**Openintro Questions 5.1 to 25**

5.1 a. n=6 CL = 90%

df = n- 1 = 5 qt(0.05, df = 5, lower.tail = TRUE) = 2.01

b. qt(0.01, df = 20, lower.tail = TRUE) = -2.52

c. qt(0.025, df = 28, lower.tail = TRUE) = 2.04 =

d. qt(0.005, df = 11, lower.tail = TRUE) = 3.10

5.2 Normal – solid curve

Df = 5 n =6 df= 1 n=2

Dashed line = t-distribution with df=5

Dotted line = t-distribution with df=1

5.3 a. pt(1.91,df=10,lower.tail = FALSE) = 0.04260244 < 0.05 – Reject the null hypothesis

b. pt(-3.45,df=16,lower.tail = TRUE) = 0.001 < 0.05 – Reject the null hypothesis

c. pt(0.83,df=6,lower.tail = TRUE) > Greater than 0.05 – Fail to reject null hypothesis

d. pt(2.13,df=27,lower.tail = FALSE) = 0.02 < 0.05 – Reject the null hypothesis

5.4 a. pt(2.485,df=25,lower.tail = FALSE) = 0.0100024 > 0.01 Fail to reject null hypothesis

b. pt(0.5,df=17,lower.tail = TRUE) = 0.68 > 0.01 Fail to reject null hypothesis

5.5 95% CI = (18.985,21.015) n = 36 sample mean? S?

18.985 = Mean – 2.03\*SE 21.015 = Mean + 2.03 \* SE

Using both equations : Mean = 20 SE = 0.5 s=3

5.7

a. H0: mu = 8 HA: mu < 8

b. DF = 24 SE = s/sqrt(25) = 0.77/5 = 0.154

t = 7.73-8/0.154 = -1.75

c. pt(-1.753247,df=24,lower.tail = TRUE) = 0.046 < 0.

d. If we take alpha = 0.05 then we can reject the null hypothesis.

e. 7.73 +/- 1.71\*0.154 (7.5,8) - 8 is just at the edge of the 90% confidence interval.

5.8 a. No, because sample size too small.

b.

5.9 s=8 n= 20

-1.72 = Mean – 60/1.788854

-3.076829 = Mean – 60 = Mean = 56.92317

5.11 sample mean = 4.6 s=2.2yrs n=20

a. H0: mu = 5 HA: mu /= 5 SE = 2.2/sqrt(20)

t = 4.6-5/0.491935 = -0.8 = pvalue = 0.4335999 < 0.05. Do not reject null hypothesis.

b. 4.6 + 2.09 \* 0.491935 = (3.555,5.645) – Yes they agree.

5.13 s=100 ME <= 10 95% confidence level

10 = t \* SE SE = 100/sqrt(n)

10 = t \* 100/sqrt(n)

10 = 1.96 \* 100/sqrt(n)

10/1.96 = 100/sqrt(n)

Sqrt(n) = 100/4.784689 = 20.9

N = 20.9^2 = 436.81

5.15 a. We should use two sided test.

b. We should use paired test

c. t-test should be used.

5.16 a. True

b. True, as the differences have to calculated so the number of entries should be same

c. No

d. No

5.17 a. True

b. False

c. True

d. True

5.19 a. It’s related.

b. H0: mu = There is no global warming

HA: mu=There is global warming

c. The sample is OK. So normal dist and skewness can be assumed.

d. n=51 mean=1.1 s=4.9 SE = 4.9/sqrt(51) = 0.6861372

df = 51+51-2 = 100

t = 1.1-0/0.69 = 1.1/0.69 = 1.594203

pt(1.59,df=100,lower.tail = FALSE) = 0.05749572 > alpha

e. Do not reject the null hypothesis.

f. We might have not reported the difference when actually there is difference. Type 2 error.

g. Yes

5.21

1.1 +/- 1.66\*0.69 = -0.0454, 2.2454 SE = 4.9/sqrt(51)

The difference between the temp in two years lies between this range.

Yes

5.23 a. No they are independent of each other.

b. Se = 7.5/6 = 1.25

t = 3.4-0/1.25 = 2.72

* 1. < 0.05 Reject the null hypothesis

5.25 No, since the histograms are heavily skewed and sample size is very small.

**5.31-5.37**

5.31

a. There is a considerable difference between the median weight for both groups. For horsebean group the mean weight of chickens are lower than mean weight of lineseed group.

b. H0: Mu(line) – Mu(ho) = 0

HA: Mu(line) – Mu(ho) /= 0

Mean diff in sample = 218.75-160.20=58.55 SE = sqrt(52.24^2/12 + 38.63^2/10)

Df = n1+n2-2 = 20

T = 58.55-0/ 19.40 = 3.01 df = min(9,11)

pt(3.01,df=9,lower.tail = FALSE)\*2 = 0.01471585< 0.05 – Reject the null hypothesis.

The means weights of chickens are different.

c. We might have committed Type 1 error.

d.. Fail to reject null hypothesis.

5.32 SE <- sqrt((3.58^2/26)+(4.51^2/26))

t\_score <- (19.85 - 16.12)/SE

pt(t\_score,df=25,lower.tail = FALSE)\*2 = 0.002883615 < 0.05

Reject null hypothesis.

5.33 H0: Mu(caesin) – Mu(soya) = 0

HA: Mu(caesin) – Mu(soya) /= 0

SE <- sqrt((64.43^2/12)+(54.13^2/14))

t\_score <- (323.58-246.43)/SE

pt(t\_score,df=11,lower.tail = FALSE)\*2 = 0.01 < 0.05 – Reject the null hypothesis.

5.34 98% confidence interval df = 25

Point estimate +/- tdf\*SE = 27.88 – 22.92 + 2.485107\*

SE <- sqrt((5.29^2/26)+(5.01^2/26))

mean.diff <-27.88-22.92

t\_Score <- qt(0.01,df=25)

lower <- mean.diff-t\_Score\*SE

upper <- mean.diff+t\_Score\*SE

lower=8.510922

upper=1.409078

5.35 n = 44, Mean1 = 52.1 s1 = 45.1 Mean2=27.1 S2 = 26.4

SE <- sqrt((45.1^2/22)+(26.4^2/22))

t\_score <- (52.1-27.1)/SE

pt(t\_score,df=21,lower.tail = FALSE)\*2 = 0.03575082 < 0.05 Reject the null hypothesis.

5.36 Mean1= 4.9 S1=1.8 Mean2=6.1 s2=1.8 n=22

SE <- sqrt((1.8^2/22)+(1.8^2/22))

t\_score <- (2.9-6.1)/SE

pt(t\_score,df=21,lower.tail = FALSE)\*2 = Large pvalue > 0.05. Do not reject null hypothesis

5.37.

TR1

SE <- 12.3/sqrt(14)

t\_score <- (6.21)/SE = 1.889081

pt(t\_score,df=13,lower.tail = FALSE) = 0.04 < 0.05. Very close to 0.05.

TR2

SE <- 7.94/sqrt(14)

t\_score <- (2.86)/SE = 1.347751

pt(t\_score,df=13,lower.tail = FALSE) = 0.10 > 0.05 Do not reject null hypothesis.

TR3

SE <- 8.57/sqrt(14)

t\_score <- (-3.21)/SE = -1.401484

pt(t\_score,df=13,lower.tail = FALSE) = 0.907 > 0.05 Do not reject null hypothesis

The first treatment is better since the difference between pre and post treatment scores is significant.

**5.41 – 5.49**

+

5.41 alternative

5.42 ANOVA test, as we need to see the difference between multiple groups of students.

5.43 The variance is constant amongst all the groups.

F = variability between groups/variability within groups

Dfg = k-1=6-1=5

Dft=n-1=71-1=70

Dferr=dft-dfg = 70-5=65

Sum of Squares total (SST) = 231,129.16 + 195,556.02 = 426685.2

SST <- 231129.16 + 195556.02

SSG <- 231129.16

(SSE <- SST-SSG)

MSG <- 46225.83

MSE <- 3008.55

F = MSG/MSE = 15.36482

The p-value for the F statistic is 0 < 0.05, so we reject the null hypothesis saying there at least one group for which mean weight of chickens is different.

5.44

H0: Average test scores are same for all teaching methods

HA: Average test score is different for at least one teaching method.

b. DFg = k-1 = 5-1 = 4

DFt = n-1 = 45 -1 = 44

DFerr = DFt- DFg = 44-4 = 40

c. if we consider alpha = 0.05 then 0.01<0.05 so we have strong evidence in favor of the alternative hypothesis.

5.45 Table

Description automatically generated

Diagram

Description automatically generated with medium confidence

1. H0: The mean MET hours are equal for all the groups of coffee consumption.

HA: The mean MET hours is different for at least one group pf coffee consumption.

1. The sample is large enough so we can relax the skew condition. The samples collected are independent of each other as the coffee consumption and exercise are independent for every individual women.
2. Dfg = k-1 = 5-1 = 4

DfT = n-1 = 50738

Dferr = Dft – Dfg = 50738 – 4 = 50734

SSG = SST – SSE = 25575327 – 25564819 = 10508

MSG = 2627

MSE = 504

F-value = MSG/MSE

1. Small p-value denotes that MET average hours differs for at least one pair of groups.

5.46 a. The p-value 0.07 > 0.05 so we don’t enough evidence that average score is different for at least one pair of two sections.

5.47 a. H0: The average GPA scores is same for all the groups.

HA: The average GPA score is different for at least one pair of groups.

b. The p-value is >0.05 so we don’t have enough evidence to reject the null hypothesis.

c. DFt = 2 + 195 = 197

n = DFt + 1 = 197 + 1= 198

5.48 similar question

5.49 a. False - As the number of group increases, the number of comparisons will also increase thereby decreasing the modified significance level.

b. N increases then DFt increases. DFe = DFT -DFG = True

c. True- The constant variance condition can be relaxed when the sample size is large enough .

d. False – We need observations to be independent regardless of large sample size.