Math 1180:Lab2

Kyle Gaffney

January 15th, 2014

1   Conditional Probability

Conditional probability is the probability that something happens (or is true) given that something else has happened or given some condition. Our goal with R is to be able to simulate large numbers of cases and also to create conditional simulations.

Suppose that we start off by simulating a population of 10000 with a disease. The probability of a person having the disease is p=0.05, then we can do this with the following script in R:

```
n<-10000
rn<-runif(n)
M<-0
for (i in 1:n){ifelse(rn[i]<0.05,M[i]<-1,M[i]<-0)}
ND<-sum(M)
```

Then out of the 10000 people that we generated ND will be the number that have this disease. Suppose that people that of the people that don't have the disease 65 percent are susceptible to it. Then how many people are susceptible?

```
rn<-runif(n-ND)
M<-0
for (i in 1:(n-ND)){ifelse(rn[i]<0.65,M[i]<-1,M[i]<-0)}
NS<-sum(M)
```

Then NS is the number of people that are susceptible conditional on them not being sick. How does NS compare to the actual statistics for this problem? Will NS get better or worse as we increase the population size?

2 Questions

1. Redo the sample problem in R for a population of 1 million people where additionally of the 20 percent of people who neither have the disease nor are susceptible are instead immune to it. What is the number of people that have the disease, are susceptible to it, are immune, or are normal? How do these numbers compare to the solution you get when you solve the problem by hand? Be sure to submit your code as well as your work by hand along with your explanations.

2. If you were to do a large number of simulations how would the mean number of each population compare to the probability solution determined by hand? Which population would have the largest standard deviation (relative/proportional to the mean)? Why?