Partial answer key

6.5

a.

The level of significance is α = 0.05.

Confidence interval method:

1.

Ho: μ1 - μ2 ≥ 0  
Ha: μ1 - μ2 < 0  
where population #1 corresponds to 26°C and population #2 corresponds to 5°C.  
2.

The “one-sided” C.I. for μ1 - μ2 is (-∞, -191.4)

> ex6.5<-read.table("ex6-5.txt", header=T)

> BP26 <- ex6.5$BloodPressure[ex6.5$Temp == 26]

> BP5 <- ex6.5$BloodPressure[ex6.5$Temp == 5]

> t.test(x = BP26, y = BP5, data = ex6.5, var.equal = FALSE,

conf.level = 0.95, alternative = "less")

Welch Two Sample t-test

data: BP26 and BP5

t = -19, df = 8.3, p-value = 2e-08

alternative hypothesis: true difference in means is less than 0

95 percent confidence interval:

-Inf -191.4

sample estimates:

mean of x mean of y

165.8 378.5

3. Since 0 is outside of this interval, reject Ho.

4. There is enough evidence indicating that rats exposed to 5°C have a higher mean blood pressure than rats exposed to 26°C.

Test statistic method:

1.

Ho: μ1 - μ2 ≥ 0  
Ha: μ1 - μ2 < 0

2.

save.it <- t.test(x = BP26, y = BP5, data = ex6.5, var.equal =

FALSE, alternative = "less")

names(save.it)

save.it$statistic

t=-18.51 (R rounds some of the values in the output)

3.

> save.it$parameter

df

8.321

-t0.05, 8.321 = -1.85

> qt(p = 0.05, df = 8.321)

[1] -1.850291



4. Because -18.51 < -1.85, reject Ho.

5. There is enough evidence indicating that rats exposed to 5°C have a higher mean blood pressure than rats exposed to 26°C.

P-value method:

1.

Ho: μ1 - μ2 ≥ 0  
Ha: μ1 - μ2 < 0

2.

t=-18.51.

The p-value is P(T < -18.51) which is very small (<0.0001).

(using the same R code as above)

3.

α = 0.05

4. Since p-value<0.0001<0.05, reject Ho.

5. There is sufficient evidence that rats exposed to 5°C have a higher mean blood pressure than rats exposed to 26°C.

b. The power here is the P(T < -t0.05, ν)

If the effect size is 10 and σ1 = σ2 = 20, to achieve a power of 0.9, the sample size for each group should be at least 70.

> effect.size <- 10

> alpha <- 0.05

> sigma <- 20

> power.t.test(n = seq(from = 10, to = 100, by

= 10), d = effect.size, sd = sigma,

sig.level = alpha, type = "two.sample", alternative =

"one.sided")

Two-sample t test power calculation

n = 10, 20, 30, 40, 50, 60, 70, 80, 90, 100

delta = 10

sd = 20

sig.level = 0.05

power = 0.2847635, 0.4633743, 0.6060253, 0.7162549, 0.7989362,

0.8594840, 0.9029656, 0.9336887, 0.9551006, 0.9698479

alternative = one.sided

NOTE: n is number in \*each\* group

This type of problem will not be on the test. I just wanted to give you some exposure to power relative to these types of hypothesis tests.

c. Note that our authors assume that a two-sided interval is of interest whenever a CI is requested.

A 95% CI on the mean difference μ1 - μ2 is (-238.98, -186.35).

> t.test(x = BP26, y = BP5, data = ex6.5, var.equal = FALSE, conf.level =

0.95)

Welch Two Sample t-test

data: BP26 and BP5

t = -19, df = 8.3, p-value = 5e-08

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-239.0 -186.3

sample estimates:

mean of x mean of y

165.8 378.5

Another way to approach the code of this problem is as follows:

> t.test(formula = BloodPressure ~ Temp, data = ex6.5, var.equal

= FALSE, conf.level = 0.95)

Welch Two Sample t-test

data: BloodPressure by Temp

t = 19, df = 8.3, p-value = 5e-08

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

186.3 239.0

sample estimates:

mean in group 5 mean in group 26

378.5 165.8

R reserves the order of the temperatures from what we want, so here is how we can get “1” to represent 26 degrees and “2” to represent 5 degrees.

> ex6.5$Tempv2 <- factor(ex6.5$Temp, levels = c("26", "5"))

> t.test(formula = BloodPressure ~ Tempv2, data = ex6.5,

var.equal = FALSE, conf.level = 0.95)

Welch Two Sample t-test

data: BloodPressure by Tempv2

t = -19, df = 8.3, p-value = 5e-08

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-239.0 -186.3

sample estimates:

mean in group 26 mean in group 5

165.8 378.5

6.11

a.

The level of significance is α = 0.05

Confidence interval method:

1.

Ho: μ1 - μ2 ≤ 0  
Ha: μ1 - μ2 > 0  
where μ1 corresponds to year 1982 and μ2 corresponds to year 1996.

2. The “one-sided” C.I. for μ1 - μ2 is (30.67, ∞)

> ex6.11 <- read.table("ex6-11.txt", header=T)

> x1 <- as.numeric(ex6.11[,2])

> y1 <- as.numeric(ex6.11[,3])

> t.test(x = x1, y = y1, var.equal = FALSE, conf.level = 0.95, alternative = "greater")

Welch Two Sample t-test

data: x1 and y1

t = 8.3522, df = 15.412, p-value = 2.074e-07

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

30.67291 Inf

sample estimates:

mean of x1 mean of y1

54.32769 15.52462

3. Since 0 is not in the interval, reject Ho.

4. There is enough evidence to indicate that there has been a significant decrease in mean PCB content.

Test statistic method:

1.

Ho: μ1 - μ2 ≤ 0  
Ha: μ1 - μ2 > 0  
2. t=8.352 (using the same R code as above)

3. t0.05, 15.412 = 1.75

> qt(p = 0.95, df = 15.412)

[1] 1.749978



4. Since 8.352 > 1.75, reject Ho.



5. There is enough evidence to indicate that there has been a significant decrease in mean PCB content.

P-value method:

1.

Ho: μ1 - μ2 ≤ 0  
Ha: μ1 - μ2 > 0

2. t=8.352. The p-value is P(T > 8.352) which is very small (<0.0001)

3. α = 0.05

4. Since p-value<0.0001<0.05, reject Ho.

5. There is enough evidence to indicate that there has been a significant decrease in mean PCB content.

b. A 95% C.I. on the difference in the mean PCB content of herring gull eggs is (28.92, 48.68).

> t.test(x = x1, y = y1, var.equal = FALSE, conf.level = 0.95)

Welch Two Sample t-test

data: x1 and y1

t = 8.3522, df = 15.412, p-value = 4.148e-07

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

28.92365 48.68251

sample estimates:

mean of x1 mean of y1

54.32769 15.52462

c.

The plots suggest there may be some departures from a normal distribution, but it is difficult to make a conclusive assessment with this small of a sample size. Without a normal distribution and this sample size, there “may” be some concern about if our inferences are correct.





> ex6.10<-data.frame(PCB=c(x1,y1), year=rep(c("1982", "1996"), each=13))

> boxplot(formula = PCB ~ year, data = ex6.10, col = "lightblue",

main = "Box and dot plot", ylab = "PCB", xlab = "year")

> stripchart(x = PCB ~ year, data = ex6.10, method = "jitter", vertical =

TRUE, pch = 1, col="red", add=T)

> par(mfrow=c(2,1))

> hist(x = x1, freq = F, main = "Year 1982", xlab = "PCB", breaks = seq(from

= 0, to = 85, by = 5))

> curve(expr = dnorm(x = x, mean = mean(x1), sd = sd(x1)), col = "red", add =

TRUE)

> hist(x = y1, freq = F, main = "Year 1996", xlab = "PCB", breaks = seq(from

= 0, to = 85, by = 5))

> curve(expr = dnorm(x = x, mean = mean(y1), sd = sd(y1)), col = "red", add =

TRUE)

d.

Because the data for 1982 and 1996 were collected at the same sites, there may be correlation between the two years. In fact, it may be of interested to use a paired sampled way of doing this type of test.

6.28

a.

The level of significance is α = 0.05.

Confidence interval method:

1.

Ho:μd ≤ 0  
Ha:μd > 0

2. The one-side 95% C.I. for μd is (-2.92, ∞)

> ex6.28<-read.table("ex6-28.txt", header=T)

> x1<-as.numeric(ex6.28[,1])

> y1<-as.numeric(ex6.28[,2])

> t.test(x = x1, y = y1, var.equal = FALSE, conf.level = 0.95, paired = TRUE,

alternative = "greater")

Paired t-test

data: x1 and y1

t = 0.861, df = 9, p-value = 0.2058

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

-2.917599 Inf

sample estimates:

mean of the differences

2.584

3. Because 0 is in this interval, do not reject Ho.

4. There is not enough evidence to show that the mean SENS value decreased.

Test statistic method:

1.

Ho:μd ≤ 0  
Ha:μd > 0

2. t=0.861.

3. t0.05, 9 = 1.83.

> qt(p = 0.95, df = 9)

[1] 1.833113



4. Since 0.861 < 1.83, do not reject Ho.

5. There is not enough evidence to show that the mean SENS value decreased.

P-value method:

1.

Ho:μd ≤ 0  
Ha:μd > 0

2. t=0.861. The p-value is P(T > 0.2058) = 0.2.

3. α = 0.05

4. Because 0.2058 > 0.05, do not reject Ho.

5. There is not enough evidence that the mean SENS value decreased.

b.

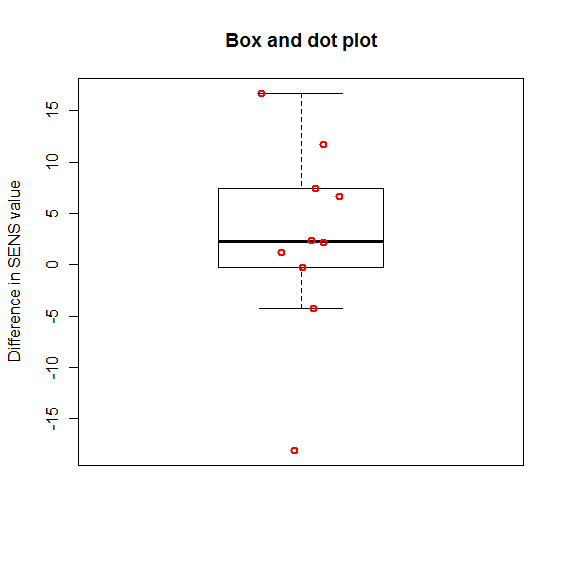
The mean change in SENS value is 2.584.

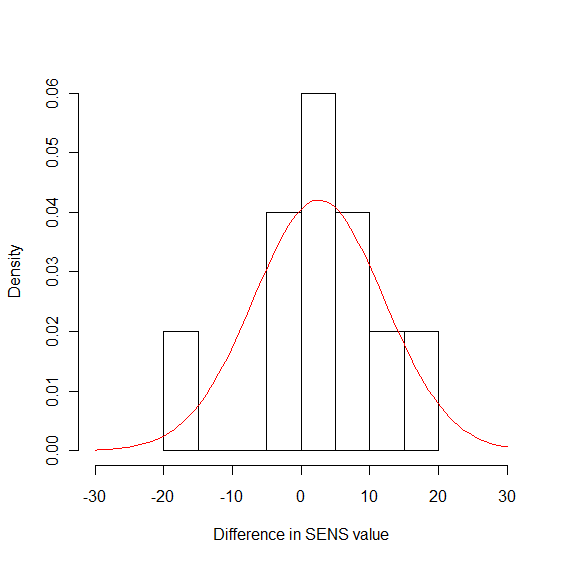
> mean(x1-y1)

[1] 2.584

c.

The box plot indicates that there is one outlier far below all other values. Other values seem to follow normal distribution. This outlier (corresponds to patient 9) should be carefully checked.





> boxplot(x = x1 - y1, main = "Box and dot plot", ylab = "Difference in SENS

value", xlab = "", pars = list(outpch=NA))

> stripchart(x = x1 - y1, lwd = 2, col = "red", method = "jitter", vertical =

TRUE, pch = 1, add = TRUE)

> hist(x = x1 - y1, main = "", xlab = "Difference in SENS value", freq =

FALSE, xlim=c(-30, 30))

> curve(expr = dnorm(x = x, mean = mean(x1-y1), sd = sd(x1-y1)), col =

"red", add = TRUE)

6.42

The level of significance is α = 0.05.

Confidence interval method:

1.

Ho:μd ≥ 0  
Ha:μd < 0

2.

The one-side 95% C.I. for μd is (-∞, -0.65)

> ex6.42<-read.table("ex6-42.txt", header=T, sep = "")

> x1<-as.numeric(ex6.42[,2])

> y1<-as.numeric(ex6.42[,3])> t.test(x = x1, y = y1, var.equal = FALSE, conf.level = 0.95, paired = TRUE,

alternative = "less")

Paired t-test

data: x1 and y1

t = -3.885, df = 11, p-value = 0.001271

alternative hypothesis: true difference in means is less than 0

95 percent confidence interval:

-Inf -0.6497695

sample estimates:

mean of the differences

-1.208333

3. Because 0 is not in this interval, reject Ho.

4. There is enough evidence that the exposure has increased mean lung capacity.

Test statistic method:

1.

Ho:μd ≥ 0  
Ha:μd < 0

2. t=-3.885.

3. -t0.05, 11 = -1.796

> qt(p = 0.05, df = 11)

[1] -1.795885



4. Because -3.885 < -1.796, reject Ho.

5. There is enough evidence that the exposure has increased mean lung capacity.

P-value method:

1.

Ho:μd ≥ 0  
Ha:μd < 0

2. t=-3.885. The p-value is P(T < -3.885) = 0.0013.

3. α = 0.05

4. Since 0.0013 < 0.05, reject Ho.

5. There is enough evidence that the exposure has increased mean lung capacity.

b.

The mean difference in lung capacity is -1.208 and the 95% CI is (-1.89, -0.52).

> t.test(x = x1, y = y1, var.equal = FALSE, conf.level = 0.95, paired = TRUE)

Paired t-test

data: x1 and y1

t = -3.885, df = 11, p-value = 0.002541

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-1.8928932 -0.5237735

sample estimates:

mean of the differences

-1.208333

c.

The researcher is justified in making the statement about the population of rats from which the rats in the study were randomly selected. However the researcher shouldn’t be extending the claim to larger populations (e.g., humans).