Grade sheet for Problem Two Part 2:
Two Sample Equal Variance Independent T Test For Mean Weights on Day 1 for High and Low Exposure Groups

---

Part 2: Is there a difference in weight between the two exposure groups (low and high) at Day 1? 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 1 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 1
Description of Predictor: Low or high exposure of this pesticide

Description of Data Summary:

Weights on Day 1-Low Exposure	Weights on Day 1-High Exposure
Mean: 98.01	Mean: 97.16
Sample Size: 49	Sample Size: 61
Standard Deviation: 9.57	Standard Deviation: 10.78
95% CI on Mean: 95.2623,100.7622	95% CI on Mean: 94.4011, 99.9218

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 1 of the low exposure group and mean body weights of day 1 of the high exposure groups are equal (H_0: μ_L-μ_H=0)

Statement of Alternative Hypothesis:
Means of body weight of day 1 low exposure group and means of body weight of day 1 high exposure groups are not equal (H_A: μ_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 difference in Table 2.0:

Table 2.0 Data Summary: For Day 1 Body Weights of Low and High Exposure Groups

Groups	*n*	mean	SD	95% CI
Low Exposure: Day 1 Weights	49	98.01	9.57	95.2623, 100.7622
High Exposure: Day 1 Weights	61	97.16	10.78	94.4011, 99.9218
		Mean	SE	95% CI
Difference (H-L)		0.85	1.97	-3.0508 4.7523

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 = 0.3975

Test Results of Equal Variance Two Sample Mean T-Test:
t = 0.43223, df = 108, p-value = 0.6664

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 reject 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 1 and mean body weight of high exposure group on day 1 is approximately equal to zero. Therefore, the results suggest there is no significant difference in means weights between the day 1 low exposure group and day 1 high exposure group. Meaning that, on day 1 there was no difference in weights between high exposure and low exposure.

Implication of Results

Based on the evidence, we conclude that neither high nor low levels of this pesticide cause a difference in weight gain on day 1 of this trial. We conclude that subjects who use this pesticide should not be concerned about a one time exposure for either low or high level amounts and weight gain, a symptom of thyroid disease.

st

R-code

```
#@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@question
2@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
#is there a difference in weights between low and high exposure groups at day 1?
#anova test
#assumtions: subjects are  represenative, subjects are indepedent of eachother, groups are
#represenative and indepdent, sample size is large enough
#normality is assess with sample size, per the CLT sample size is over 30, therefore we can
#assume that the sample mean will take on a normal distrubution
#Ho: there is no mean differene
#Ha: there is a mean difference
#barlett test tests for equality of variances
#anova testing difference in means: if f test yields a p value smaller than the alpha we reject the null hypothesis
group_lo<-T4_Problem2[T4_Problem2$Exposure=='low' & T4_Problem2$BWDay1,1]
group_hi<-T4_Problem2[T4_Problem2$Exposure=='high' & T4_Problem2$BWDay1,1]
#@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@summery@@@@@@@
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
CI(group_lo,ci=0.95)
CI(group_hi, ci=0.95)
#> CI(group_lo,ci=0.95)
#upper      mean     lower
#100.76221  98.01224  95.26228
#> CI(group_hi, ci=0.95)
#upper     mean    lower
#99.92182 97.16148 94.40113

#Standard deviation
#Exposure    BWDay1
#     high 10.777890
#      low  9.573986

sd_low<-9.57
sd_hi<-10.78
n_low<-49
n_hi<-61
sp_for_group<-((n_low-1)*sd_low*sd_low+(n_hi-1)*sd_hi*sd_hi)/(n_low+n_hi-2)
#105.2646
standard_error_pooled<-sqrt(sp_for_group*(1/n_low+1/n_hi))
#1.968224


std.error(group_lo,group_hi)
std.error(group_hi)
#[1] 1.367712
#> std.error(group_hi)
#[1] 1.379967


#@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@barlett
test@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
```

```
lo_Label_day1 <- rep("low", 49)
hi_label_day1 <- rep("high", 61)

TreatQ2 <- c(lo_Label_day1, hi_label_day1)
ScoreQ2 <- c(Day1_lo,Day1_hi)
bartlett.test(ScoreQ2~TreatQ2)
#Bartlett test of homogeneity of variances

#data:  ScoreT by TreatT
#Bartlett's K-squared = 0.73314, df = 1, p-value = 0.3919
# if p-value is greater than the CV there is no difference in variances
#@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@two sample t
test@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
var.test(group_lo, group_hi, ratio = 1,
         alternative = c("two.sided"),
         conf.level = 0.95)
#F test to compare two variances

#data:  group_lo and group_hi
#F = 0.78907, num df = 48, denom df = 60, p-value = 0.3975
#alternative hypothesis: true ratio of variances is not equal to 1
#95 percent confidence interval:
#  0.4624478 1.3688327
#sample estimates:
#  ratio of variances
#0.7890747

t.test(group_lo,group_hi, mu=0, alternative="two.sided", var.equal = TRUE)
#Two Sample t-test
#data:  group_lo and group_hi
#t = 0.43223, df = 108, p-value = 0.6664
#alternative hypothesis: true difference in means is not equal to 0
#95 percent confidence interval:
#  -3.050753  4.752292
#sample estimates:
#  mean of x mean of y
#98.01224  97.16148


#data_Q2 <- data.frame(Treatment_2 = TreatQ2, Score_2 = ScoreQ2)
#Score.aov <- aov( Score_2 ~ Treatment_2, data=data_Q2)
#summary(Score.aov, order=TRUE)
#            3Df Sum Sq Mean Sq F value Pr(>F)
#Treatment_2   1     20   19.67   0.187  0.666
#Residuals   108  11370  105.27

#results: p value (0.666) is above 0.05. therefore we fail to reject the null
# no need for multiple comparions (tukey)
```