InClassScript5

Qingqing Yang

04/06/2022

## sampling from normal distribution

help(rnorm)

## starting httpd help server ... done

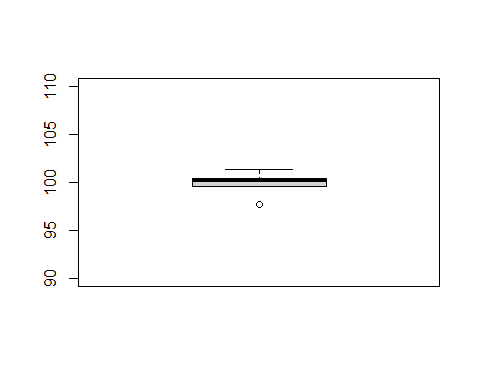
# rnorm generate random deviates from a normal dis  
q1\_1=rnorm(10,mean=100,sd=1)  
q1\_2=rnorm(10,mean=100,sd=10)  
summary(q1\_1)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 97.74 99.75 100.20 100.07 100.39 101.36

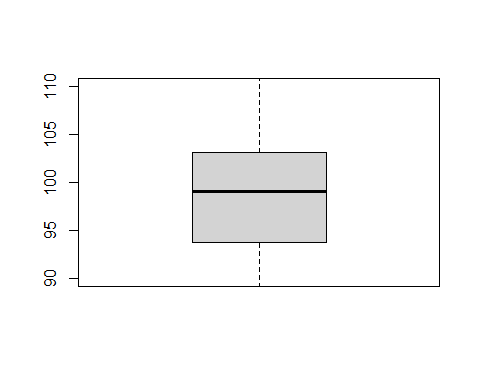
summary(q1\_2)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 78.31 94.26 99.09 97.89 102.50 116.08

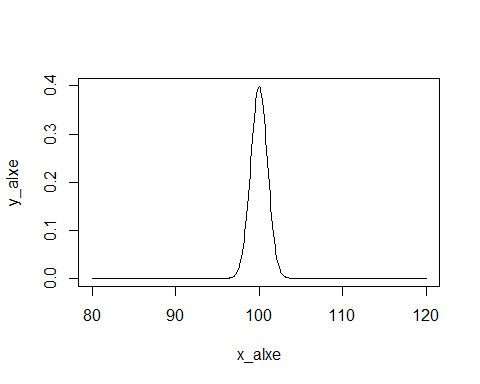
# plot the x and y  
boxplot(q1\_1,ylim=c(90,110))



boxplot(q1\_2,ylim=c(90,110))



# q1\_1 was drawn from the dis as following:  
x\_alxe=seq(80, 120, by = .1)  
y\_alxe=dnorm(x\_alxe, mean = 100, sd = 1)  
plot(x\_alxe,y\_alxe,type="l")



## t-tests

# one sample t-test

q2\_df=rnorm(10,mean=45,sd=1)  
q2\_df

## [1] 46.51209 47.21041 45.53811 43.70306 46.06508 45.84592 45.00750 44.74327  
## [9] 45.70978 45.94202

# compare with mu by code  
t.test(q2\_df, mu = 50)

##   
## One Sample t-test  
##   
## data: q2\_df  
## t = -14.218, df = 9, p-value = 1.794e-07  
## alternative hypothesis: true mean is not equal to 50  
## 95 percent confidence interval:  
## 44.93208 46.32336  
## sample estimates:  
## mean of x   
## 45.62772

#df =n-1= 9

# compare with mu by hand  
q2\_mean=mean(q2\_df)  
q2\_se=sd(q2\_df)/sqrt(10)  
p5\_2tail=qt(c(.025, .975), df=9)  
CI95\_lower=q2\_mean+p5\_2tail[1]\*q2\_se  
CI95\_upper=q2\_mean+p5\_2tail[2]\*q2\_se  
CI95\_lower

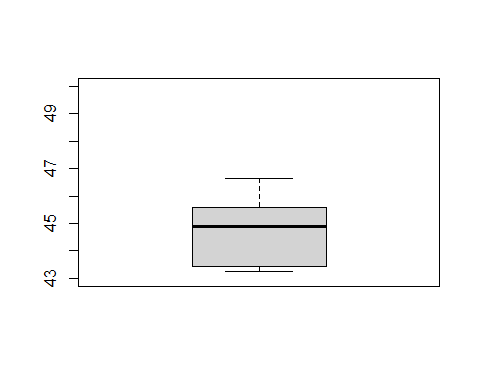
## [1] 44.93208

CI95\_upper

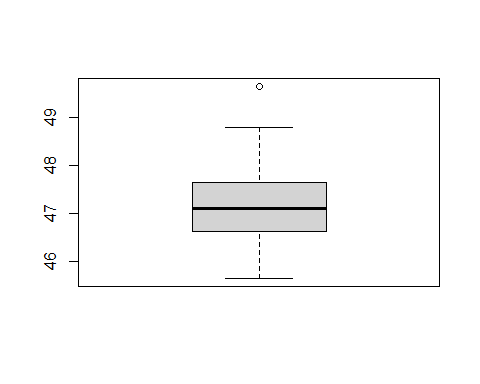
## [1] 46.32336

# two samples independent t-test

q2\_df=rnorm(10,mean=45,sd=1)  
q3\_df=rnorm(10,mean=47,sd=1)  
boxplot(q2\_df,ylim=c(43,50))



boxplot(q3\_df)



t.test(q2\_df,q3\_df,var.equal = TRUE)

##   
## Two Sample t-test  
##   
## data: q2\_df and q3\_df  
## t = -5.0619, df = 18, p-value = 8.122e-05  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -3.638750 -1.504193  
## sample estimates:  
## mean of x mean of y   
## 44.76139 47.33286

#df =N-2=2n-2=18

# two samples dependent t-test

t.test(q2\_df,q3\_df, paired = T)

##   
## Paired t-test  
##   
## data: q2\_df and q3\_df  
## t = -5.6455, df = 9, p-value = 0.0003154  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -3.601855 -1.541087  
## sample estimates:  
## mean of the differences   
## -2.571471

# t = -2.3333, df =n-1=9, p-value = 0.0445