Topic 1 : Parametric Inference

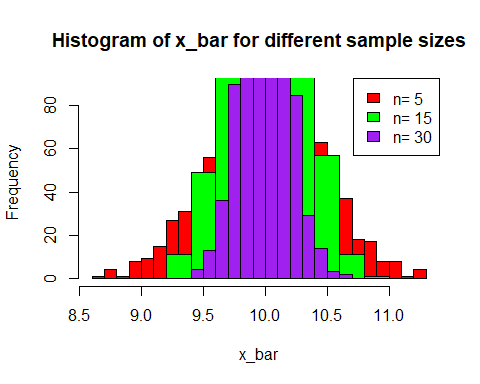
Shubhangi Mistari

2024-12-31

#Topic 1

1.Example  
To conduct the simulation study from N(=1,=1) distribution. Demonstrate the distribution and it’s descriptive statistics for different sample sizes of following statistic/estimators. ##1.sample mean

n=c(5,15,30)  
sam\_mean=sam\_var=numeric(length(n))  
x\_bar\_list=list()  
for (i in 1:length(n)){  
 x=matrix(rnorm(1000\*n[i],10,1),nrow = 1000,ncol = n[i])  
 x\_bar=apply(x,1,mean)  
 sam\_mean[i]=mean(x\_bar)  
 sam\_var[i]=var(x\_bar)  
 x\_bar\_list[[i]]=x\_bar  
   
}  
  
hist(x\_bar\_list[[1]],breaks="FD",col="red",main="Histogram of x\_bar for different sample sizes",xlab="x\_bar")  
hist(x\_bar\_list[[2]],col="green",add=T)  
hist(x\_bar\_list[[3]],col="purple",add=T)  
legend("topright",legend=paste("n=",n),  
 fill=c("red","green","purple"))



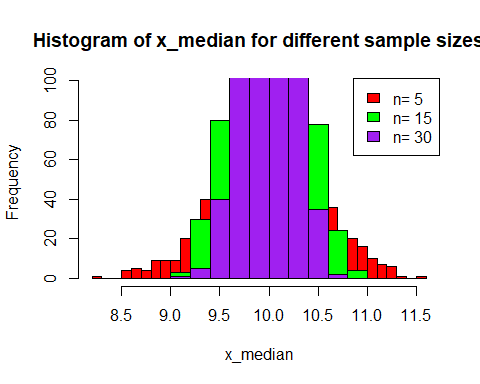
cbind(n,sam\_mean,sam\_var)

## n sam\_mean sam\_var  
## [1,] 5 10.010357 0.20344370  
## [2,] 15 10.006069 0.06906252  
## [3,] 30 9.995775 0.03472507

From this output, we conclude that as sample size increases,variability in the sample mean is decreases. Also, sample mean tends to population mean as n increases.

##2.Sample median

n=c(5,15,30)  
sam\_median\_mean=sam\_median\_var=numeric(length(n))  
x\_median\_list=list()  
for (i in 1:length(n)){  
 x=matrix(rnorm(1000\*n[i],10,1),nrow = 1000,ncol = n[i])  
 x\_median=apply(x,1,median)  
 sam\_median\_mean[i]=mean(x\_median)  
 sam\_median\_var[i]=var(x\_median)  
 x\_median\_list[[i]]=x\_median  
}  
hist(x\_median\_list[[1]],breaks="FD",col="red",main="Histogram of x\_median for different sample sizes",xlab="x\_median")  
hist(x\_median\_list[[2]],col="green",add=T)  
hist(x\_median\_list[[3]],col="purple",add=T)  
legend("topright",legend=paste("n=",n),  
 fill=c("red","green","purple"))



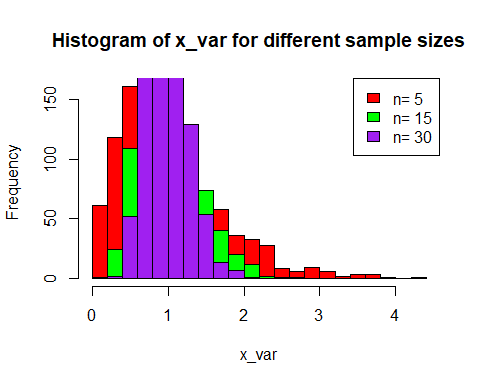
cbind(n,sam\_median\_mean,sam\_median\_var)

## n sam\_median\_mean sam\_median\_var  
## [1,] 5 9.989167 0.26284511  
## [2,] 15 9.987423 0.09937219  
## [3,] 30 9.991454 0.05201575

From this output, we conclude that as sample size increases,variability in the sample median is decreases.

##3.Sample variance

n=c(5,15,30)  
sam\_var\_mean=sam\_var\_var=numeric(length(n))  
x\_var\_list=list()  
for (i in 1:length(n)){  
 x=matrix(rnorm(1000\*n[i],10,1),nrow = 1000,ncol = n[i])  
 x\_var=apply(x,1,var)  
 sam\_var\_mean[i]=mean(x\_var)  
 sam\_var\_var[i]=var(x\_var)  
 x\_var\_list[[i]]=x\_var  
   
}  
  
hist(x\_var\_list[[1]],breaks="FD",col="red",main="Histogram of x\_var for different sample sizes",xlab="x\_var")  
hist(x\_var\_list[[2]],col="green",add=T)  
hist(x\_var\_list[[3]],col="purple",add=T)  
legend("topright",legend=paste("n=",n),  
 fill=c("red","green","purple"))



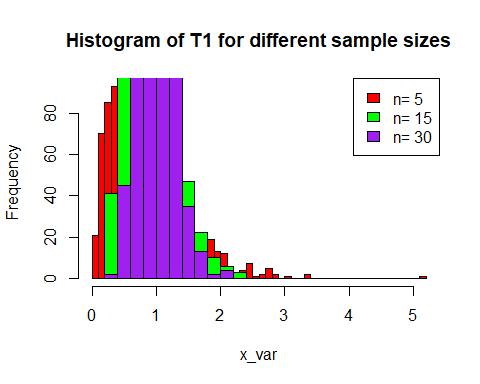
cbind(n,sam\_var\_mean,sam\_var\_var)

## n sam\_var\_mean sam\_var\_var  
## [1,] 5 1.0135510 0.49156537  
## [2,] 15 0.9917855 0.14672027  
## [3,] 30 0.9968313 0.07143448

From this output, we conclude that as sample size increases,variability in the sample variance is decreases.

##population variance

n=c(5,15,30)  
sam\_var\_mean=sam\_var\_var=numeric(length(n))  
T1\_list=list()  
for (i in 1:length(n)){  
 x=matrix(rnorm(1000\*n[i],10,1),nrow = 1000,ncol = n[i])  
 x\_var=apply(x,1,var)  
 T1=((n[i]-1)/n[i])\*x\_var  
 sam\_var\_mean[i]=mean(T1)  
 sam\_var\_var[i]=var(T1)  
 T1\_list[[i]]=T1  
   
}  
  
hist(T1\_list[[1]],breaks="FD",col="red",main="Histogram of T1 for different sample sizes",xlab="x\_var")  
hist(T1\_list[[2]],col="green",add=T)  
hist(T1\_list[[3]],col="purple",add=T)  
legend("topright",legend=paste("n=",n),  
 fill=c("red","green","purple"))



cbind(n,sam\_var\_mean,sam\_var\_var)

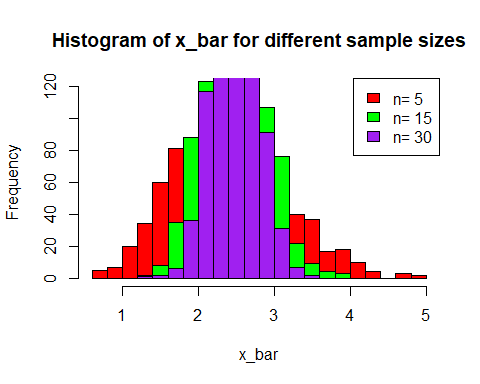
## n sam\_var\_mean sam\_var\_var  
## [1,] 5 0.8098024 0.34121988  
## [2,] 15 0.9251474 0.12401232  
## [3,] 30 0.9760031 0.06379236

2.Example

To conduct the simulation study from Poisson(=2.5) distribution. Demonstrate the distribution and it’s descriptive statistics for different sample sizes of following statistic/estimators.

## sample mean & sample variance

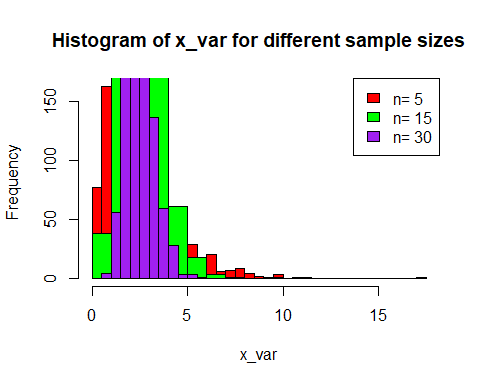
n=c(5,15,30)  
sam\_mean=sam\_var=numeric(length(n))  
sam\_var\_mean=sam\_var\_var=numeric(length(n))  
x\_bar\_list1=list()  
x\_var\_list2=list()  
for (i in 1:length(n)){  
 x=matrix(rpois(1000\*n[i],2.5),nrow = 1000,ncol = n[i])  
   
 x\_bar=apply(x,1,mean)  
 x\_var=apply(x,1,var)  
   
 sam\_mean[i]=mean(x\_bar)  
 sam\_var\_mean[i]=mean(x\_var)  
   
   
 sam\_var[i]=var(x\_bar)  
 sam\_var\_var[i]=var(x\_var)  
   
 x\_bar\_list[[i]]=x\_bar  
 x\_var\_list[[i]]=x\_var  
   
}  
  
hist(x\_bar\_list[[1]],breaks="FD",col="red",main="Histogram of x\_bar for different sample sizes",xlab="x\_bar")  
hist(x\_bar\_list[[2]],col="green",add=T)  
hist(x\_bar\_list[[3]],col="purple",add=T)  
legend("topright",legend=paste("n=",n),  
 fill=c("red","green","purple"))



cbind(n,sam\_mean,sam\_var)

## n sam\_mean sam\_var  
## [1,] 5 2.504400 0.5036042  
## [2,] 15 2.512733 0.1719654  
## [3,] 30 2.499433 0.0837868

hist(x\_var\_list[[1]],breaks="FD",col="red",main="Histogram of x\_var for different sample sizes",xlab="x\_var")  
hist(x\_var\_list[[2]],col="green",add=T)  
hist(x\_var\_list[[3]],col="purple",add=T)  
legend("topright",legend=paste("n=",n),  
 fill=c("red","green","purple"))



cbind(n,sam\_var\_mean,sam\_var\_var)

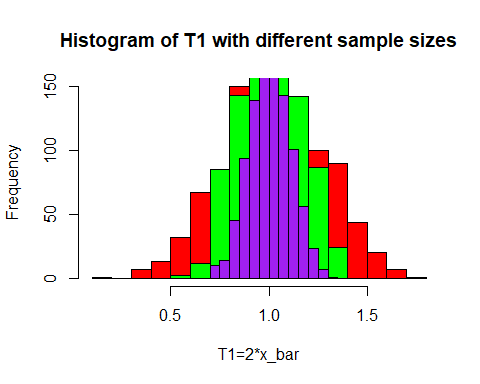
## n sam\_var\_mean sam\_var\_var  
## [1,] 5 2.486600 3.5033638  
## [2,] 15 2.515333 1.0751809  
## [3,] 30 2.515066 0.5280931

From this output, we conclude that as sample size increases,variability in the sample mean and sample variance is decreases. Also, sample mean tends to population mean as n increases and sample variance tends to population variance as n increases.

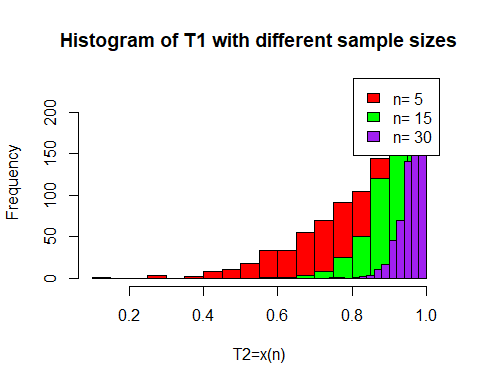
1. Example To conduct the simulation study from U(0,),>0 distribution. Demonstrate the distribution and it’s descriptive statistics for different sample sizes of following statistic/estimators.

T1=2\*() & T2=

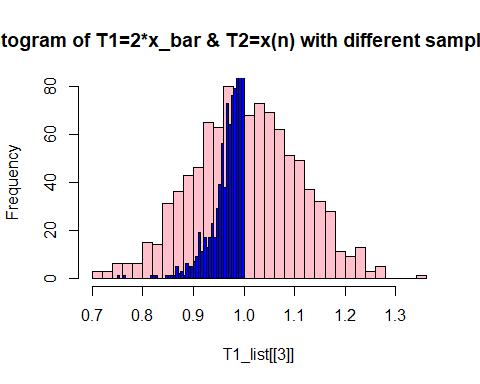
n=c(5,15,30)  
sam\_mean\_T1=sam\_var\_T1=numeric(length(n))  
sam\_mean\_T2=sam\_var\_T2=numeric(length(n))  
  
T1\_list=list()  
T2\_list=list()  
  
for (i in 1:length(n)){  
 x=matrix(runif(1000\*n[i],0,1),nrow = 1000,ncol = n[i])  
   
 x\_bar=apply(x,1,mean)  
 T1=2\*x\_bar  
 T2=apply(x,1,max)  
   
 sam\_mean\_T1[i]=mean(T1)   
 sam\_mean\_T2[i]=mean(T2)  
   
 sam\_var\_T1[i]=var(T1)  
 sam\_var\_T2[i]=var(T2)  
   
 T1\_list[[i]]=T1  
 T2\_list[[i]]=T2  
   
}  
  
hist(T1\_list[[1]],breaks="FD",col="red",main="Histogram of T1 with different sample sizes",xlab="T1=2\*x\_bar")  
hist(T1\_list[[2]],col="green",add=T)  
hist(T1\_list[[3]],col="purple",add=T)



hist(T2\_list[[1]],breaks="FD",col="red",main="Histogram of T1 with different sample sizes",xlab="T2=x(n)")  
hist(T2\_list[[2]],col="green",add=T)  
hist(T2\_list[[3]],col="purple",add=T)  
  
legend("topright",legend=paste("n=",n),  
 fill=c("red","green","purple"))



hist(T1\_list[[3]],breaks=35,col="pink",main="Histogram of T1=2\*x\_bar & T2=x(n) with different sample sizes")  
hist(T2\_list[[3]],col="blue",add=T,breaks = 50)



cbind(n,sam\_mean\_T1,sam\_var\_T1)

## n sam\_mean\_T1 sam\_var\_T1  
## [1,] 5 1.015083 0.06893629  
## [2,] 15 1.003221 0.02200900  
## [3,] 30 1.004161 0.01097508

cbind(n,sam\_mean\_T2,sam\_var\_T2)

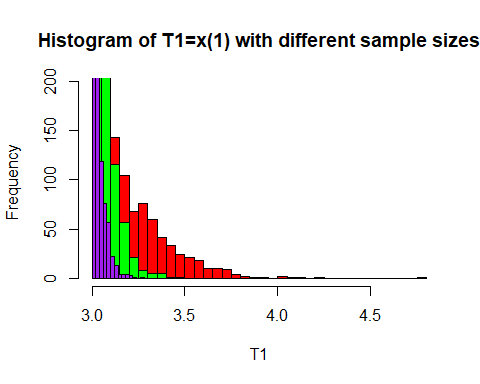
## n sam\_mean\_T2 sam\_var\_T2  
## [1,] 5 0.8376709 0.019157080  
## [2,] 15 0.9370044 0.003627741  
## [3,] 30 0.9683502 0.000952815

From this output, we conclude that as sample size increases variability in T1 and T2 decreases.

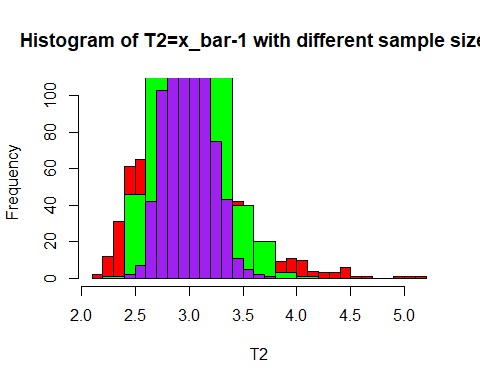
1. Example To conduct the simulation study from Exponential distribution with location parameter . Demonstrate the distribution and it’s descriptive statistics for different sample sizes of following statistic/estimators.

T1= & T2=

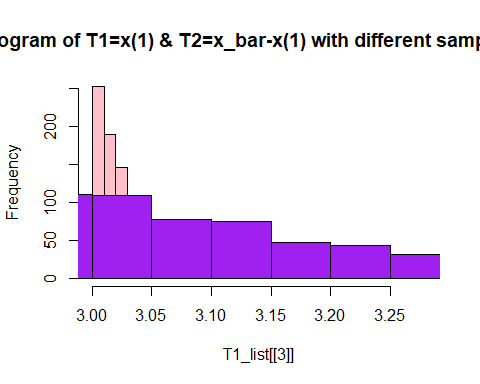
n=c(5,15,30)  
mu=3  
sam\_mean\_T1=sam\_var\_T1=numeric(length(n))  
sam\_mean\_T2=sam\_var\_T2=numeric(length(n))  
  
T1\_list=list()  
T2\_list=list()  
  
for (i in 1:length(n)){  
 x=matrix((rexp(1000\*n[i],1)+mu),nrow = 1000,ncol = n[i])  
   
 T1=apply(x,1,min)  
 T2=(apply(x,1,mean)-1)  
   
 sam\_mean\_T1[i]=mean(T1)  
 sam\_mean\_T2[i]=mean(T2)  
   
 sam\_var\_T1[i]=var(T1)  
 sam\_var\_T2[i]=var(T2)  
   
 T1\_list[[i]]=T1  
 T2\_list[[i]]=T2  
}  
  
hist(T1\_list[[1]],breaks="FD",col="red",main="Histogram of T1=x(1) with different sample sizes",xlab="T1")  
hist(T1\_list[[2]],col="green",add=T)  
hist(T1\_list[[3]],col="purple",add=T)



hist(T2\_list[[1]],breaks="FD",col="red",main="Histogram of T2=x\_bar-1 with different sample sizes",xlab="T2")  
hist(T2\_list[[2]],col="green",add=T)  
hist(T2\_list[[3]],col="purple",add=T)



hist(T1\_list[[3]],breaks=35,col="pink",main="Histogram of T1=x(1) & T2=x\_bar-x(1) with different sample sizes")  
hist(T2\_list[[3]],col="purple",add=T,breaks = 40)



cbind(n,sam\_mean\_T1,sam\_var\_T1)

## n sam\_mean\_T1 sam\_var\_T1  
## [1,] 5 3.204424 0.036883214  
## [2,] 15 3.064248 0.004168304  
## [3,] 30 3.034695 0.001258468

cbind(n,sam\_mean\_T2,sam\_var\_T2)

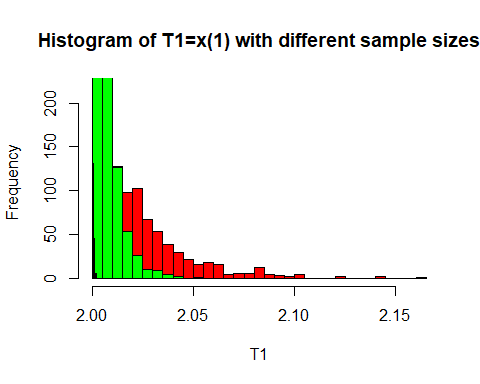
## n sam\_mean\_T2 sam\_var\_T2  
## [1,] 5 3.004824 0.19410768  
## [2,] 15 2.992139 0.06774656  
## [3,] 30 2.990118 0.03540754

From this output, we conclude that as sample size increases variability in T1 and T2 decreases.

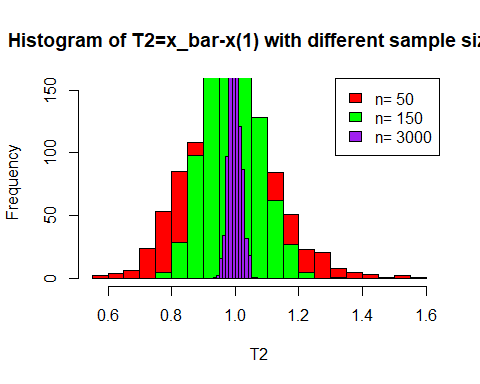
5.Example To conduct the simulation study from Exponential distribution with location parameter and scale parameter . Demonstrate the distribution and it’s descriptive statistics for different sample sizes of following statistic/estimators.

T1= & T2=-

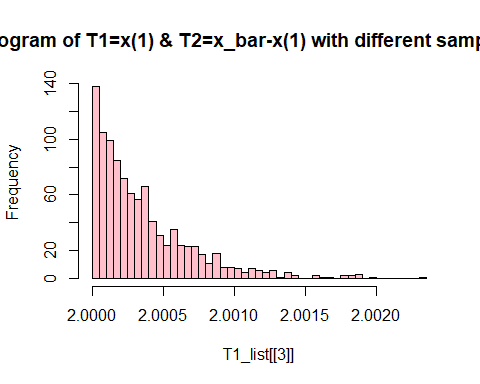
n=c(50,150,3000)  
sam\_mean\_T1=sam\_var\_T1=numeric(length(n))  
sam\_mean\_T2=sam\_var\_T2=numeric(length(n))  
mu=2  
sigma=1  
  
T1\_list=list()  
T2\_list=list()  
  
for (i in 1:length(n)){  
 y=matrix(runif(1000\*n[i],0,1),nrow = 1000,ncol = n[i])  
 x=mu-(sigma\*log(1-y))  
   
 T1=apply(x,1,min)  
 x\_bar=apply(x,1,mean)  
 T2=x\_bar-T1  
   
 sam\_mean\_T1[i]=mean(T1)   
 sam\_mean\_T2[i]=mean(T2)  
   
 sam\_var\_T1[i]=var(T1)  
 sam\_var\_T2[i]=var(T2)  
   
 T1\_list[[i]]=T1  
 T2\_list[[i]]=T2  
   
}  
  
hist(T1\_list[[1]],breaks="FD",col="red",main="Histogram of T1=x(1) with different sample sizes",xlab="T1")  
hist(T1\_list[[2]],col="green",add=T)  
hist(T1\_list[[3]],col="purple",add=T)



hist(T2\_list[[1]],breaks="FD",col="red",main="Histogram of T2=x\_bar-x(1) with different sample sizes",xlab="T2")  
hist(T2\_list[[2]],col="green",add=T)  
hist(T2\_list[[3]],col="purple",add=T)  
  
legend("topright",legend=paste("n=",n),  
 fill=c("red","green","purple"))



hist(T1\_list[[3]],breaks=35,col="pink",main="Histogram of T1=x(1) & T2=x\_bar-x(1) with different sample sizes")  
hist(T2\_list[[3]],col="purple",add=T,breaks = 40)



cbind(n,sam\_mean\_T1,sam\_var\_T1)

## n sam\_mean\_T1 sam\_var\_T1  
## [1,] 50 2.021499 4.642287e-04  
## [2,] 150 2.006874 4.686284e-05  
## [3,] 3000 2.000351 1.158359e-07

cbind(n,sam\_mean\_T2,sam\_var\_T2)

## n sam\_mean\_T2 sam\_var\_T2  
## [1,] 50 0.9823792 0.0201109863  
## [2,] 150 0.9912014 0.0064092467  
## [3,] 3000 0.9988754 0.0003488897

From this output, we conclude that as sample size increases variability in T1 and T2 decreases.