Project1

Jongwoo Choi, Devon Rhodes

March 27, 2017

Abstract: Suppose one wanted to know how many trees were on a plantation. It would be nonsensical to count all trees on a large plantation. One could take plot samples to deduce the total number of trees on an entire plantation. We examine how varying both the number of plots and the plot sizes affects the estimation of total trees on a plantation.

Introduction: In this study we look at how accurately different plot sizes estimate the number of trees on a plantation. We have two plantations both of size 1000 by 1000 meters. We seek to know if plot size matters depending whether the entire plantation is densely or sparsely populated. It would obviously take longer and cost more to sample larger plots. In order to be realistic, when the plot size is large, we take fewer samples and vice versa. We consider 500 five by five meter plots, 250 ten by ten meter plots, 125 twenty by twenty meter plots, and 62 forty by forty meter plots.

```
> ###Create trees placed randomly in forest and get coordinates for them.
> ###Assume a 1000 by 1000 unit plot
> num.trees=5000
> x.tree.cord=runif(num.trees,0,1000)
> y.tree.cord=runif(num.trees,0,1000)
> qplot(x.tree.cord,y.tree.cord,main ="1000x1000 meter plantation with 5000 trees" )
>

> ##Function to randomly select a plot and calculate number
> ##of trees in plot
> get.count.fun=function(num.trees,plot.size,x.tree.cord,y.tree.cord){
+   x.cord=round(runif(1,1,1000/plot.size),0)*plot.size
+   y.cord=round(runif(1,1,1000/plot.size),0)*plot.size
+   trees.in.sample=sum(as.numeric(x.cord-plot.size<x.tree.cord &x.tree.cord<x.cord&y.cord-plot.size<y.
+   trees.in.sample}

> ### Function to repeat above function n times (the sample size which you
> ### should vary)
> get.sample=function(n,num.trees,plot.size,x.tree.cord,y.tree.cord){
+   count.per.plot=rep(0,n)
+   for (i in 1:n)
+     count.per.plot[i]=get.count.fun(num.trees,plot.size,x.tree.cord,y.tree.cord)
+   count.per.plot
+ }
>  d0=do(100)*mean(get.sample(100,num.trees,5,x.tree.cord,y.tree.cord))
>  d1=do(100)*mean(get.sample(50,num.trees,10,x.tree.cord,y.tree.cord))
>  d2=do(100)*mean(get.sample(25,num.trees,20,x.tree.cord,y.tree.cord))
>  d3=do(100)*mean(get.sample(10,num.trees,40,x.tree.cord,y.tree.cord))
> m0=mean(d0[,1])*40000
> m1=mean(d1[,1])*10000
> m2=mean(d2[,1])*2500
> m3=mean(d3[,1])*625
```

```
> num.trees.s=1000
> x.tree.cord=runif(num.trees.s,0,1000)
> y.tree.cord=runif(num.trees.s,0,1000)
> qplot(x.tree.cord,y.tree.cord,main ="1000x1000 meter plantation with 1000 trees")
>  d4=do(100)*mean(get.sample(100,num.trees.s,5,x.tree.cord,y.tree.cord))
>  d5=do(100)*mean(get.sample(50,num.trees.s,10,x.tree.cord,y.tree.cord))
>  d6=do(100)*mean(get.sample(25,num.trees.s,20,x.tree.cord,y.tree.cord))
>  d7=do(100)*mean(get.sample(10,num.trees.s,40,x.tree.cord,y.tree.cord))
> m4=mean(d4[,1])*40000
> m5=mean(d5[,1])*10000
> m6=mean(d6[,1])*2500
> m7=mean(d7[,1])*625
```

Methods: We begin by creating a 1000 meter by 1000 meter plantation and uniformly distributing trees on that plantation. The trees are represented as points on a plane. The plantations have 5000 and 1000 trees. A function is used to randomly select a five by five meter plot and count the number of trees found in that plot. We sample 500 plots when using five by five meter plots. We simulate this process 10,000 times and calculate the mean total trees in each plot. Doing this, we can estimate how many trees are in the entire plantation. The same technique is used with the ten by ten meter plots, the twenty by twenty plots, and the forty by forty plots.

```
> r1=c(m0,m1,m2,m3)
> r2=c(m4,m5,m6,m7)

> tm1=matrix(c(r1),nrow=1,byrow=T)
> tree.count=c("5x5 plot with sample size 500", "10x10 plot with sample size 250" ,"20x20 plot with sam
> colnames(tm1)=tree.count
> tm1

     5x5 plot with sample size 500 10x10 plot with sample size 250
[1,]                          5056                            4908
     20x20 plot with sample size 125 40x40 plot with sample size 62.5
[1,]                            5057                          5036.25

> tm2=matrix(c(r2),nrow=1,byrow=T)
> tree.count=c("5x5 plot with sample size 500", "10x10 plot with sample size 250" ,"20x20 plot with sam
> colnames(tm2)=tree.count
> tm2

     5x5 plot with sample size 500 10x10 plot with sample size 250
[1,]                          1068                            1048
     20x20 plot with sample size 125 40x40 plot with sample size 62.5
[1,]                             975                          1011.25
```

Conclusion: From our simulation study, we conclude that 250 samples of ten by ten meter plots estimates the number of trees on a plantation most accurately of the scenarios we tried. From this we can conclude that estimating the number of trees on a plantation is done more accurately using many small plots rather than less large plots.