LAB7-RMD

Clayton Glenn

March 1, 2018

Table of Contents

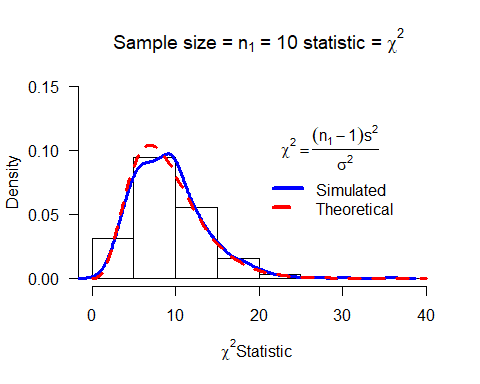
# Get working Directory

getwd()

## [1] "C:/Users/cglen/Documents/Stat Methods/Labs/LAB7"

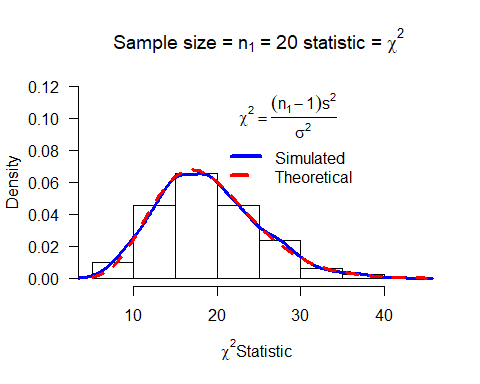
# Chi-Square Distribution for 1 population

mychisim<-function(n1=10,sigma1=3,mean1=5,iter=1000,ymax=0.1,x=20, y=0.1){ # adjust ymax to make graph fit  
   
 y1=rnorm(n1\*iter,mean=mean1,sd=sigma1)# generate iter samples of size n1  
   
 data1.mat=matrix(y1,nrow=n1,ncol=iter,byrow=TRUE) # Each column is a sample size n1  
   
 ssq1=apply(data1.mat,2,var) # ssq1 is s squared  
   
 w=(n1-1)\*ssq1/sigma1^2 #chi-sq stat  
   
 hist(w,freq=FALSE, ylim=c(0,ymax), # Histogram with annotation  
 main=substitute(paste("Sample size = ",n[1]," = ",n1," statistic = ",chi^2)),  
 xlab=expression(paste(chi^2, "Statistic",sep=" ")), las=1)  
 lines(density(w),col="Blue",lwd=3) # add a density plot  
 curve(dchisq(x,n1-1),add=TRUE,col="Red",lty=2,lwd=3) # add a theoretical curve  
 title=expression(chi^2==frac((n[1]-1)\*s^2,sigma^2)) #mathematical annotation -see ?plotmath  
 legend(x,y,c("Simulated","Theoretical"),col=c("Blue","Red"),lwd=4,lty=1:2,bty="n",title=title) # Legend #  
 return()#list(w=w,summary=summary(w),sd=sd(w),fun="Chi-sq")) # some output to use if needed  
}  
mychisim(n1 = 10, iter=1000,mean = 10, sigma = 4,ymax=0.15)



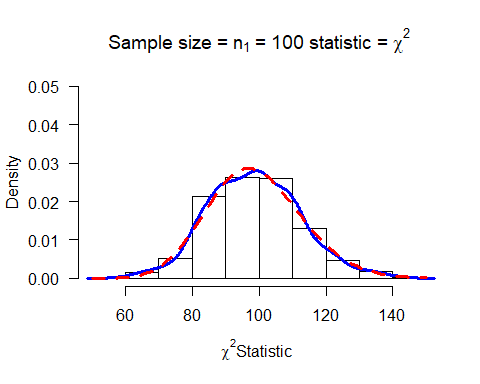
## NULL

mychisim(n1 = 20, iter=1000,mean = 10, sigma = 4,ymax=0.12)



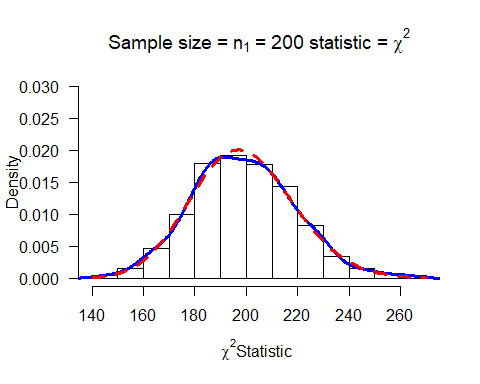
## NULL

mychisim(n1 = 100, iter=1000,mean = 10, sigma = 4,ymax=0.05)



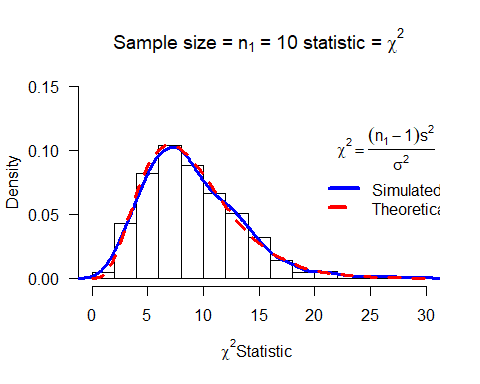
## NULL

mychisim(n1 = 200, iter=1000,mean = 10, sigma = 4,ymax=0.03)



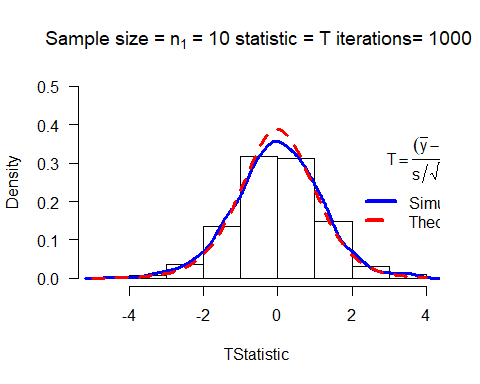
## NULL

chisq=mychisim(n1 = 10, iter=1500,mean = 20, sigma = 10,ymax=0.15)



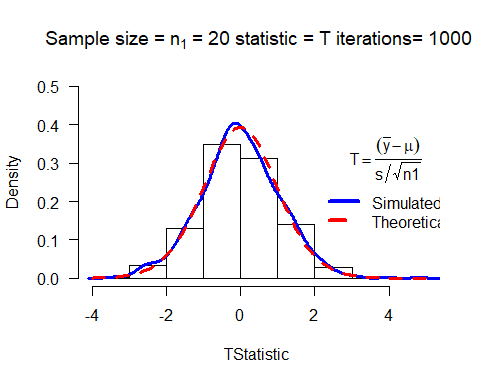
# T Distribution for 1 population

myTsim<-function(n1=10,sigma1=3,mean1=5,iter=1000,ymax=0.1,x=2,y=0.3,...){ # adjust ymax to make graph fit  
   
 y1=rnorm(n1\*iter,mean=mean1,sd=sigma1)# generate iter samples of size n1  
   
 data1.mat=matrix(y1,nrow=n1,ncol=iter,byrow=TRUE) # Each column is a sample size n1  
   
 sd1=apply(data1.mat,2,sd) # sd  
 ybar=apply(data1.mat,2,mean) # mean  
   
 w=(ybar-mean1)/(sd1/sqrt(n1)) #T stat  
   
 hist(w,freq=FALSE, ylim=c(0,ymax), # Histogram with annotation  
 main=substitute(paste("Sample size = ",n[1]," = ",n1," statistic = ",T," iterations= ",iter)),  
 xlab=expression(paste(T, "Statistic",sep=" ")), las=1)  
 lines(density(w),col="Blue",lwd=3) # add a density plot  
 curve(dt(x,n1-1),add=TRUE,col="Red",lty=2,lwd=3) # add a theoretical curve  
 title=expression(T==frac((bar(y)-mu),s/sqrt(n1))) #mathematical annotation -see ?plotmath  
 legend(x,y,c("Simulated","Theoretical"),col=c("Blue","Red"),lwd=4,lty=1:2,bty="n",title=title) # Legend #  
 return()#list(w=w,summary=summary(w),sd=sd(w),fun="T")) # some output to use if needed  
}  
myTsim(n1 = 10, iter=1000,mean = 10, sigma = 4,ymax=0.5)



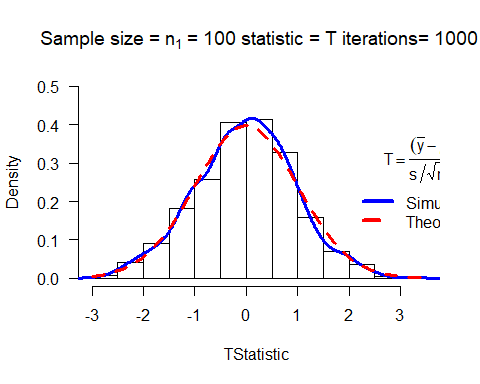
## NULL

myTsim(n1 = 20, iter=1000,mean = 10, sigma = 4,ymax=0.5)



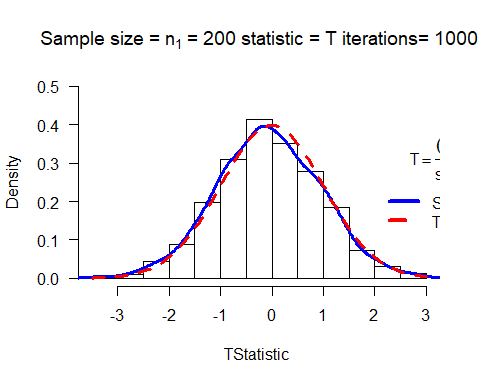
## NULL

myTsim(n1 = 100, iter=1000,mean = 10, sigma = 4,ymax=0.5)



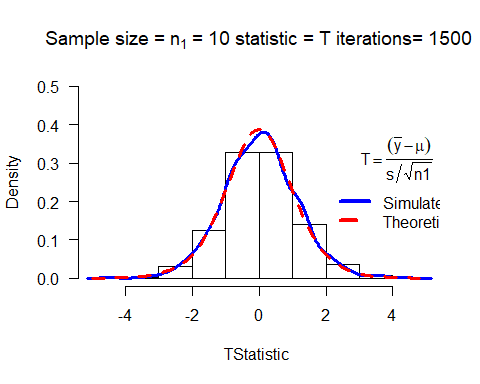
## NULL

myTsim(n1 = 200, iter=1000,mean = 10, sigma = 4,ymax=0.5)



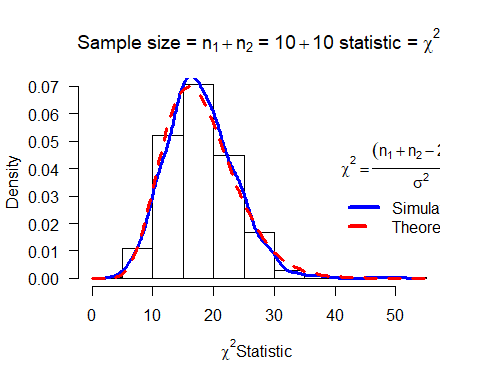
## NULL

T<-myTsim(n1 = 10, iter=1500,mean = 20, sigma = 10,ymax=0.50)



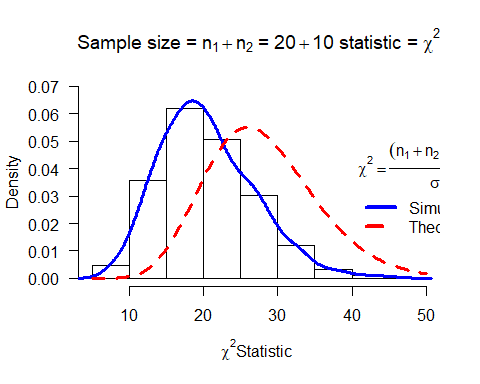
# Chi-Squared Distribution with 2 populations

mychisim2<-function(n1=10,n2=14,sigma1=3,sigma2=3,mean1=5,mean2=10,iter=1000,ymax=0.07,x=40,y=0.04,...){ # adjust ymax to make graph fit  
   
 y1=rnorm(n1\*iter,mean=mean1,sd=sigma1)# generate iter samples of size n1  
 y2=rnorm(n2\*iter,mean=mean2,sd=sigma2)  
   
 data1.mat=matrix(y1,nrow=n1,ncol=iter,byrow=TRUE) # Each column is a sample size n1  
 data2.mat=matrix(y2,nrow=n2,ncol=iter,byrow=TRUE)  
   
 ssq1=apply(data1.mat,2,var) # ssq1 is s squared  
 ssq2=apply(data2.mat,2,var)  
   
 spsq=((n1-1)\*ssq1 + (n2-1)\*ssq2)/(n1+n2-2) # pooled s squared  
   
 w=(n1+n2-2)\*spsq/(sigma1^2)#sigma1=sigma2, Chi square stat  
   
 hist(w,freq=FALSE, ylim=c(0,ymax), # Histogram with annotation  
 main=substitute(paste("Sample size = ",n[1]+n[2]," = ",n1+n2," statistic = ",chi^2)),  
 xlab=expression(paste(chi^2, "Statistic",sep=" ")), las=1)  
 lines(density(w),col="Blue",lwd=3) # add a density plot  
 curve(dchisq(x,n1+n2-2),add=TRUE,col="Red",lty=2,lwd=3) # add a theoretical curve  
 title=expression(chi^2==frac((n[1]+n[2]-2)\*S[p]^2,sigma^2)) #mathematical annotation -see ?plotmath  
 legend(x,y,c("Simulated","Theoretical"),col=c("Blue","Red"),lwd=4,lty=1:2,bty="n",title=title) # Legend #  
 return()#list(w=w,summary=summary(w),sd=sd(w),fun="Chi-sq")) # some output to use if needed  
}  
mychisim2(n1=10,n2=10,mean1=5,mean2=10,sigma1=4,sigma2=4,iter=1000)



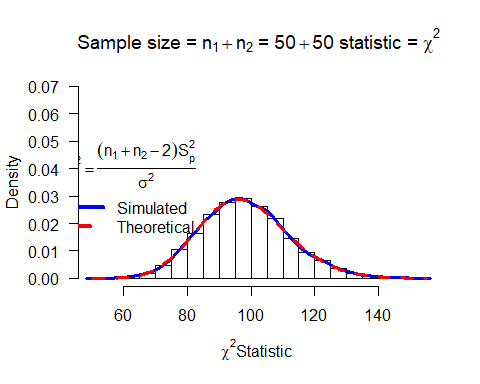
## NULL

mychisim2(n1=20,n2=10,mean1=3,mean2=5,sigma1=10,sigma2=4,iter=1000)



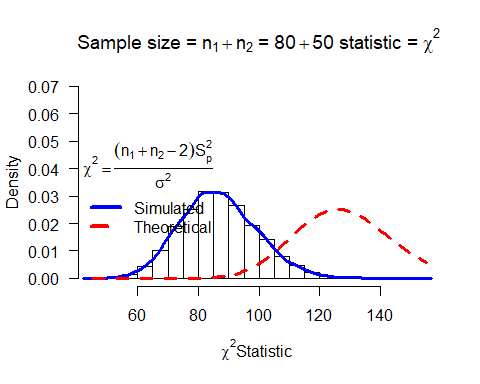
## NULL

mychisim2(n1=50,n2=50,mean1=5,mean2=10,sigma1=4,sigma2=4,iter=10000)



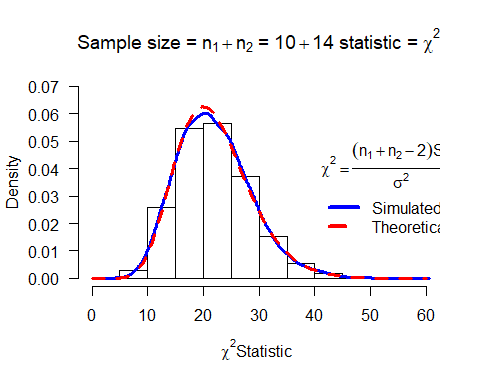
## NULL

mychisim2(n1=80,n2=50,mean1=3,mean2=5,sigma1=10,sigma2=4,iter=10000)



## NULL

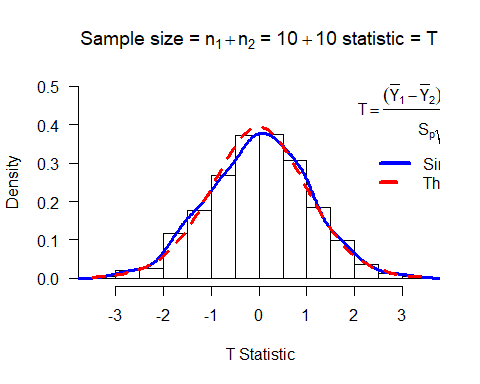
mychisim2(iter=10000)



## NULL

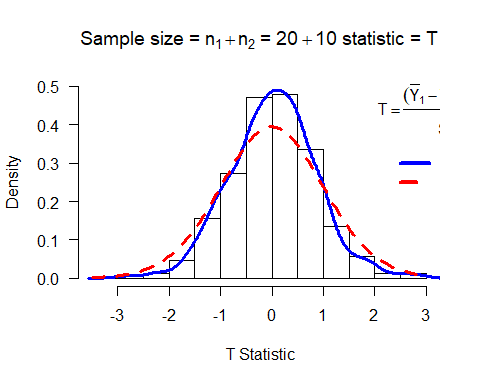
# T Distribution with 2 populations

myTsim2<-function(n1=10,n2=14,sigma1=3,sigma2=3,mean1=5,mean2=10,iter=1000,ymax=0.5,x=2,y=0.4,...){  
   
 y1=rnorm(n1\*iter,mean=mean1,sd=sigma1)# generate iter samples of size n1  
 y2=rnorm(n2\*iter,mean=mean2,sd=sigma2)  
   
 data1.mat=matrix(y1,nrow=n1,ncol=iter,byrow=TRUE) # Each column is a sample size n1  
 data2.mat=matrix(y2,nrow=n2,ncol=iter,byrow=TRUE)  
   
 ssq1=apply(data1.mat,2,var) # ssq1 is s squared  
 ssq2=apply(data2.mat,2,var)  
   
 ybar1= apply(data1.mat,2,mean)  
 ybar2=apply(data2.mat,2,mean)  
   
 spsq=((n1-1)\*ssq1 + (n2-1)\*ssq2)/(n1+n2-2) # pooled s squared  
   
 w=((ybar1-ybar2)-(mean1-mean2))/sqrt(spsq\*(1/n1+1/n2))#sigma1=sigma2, Chi square stat  
   
 hist(w,freq=FALSE, ylim=c(0,ymax), # Histogram with annotation  
 main=substitute(paste("Sample size = ",n[1]+n[2]," = ",n1+n2," statistic = ",T)),  
 xlab=paste(" T Statistic",sep=""), las=1)  
 lines(density(w),col="Blue",lwd=3) # add a density plot  
 curve(dt(x,n1+n2-2),add=TRUE,col="Red",lty=2,lwd=3) # add a theoretical curve  
 title=expression(T==frac((bar(Y)[1]-bar(Y)[2])-(mu[1]-mu[2]),S[p]\*sqrt(frac(1,n[1])+frac(1,n[2])))) #mathematical annotation -see ?plotmath  
 legend(x,y,c("Simulated","Theoretical"),col=c("Blue","Red"),lwd=4,lty=1:2,bty="n",title=title)# Legend #  
 return()#list(w=w,summary=summary(w),sdw=sd(w),fun="T")) # some output to use if needed  
}  
myTsim2(n1=10,n2=10,mean1=5,mean2=10,sigma1=4,sigma2=4,iter=1000)



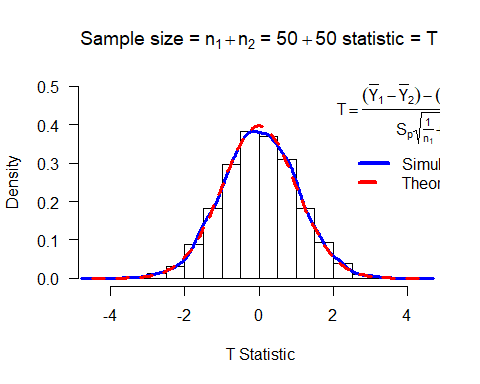
## NULL

myTsim2(n1=20,n2=10,mean1=3,mean2=5,sigma1=10,sigma2=4,iter=1000)



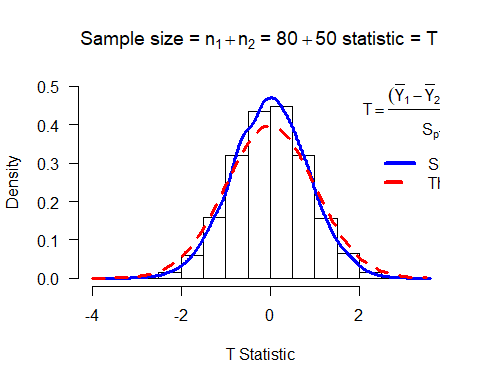
## NULL

myTsim2(n1=50,n2=50,mean1=5,mean2=10,sigma1=4,sigma2=4,iter=10000)



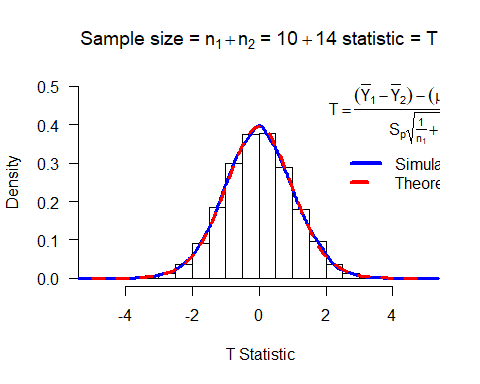
## NULL

myTsim2(n1=80,n2=50,mean1=3,mean2=5,sigma1=10,sigma2=4,iter=10000)



## NULL

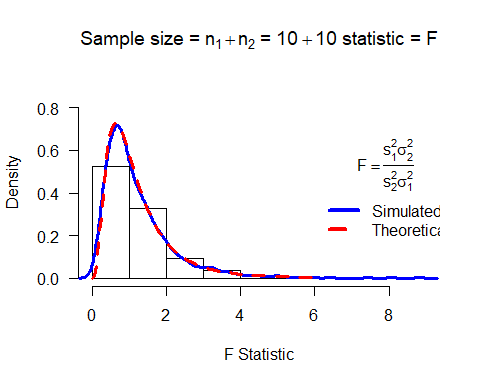
myTsim2(iter=10000)



## NULL

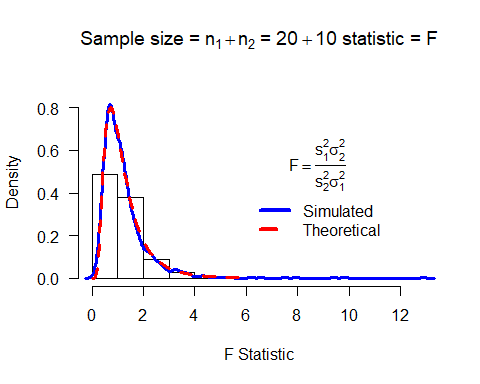
# F Distribution with 2 populations

myFsim2<-function(n1=10,n2=14,sigma1=3,sigma2=2,mean1=5,mean2=10,iter=1000,ymax=0.9,x=6,y=0.5,...){  
   
 y1=rnorm(n1\*iter,mean=mean1,sd=sigma1)# generate iter samples of size n1  
 y2=rnorm(n2\*iter,mean=mean2,sd=sigma2)  
   
 data1.mat=matrix(y1,nrow=n1,ncol=iter,byrow=TRUE) # Each column is a sample size n1  
 data2.mat=matrix(y2,nrow=n2,ncol=iter,byrow=TRUE)  
   
 ssq1=apply(data1.mat,2,var) # ssq1 is s squared  
 ssq2=apply(data2.mat,2,var)  
   
 #spsq=((n1-1)\*ssq1 + (n2-1)\*ssq2)/(n1+n2-2) # pooled s squared  
   
 w=ssq1\*sigma2^2/(ssq2\*sigma1^2) #  
   
 hist(w,freq=FALSE, ylim=c(0,ymax), # Histogram with annotation  
 main=substitute(paste("Sample size = ",n[1]+n[2]," = ",n1+n2," statistic = ",F)),  
 xlab=paste("F Statistic",sep=""), las=1)  
 lines(density(w),col="Blue",lwd=3) # add a density plot  
 curve(df(x,n1-1,n2-1),xlim=c(0,6),add=TRUE,col="Red",lty=2,lwd=3) # add a theoretical curve  
 title=expression(F==frac(s[1]^2,s[2]^2)\*frac(sigma[2]^2,sigma[1]^2)) #mathematical annotation -see ?plotmath  
 legend(x,y,c("Simulated","Theoretical"),col=c("Blue","Red"),lwd=4,lty=1:2,bty="n",title=title)# Legend #  
 return()#list(w=w,summary=summary(w),sd=sd(w),fun="F")) # some output to use if needed  
}  
myFsim2(n1=10,n2=10,mean1=5,mean2=10,sigma1=4,sigma2=4,iter=1000)



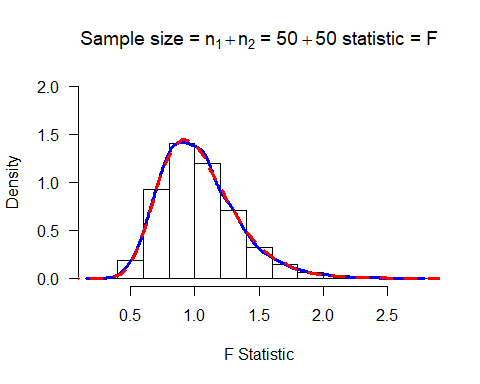
## NULL

myFsim2(n1=20,n2=10,mean1=3,mean2=5,sigma1=10,sigma2=4,iter=1000)



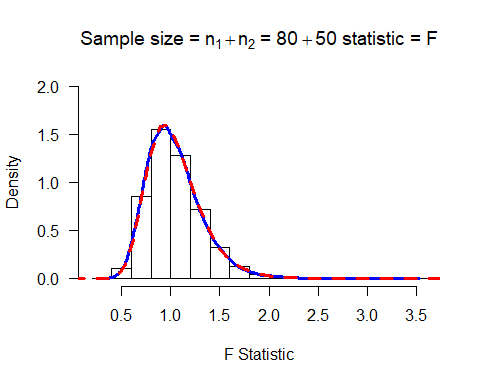
## NULL

myFsim2(n1=50,n2=50,mean1=5,mean2=10,sigma1=4,sigma2=4,iter=10000, ymax=2)



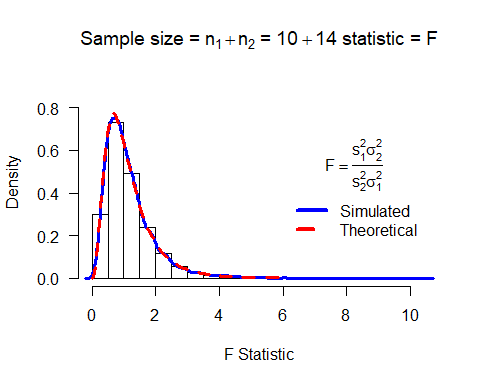
## NULL

myFsim2(n1=80,n2=50,mean1=3,mean2=5,sigma1=10,sigma2=4,iter=10000, ymax=2)



## NULL

myFsim2(iter=10000)



## NULL

In F distribution, we assume the data points are independent from one another.