Grade sheet for Problem Two Part 3:
Two Sample Equal Variance Independent T Test For Mean Weights on Day 15 for High and Low Exposure Groups

---

Part 3: Is there a difference in weight between the two exposure groups (low and high) at Day 15? Justify

---

Topic:
Previous studies have shown a link between pesticide exposure and thyroid disease, which can lead to increased weight gain, a symptom of thyroid disease.

Population:
Rats who are exposed to the pesticide of interest.

Research Question:
Whether or not there is a difference in mean weights on day 15 between rats who received a low exposure to the pesticide and rats who received a high exposure to the pesticide

Methods

Description of Outcome: The weights of the rats on day 15
Description of Predictor: Low or high exposure of this pesticide

Description of Data Summary on Weights:

Weights on Day 15 -Low Exposure	Weights on Day 15-High Exposure
Mean: 126.58	Mean: 127.79
Sample Size: 49	Sample Size: 61
Standard Deviation: 10.32	Standard Deviation: 10.72
95% CI on Mean: 123.6145,129.5434	95% CI on Mean: 125.0412, 130.5319

Description of Data Summary for Each Variable:
110 rats were a part of a research study to test the effects of a certain pesticide on weight gain. The rats were randomly selected, and divide among two groups: rats that either receive low exposure of this pesticide or high exposure of this pesticide. The trial lasted for 15 days and the weights of the rats were recorded on day 1 and day 15.

Verification of normality:
Per the central limit theorem, if the sample size is above 30 the sample mean will follow a normal distribution.

Statement of Null Hypothesis:
Mean body weight of day 15 of the low exposure group and mean body weights of day 15 of the high exposure groups are equal (H_0: $\mu_L-\mu_H=0$)

Statement of Alternative Hypothesis:
Means of body weight of day 15 low exposure group and means of body weight of day 15 high exposure groups are not equal (H_A: $\mu_L-\mu_H \neq 0$)

Statistical Method for Test:
Equal Variance or Unequal Variance Independent Two sample T-test

Statistical Method for testing for testing equality of variances:

F-test

Decision Rule:
Reject H_O in favor of H_A if p-value is less than alpha (p-value<α) otherwise fail to reject the null H_O

Method of Computation:
R statistical software version 2.11.1

Significance Level:
α=0.05

Results

Data summary of the difference in Table 3.0

Table 3.0 Data Summary: For Day 15 Body Weights of Low and High Exposure Groups

Groups	*n*	mean	SD	95% CI
Low Exposure: Day 15 Weights	49	126.58	10.32	123.6145, 129.5434
High Exposure: Day 15 Weights	61	127.79	10.72	125.0412, 130.5319
		Mean	SE	95% CI
Difference (H-L)		-1.21	2.02	-5.2170 2.8019

Normality is assumed per the CLT based on adequate sample size for each group

Assumptions:
Sample is representative
Sample is large enough
Sample measurements for each group are independent of each other
Subjects were randomized into each group, therefore we assume measurements of each group are independent of each other
Variances are equal

F-Test Results:
F-Test p-value = 0.791

Test Results of Equal Variance Two Sample Mean T-Test:
t = -0.597, df = 108, p-value = 0.5518

P-value Results
Since p-value > 0.05, we fail to reject the H_O

Description of Results:

Since the F value is higher than alpha level we assume equal variances. For the equal variance independent two sample t-test yields a p-value higher than the stated alpha level, therefore we fail to reject the null hypothesis.

Discussion

Based on the observed data, evidence suggests that the difference in mean body weight for the low exposure group on day 15 and mean body weight of high exposure group on day 15 is approximately equal to zero. Therefore, the results suggest there is no significant difference in means weights between the day 15 low exposure group and day 15 high exposure group.

Implication of Results

Based on the evidence, neither high nor low levels of this pesticide cause a difference in weight gain on day 15 of this trial. We conclude that subjects who use this pesticide should not be concerned about either low or high level exposure and weight gain, a symptom of thyroid disease.

R-code

```
#$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$question 3 $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
#is there a difference in weights between day 15 of low and high exposure groups?
group_lo_day15<-T4_Problem2[T4_Problem2$Exposure=='low' & T4_Problem2$BWDay15,2]
group_hi_day15<-T4_Problem2[T4_Problem2$Exposure=='high' & T4_Problem2$BWDay15,2]

var.test(group_lo_day15, group_hi_day15, ratio = 1,
        alternative = c("two.sided"),
        conf.level = 0.95)

#F test to compare two variances

#data:  group_lo_day15 and group_hi_day15
#F = 0.92697, num df = 48, denom df = 60, p-value = 0.791
#alternative hypothesis: true ratio of variances is not equal to 1
#95 percent confidence interval:
#  0.5432618 1.6080400
#sample estimates:
#  ratio of variances
#0.9269677

CI(group_lo_day15,ci=0.95)
CI(group_hi_day15, ci=0.95)
#upper     mean    lower
#129.5434 126.5790 123.6145
#> CI(group_hi_day15, ci=0.95)
#upper     mean    lower
#130.5319 127.7866 125.0412

#standard deviation
#Exposure  BWDay15
#     high 10.71949
#    low 10.32063

sd__day15_low<-10.32
sd__day15_hi<-10.72
n_low<-49
n_hi<-61
sp_for_group_day_15<-((n_low-1)*sd_low*sd_low+(n_hi-1)*sd_hi*sd_hi)/(n_low+n_hi-2)
#105.2646
standard_error_pooled_day_15<-sqrt(sp_for_group_day_15*(1/n_low+1/n_hi))
#1.968224



std.error(group_lo_day15)
std.error(group_hi_day15)
#[1] 1.474376
#> std.error(group_hi_day15)
#[1] 1.37249
#$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$barlett test$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
```

```
lo_Label_day15 <- rep("low", 49)
hi_label_day15 <- rep("high", 61)

TreatQ3 <- c(lo_Label_day15, hi_label_day15)
ScoreQ3 <- c(Day15_lo,Day15_hi)
bartlett.test(ScoreQ3~TreatQ3)
#Bartlett test of homogeneity of variances

#data:  ScoreQ3 by TreatQ3
#Bartlett's K-squared = 0.075737, df = 1, p-value = 0.7832
# if p-value is greater than the alpha, therefore there is no difference in variances
#$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$t sample t test$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
t.test(group_lo_day15,group_hi_day15, mu=0, alternative="two.sided", var.equal = TRUE)
#Two Sample t-test

#data:  group_lo_day15 and group_hi_day15
#t = -0.597, df = 108, p-value = 0.5518
#alternative hypothesis: true difference in means is not equal to 0
#95 percent confidence interval:
#  -5.217019  2.801863
#sample estimates:
#  mean of x mean of y
#126.5790  127.7866


#data_Q3 <- data.frame(Treatment_3 = TreatQ3, Score_3 = ScoreQ3)
#Score.aov <- aov( Score_3 ~ Treatment_3, data=data_Q3)
#summary(Score.aov, order=TRUE)
#             Df Sum Sq Mean Sq F value Pr(>F)
#Treatment_3   1     40   39.62   0.356  0.552
#Residuals   108  12007  111.18
#results: p value (0.552) is above 0.05. therefore we fail to reject the null
# no need for multiple comparions (tukey)
```