, .21, and .3, respectively. Which runner has the greatest chance of winning the race? Intuitively, it's E, but this is difficult to show. Let p_A, p_B, p_C, p_D, and p_E be the probabilities that runners A, B, C, D, and E win the race, respectively. Carry out a Monte Carlo experiment that will gives estimates of these five probabilities.

Solution Let A, B, C, D, E be independent random variables, all normally distributed with mean 12; the standard deviations are .1, .19, .2, .21, and .3, respectively. Following is R code to estimate the probability that the min of the five variables is A; that it is B; $\ldots$; that it is E. Note that we use the function `pmin` (parallel min), which takes as arguments several vectors, and returns the *componentwise* minima. Note: if `xa` and `xb` are two vectors, `min(xa, xb)` will first concatenate `xa` and `xb` and then take the min of that. The reason for using the `pmin` function is that we want to avoid looping whenever possible, both for speed of execution and for memory use. The code as written below runs in much less than 1 second on a "current machine."

```
set.seed(1)
ss <- 100000

xa <- rnorm(n=ss, mean=12, sd=.10)
xb <- rnorm(n=ss, mean=12, sd=.19)
xc <- rnorm(n=ss, mean=12, sd=.20)
xd <- rnorm(n=ss, mean=12, sd=.21)
xe <- rnorm(n=ss, mean=12, sd=.30)

xmin <- pmin(xa,xb,xc,xd,xe)

p.a.wins <- sum(xmin==xa)/ss ; p.a.wins # gives 0.10479
p.b.wins <- sum(xmin==xb)/ss ; p.b.wins # gives 0.19277
p.c.wins <- sum(xmin==xc)/ss ; p.c.wins # gives 0.20224
p.d.wins <- sum(xmin==xd)/ss ; p.d.wins # gives 0.21146
p.e.wins <- sum(xmin==xe)/ss ; p.e.wins # gives 0.28874

# Also get confidence intervals
binom.test(p.a.wins*ss, ss, conf.level=.95)$conf.int
binom.test(p.b.wins*ss, ss, conf.level=.95)$conf.int
binom.test(p.c.wins*ss, ss, conf.level=.95)$conf.int
binom.test(p.d.wins*ss, ss, conf.level=.95)$conf.int
binom.test(p.e.wins*ss, ss, conf.level=.95)$conf.int
```