Assignment2

Syed Muhammad Adeel Ibrahim

April 2, 2019

## Including Q1

1. {28.88,7.31,2.47,1.74,20.62,17.38,8.52,6.72,29.35,42.86} is a sample from a normal population with mean Âµ and standard deviation sigma. It is already known that the standard deviation is either 15, 20 or 25. By the maximum likelihood (ML) method estimate µ and sigma.

x = c(28.88, 7.31, 2.47, 1.74, 20.62, 17.38, 8.52, 6.72, 29.35, 42.86)  
mu <- seq(0, 60, .01)  
sigma <- seq(15, 25, .01)  
  
ll2 <- function(theta1, theta2) {   
   
 if(length(theta1) < length(theta2)) {  
 theta1 <- rep(theta1, len=length(theta2))  
 }  
   
 if(length(theta2) < length(theta1)) {  
 theta2 <- rep(theta2, len=length(theta1))   
 }  
   
 ans <- numeric(length(theta1))   
   
 for(i in 1:length(ans)) {  
 density <- dnorm(x, theta1[i], theta2[i])  
 ans[i] <- prod(density)   
 }   
   
 ans   
}  
  
  
z <- outer(mu, sigma, ll2)  
i <- which(z == max(z), TRUE)  
  
theta <- c(mu[i[1]], sigma[i[2]])  
  
ll <- function(theta) {   
 prod(dnorm(x, theta[1], theta[2]))  
}  
  
z <- suppressWarnings(optim(theta, ll, control = list(fnscale=-1), method="BFGS", hessian = TRUE))  
  
muSigmaMLE <- z$par  
muSigmaMLE

## [1] 16.58 15.00

## Including Q2

1. A small lawnmower company produced 1500 lawnmowers in 1998. In an effort to determine how maintenance-free these units were, the company decided to conduct a multi-year study of the 1998 lawnmowers. A sample of 200 owners of these lawnmowers was drawn randomly from company records and contacted. The owners were given an 800 number and asked to call the company when the frst major repair was required for the lawnmowers. Owners who no longer used the lawnmower to cut their grass were disqualified. After many years, 187 of the owners had reported. The other 13 disqualified themselves. The average number of years until the first major repair was 5.3 for the 187 owners reporting. It is believed that the population standard deviation was 1.28 years. If the company wants to advertise and average number of years of repair-free lawn mowing for this lawnmower, what is the point estimate? Construct a 89% confidence interval for the average number of years until the first major repair.

n = 187  
xbar = 5.3  
sigma = 1.28  
  
q = qnorm(1-.11/2, 0, 1)  
  
(lvalue = xbar - (q \* sigma/sqrt(n)))

## [1] 5.150405

(uvalue = xbar + (q \* sigma/sqrt(n)))

## [1] 5.449595

## Including Q3

1. To compare two programs for training industrial workers to perform a skilled job, 20 workers are included in an experiment. Of these 10 are selected at random to be trained by method 1; the remaining 10 workers are to be trained by method 2. After completion of training, all the workers are subjected to a time-and-motion test that records the speed of performance of a skilled job. The following data are obtained:

Time (in minutes) Method 1 | 15 20 11 23 16 21 18 16 27 24 Method 2 | 23 31 13 19 23 17 28 26 25 28

method1 = c (15, 20, 11, 23, 16, 21, 18, 16, 27, 24)  
method2 = c (23, 31, 13, 19, 23, 17, 28, 26, 25, 28)

1. Can you conclude from the data that the main job time is significantly less after training with method 1 than after training with method 2? (Test with ?? = 0.05)

# H0 : mu1 = mu2  
# HA : mu1 < mu2  
  
xbar = mean(method1)  
ybar = mean(method2)  
  
S1 = sd(method1)  
S2 = sd(method2)  
n1 = length(method1)  
n2 = length(method2)  
  
Spooled = sqrt(((n1-1) \* (S1^2) + (n2-1) \* (S2^2)) / (n1 + n2 - 2))  
  
t = (xbar - ybar) / (Spooled \* sqrt((1/n1) + (1/n2)))  
  
c1 = qt(1-.05, n1 + n2 - 2)  
  
if (t < -c1) {  
 message("T is in rejection region, so we reject hypothesis.")  
} else {  
 message("T is not in rejection region, so we don't reject hypothesis.")   
}

## T is in rejection region, so we reject hypothesis.

1. State the assumptions you make for the population distributions.
2. Both distributions are normal
3. The population variances are equal
4. Construct a 95% confidence interval for the population mean different in job times between the two methods.

c = qt(1-.05/2, n1 + n2 - 2)  
(U = (xbar - ybar) + c \* (Spooled \* sqrt((1/n1) + (1/n2))))

## [1] 0.6871218

(L = (xbar - ybar) - c \* (Spooled \* sqrt((1/n1) + (1/n2))))

## [1] -9.087122

## Including Q4

1. To determine how an experimental dose of a dental anesthesia affects male and female patients, random samples of 15 male and 16 female patients are selected and their reaction times are recorded in minutes. The mean and the standard deviations obtained from the data sets are: | Male | Female Mean | 4.8 | 4.4 sd | 0.8 | 0.9
2. Devise a test with ?? = 0.1 to determine whether there is a significant difference in the mean reaction times between the males and females.

n1 = 15  
Xbar = 4.8  
s1 = .8  
  
n2 = 16  
Ybar = 4.4  
s2 = .9  
  
# H0 : mu1 = mu2  
# HA : mu1 != mu2  
  
(c1 = qnorm(1 - .01/2, 0, 1))

## [1] 2.575829

(c2 = qnorm(.01/2, 0, 1))

## [1] -2.575829

(z = (Xbar - Ybar) / sqrt((s1^2/n1) + (s2^2/n2)))

## [1] 1.3096

if (z > c2 && z < c1) {  
 message("Z is not in rejection region, therefore there is no significant difference in the mean in between males and females")  
} else {  
 message("Z is in rejection region, therefore null hypothesis is rejected.")  
}

## Z is not in rejection region, therefore there is no significant difference in the mean in between males and females

1. Construct a 95% confidence interval for the difference between the mean reaction times of males and females.

c = qnorm(1 -.05/2, 0, 1)  
  
(Lval = (Xbar - Ybar) - c \* sqrt((s1^2/n1) + (s2^2/n2)))

## [1] -0.1986452

(Uval = (Xbar - Ybar) + c \* sqrt((s1^2/n1) + (s2^2/n2)))

## [1] 0.9986452

1. Construct a 99% confidence interval for the mean reaction time of each group individually.

c = qnorm(1 -.01/2, 0, 1)  
  
(LMale = Xbar - c \* s1 / sqrt(n1))

## [1] 4.267939

(UMale = Xbar + c \* s1 / sqrt(n1))

## [1] 5.332061

(LFemale = Ybar - c \* s2 / sqrt(n2))

## [1] 3.820438

(UFemale = Ybar + c \* s2 / sqrt(n2))

## [1] 4.979562

## Including Q5

1. A sociologist wishes to compare the fertility rates of women in two tribal sects A and B of eastern Africa. From each sect a random sample of 100 women in the age group 50-60 years is selected, and the number of children born to each is recorded. The following frequency distributions are obtained:

Number of children | 0 1 2 3 4 5 6 7 8 | Total A | 6 14 18 25 19 11 5 2 3 | 100 B | 0 3 8 18 30 19 15 5 2 | 100

1. Calculate the mean and the standard deviation for each frequency distribution.

sectA = c(6, 14, 18, 25, 19, 11, 5, 2, 3)  
sectB = c(0, 3, 8, 18, 30, 19, 15, 5, 2)  
  
(meanSectA = mean(sectA))

## [1] 11.44444

(sdSectA = sd(sectA))

## [1] 8.079466

(meanSectB = mean(sectB))

## [1] 11.11111

(sdSectB = sd(sectB))

## [1] 10.00555

1. Do the data indicate a significant difference in the mean number of children born to women in the two sects?

# H0 : mu1 = mu2  
# HA : mu1 != mu2  
  
n1 = 100  
n2 = 100  
s1 = sdSectA  
s2 = sdSectB  
Xbar = meanSectA  
Ybar = meanSectB  
  
(c1 = qnorm(1, 0, 1))

## [1] Inf

(c2 = qnorm(0, 0, 1))

## [1] -Inf

(z = (Xbar - Ybar) / sqrt((s1^2/n1) + (s2^2/n2)))

## [1] 0.2591944

if (z > c2 && z < c1) {  
 message("Z is not in rejection region, there is no significant difference in between children born in each sect")  
} else {  
 message("Z is in rejection region, therefore null hypothesis is rejected.")  
}

## Z is not in rejection region, there is no significant difference in between children born in each sect

1. Construct a 98% confidence interval for the difference between the population means.

(c = qnorm(1 -.02/2, 0, 1))

## [1] 2.326348

(Lval = (Xbar - Ybar) - c \* sqrt((s1^2/n1) + (s2^2/n2)))

## [1] -2.658434

(Uval = (Xbar - Ybar) + c \* sqrt((s1^2/n1) + (s2^2/n2)))

## [1] 3.325101

## Including Q6

1. Measurements of the left-hand and right-hand gripping strengths of 10 left-handed writers are recorded: Person | 1 2 3 4 5 6 7 8 9 10 Left hand | 140 90 125 130 95 121 85 97 131 110 Right hand | 138 87 110 132 96 120 86 90 129 100
2. Do the data provide strong evidence that people who write with the left hand have a greater gripping strength in the left hand than they do in the right hand? (?? = 0.05)

LH = c(140, 90, 125, 130, 95, 121, 85, 97, 131, 110)  
RH = c(138, 87, 110, 132, 96, 120, 86, 90, 129, 100)  
  
d = LH - RH  
  
dbar = mean(d)  
  
sdD = sd(d)  
  
n = 10  
df = n - 1  
  
# H0 : muD <= 0  
# HA : muD > 0  
  
(c1 = qt(.05/2, df))

## [1] -2.262157

(c2 = qt(1 - .05/2, df))

## [1] 2.262157

(t = dbar / (sdD \* sqrt(n)))

## [1] 0.2084647

if (t > c2 && t < c1) {  
 message(bquote("Z is not in rejection region, Left hand griping is not strong"))  
} else {  
 message("Z is in rejection region, therefore null hypothesis is rejected. Left hand gripping is stronger that right hand")  
}

## Z is in rejection region, therefore null hypothesis is rejected. Left hand gripping is stronger that right hand

1. Construct a 90% confdence interval for the mean difference.

c = qt(1 - 0.1/2, n - 1)  
  
(U = (dbar) + c \* (sdD / sqrt(n)))

## [1] 6.765623

(L = (dbar) - c \* (sdD / sqrt(n)))

## [1] 0.4343768