Random Sampling and Sampling Distributions

2304-BADGUJAR PARAG KISHOR

2023-08-08

# Random Sampling and Sampling distributions

## Problem 1- Generate random sample from the finite population using (i) SRSWR (ii) SRSWOR (iii) Unequal probabilities for population elements.

U=1:5  
Pu=c(0.1,0.4,0.1,0.2,0.2)  
n=4  
#sample(x, size, replace = FALSE, prob = NULL)  
x1=sample(U,n,replace=T)  
x2=sample(U,n,replace=F)  
x3=sample(U,n,replace=T,prob=Pu)  
cat("Original Population:",U)

## Original Population: 1 2 3 4 5

cat("\n Random Sample by using SRSWR",x1)

##   
## Random Sample by using SRSWR 4 3 5 3

cat("\n Random Sample by using SRSWOR",x2)

##   
## Random Sample by using SRSWOR 2 5 1 4

cat("\n Random Sample by using unequal probabilities",x3)

##   
## Random Sample by using unequal probabilities 2 2 3 2

## Problem2 - Generate random sample of size 10, 000 from . Treat this generated sample data as characteristic values of study variables of the population of the size 10, 000. Now, draw 5 random samples of size 100 each from this population using SRSWOR, obtain estimate of population mean based on each sample data by using sample mean as an estimator for population mean. Also find the amount of sampling error from your estimates.

N=10000  
lambda=3  
fam=1:N  
fam\_size=rpois(N,lambda)  
s=matrix(rep(0,500),nrow=5,ncol=100)  
for(i in 1:5){  
 s[i,]=fam\_size[sample(fam,100,replace=F)]  
}  
  
pop\_mean=mean(fam\_size)  
sam\_mean=rowMeans(s)  
cat("\n Population mean",pop\_mean)

##   
## Population mean 2.9893

cat("\n sample mean",sam\_mean)

##   
## sample mean 3.09 3 2.96 2.84 3.14

sm=c("s1","s2","s3","s4","s5")  
data.frame(sm,sam\_mean)

## sm sam\_mean  
## 1 s1 3.09  
## 2 s2 3.00  
## 3 s3 2.96  
## 4 s4 2.84  
## 5 s5 3.14

## Problem3- Draw 10, 000 samples of size 25 from Normal and study the sampling distribution of following statistics for = 5, 10, 25.

* \

Y=matrix(rnorm(10000\*25,100,3),nrow=10000,ncol=25)  
M1=matrix(rep(0,5\*6),nrow=5)  
colnames(M1)=c("Mean\_n5","Mean\_n10","Mean\_n25","MSE\_n5","MSE\_n10","MSE\_n25")  
row.names(M1)=c( "T1","T2","T3","T4","T5")  
  
n1=c(5,10,25)  
  
for(i in 1:3)  
{ #First Estimator  
 T1=rowMeans(Y[,1:n1[i]])  
 M1[1,i]=mean(T1)  
 M1[1,i+3]=mean((T1-100)^2)  
 #Second Estimator  
 T2=apply(Y[,1:n1[i]],1,var)  
 M1[2,i]=mean(T2)  
 M1[2,i+3]=mean((T2-9)^2)  
 #Third Estimator  
 T3=apply(Y[,1:n1[i]],1,min)  
 M1[3,i]=mean(T3)  
 M1[3,i+3]=mean((T3-100)^2)  
 #Fourth Parameter  
 T4=apply(Y[,1:n1[i]],1,max)  
 M1[4,i]=mean(T4)  
 M1[4,i+3]=mean((T4-100)^2)  
 #Fifth Parameter  
 T5=T4-T3  
 M1[5,i]=mean(T5)  
 M1[5,i+3]=mean((T5-100)^2)  
   
}  
M1

## Mean\_n5 Mean\_n10 Mean\_n25 MSE\_n5 MSE\_n10 MSE\_n25  
## T1 100.004401 100.015827 100.008455 1.826453 0.9150704 0.3636527  
## T2 9.020703 9.035015 9.044867 40.485269 17.7984555 6.7364724  
## T3 96.516626 95.387049 94.105984 16.188704 24.4244267 37.1024373  
## T4 103.503401 104.636062 105.912729 16.320250 24.5726208 37.3032541  
## T5 6.986775 9.249013 11.806744 8658.120894 8241.4230571 7782.5891672

## Problem 4 Consider the population of size with values as (138, 142, 145, 155, 143). Write down all possible samples of size = 3 using SRSWOR for each sample obtain values of

Now verify - i.e.  is unbiased estimator of (population mean). - i.e.  is biased estimator of - i.e.  is biased estimator of

Compare and in terms of their M.S.E.

Y=c(138,142,145,155,143)  
#All possible samples  
Mat=t(combn(Y,3))  
T1=rowMeans(Mat)  
T2=(Mat[,1]+Mat[,3])/2  
T3=apply(Mat,1,median)  
popmn=mean(Y)  
estimates=c(mean(T1),mean(T2),mean(T3))  
mses=c(mean(T1-popmn)^2,mean((T2-popmn)^2),mean((T3-popmn)^2))  
M5=data.frame(estimates,mses)  
M5

## estimates mses  
## 1 144.60 3.231273e-29  
## 2 143.35 9.165000e+00  
## 3 143.30 3.100000e+00

u=c(1,2,3,4)  
range(u)

## [1] 1 4