

STAT 511A Homework 6

Kathleen Wendt

10/18/2019

Load packages

```
library(tidyverse)
```

Question 1

Rat data

In an investigation of the possible influence of dietary chromium on diabetic symptoms, 14 rats were fed a low-chromium diet and 10 were fed a control diet. One response variable was activity of the liver enzyme GITH. The data is available as “RatLiver.csv”.

```
rat_data <- readr::read_csv("RatLiver.csv")
```

```
## Parsed with column specification:
## cols(
##   Trt = col_character(),
##   Enzyme = col_double()
## )
```

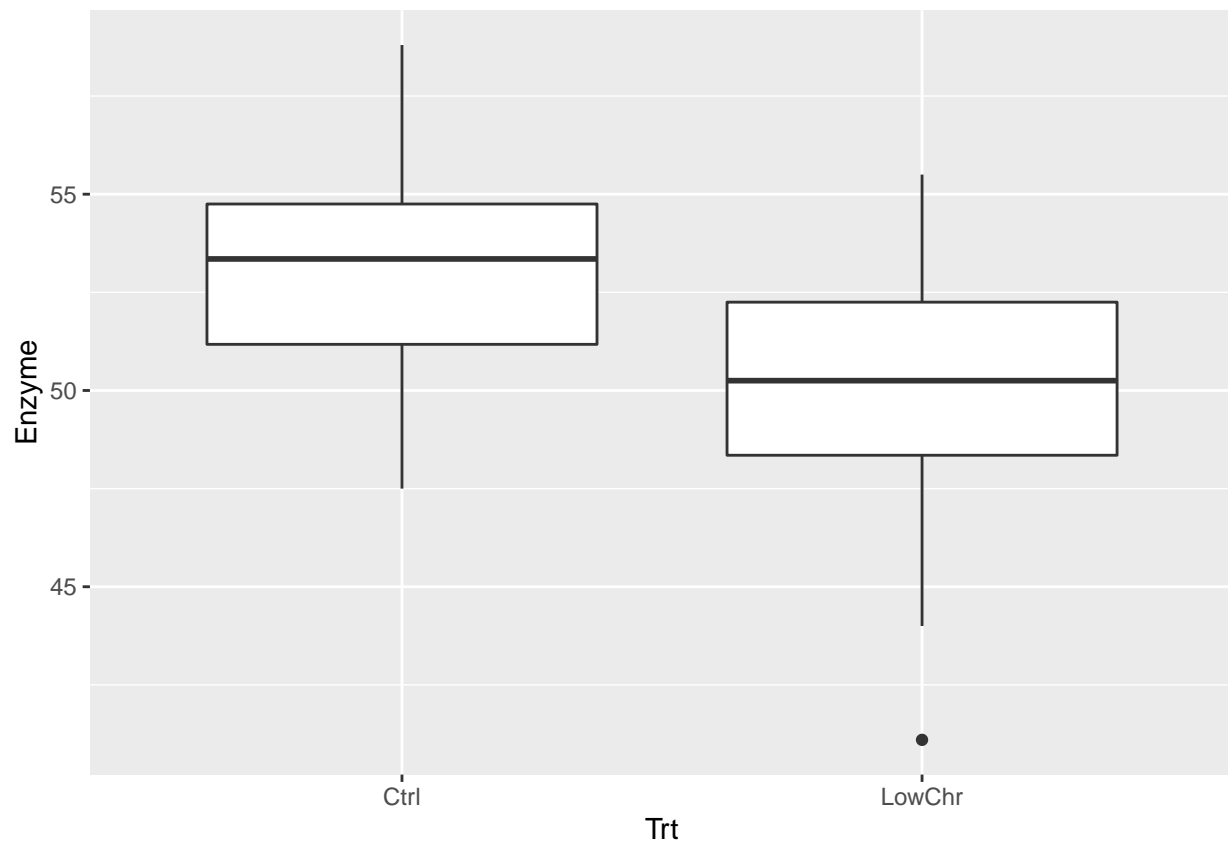
```
tibble::glimpse(rat_data)
```

```
## Observations: 24
## Variables: 2
## $ Trt      <chr> "LowChr", "LowChr", "LowChr", "LowChr", "LowChr", "LowC...
## $ Enzyme <dbl> 44.0, 48.5, 50.7, 45.0, 53.0, 52.7, 51.8, 49.8, 48.3, 5...
```

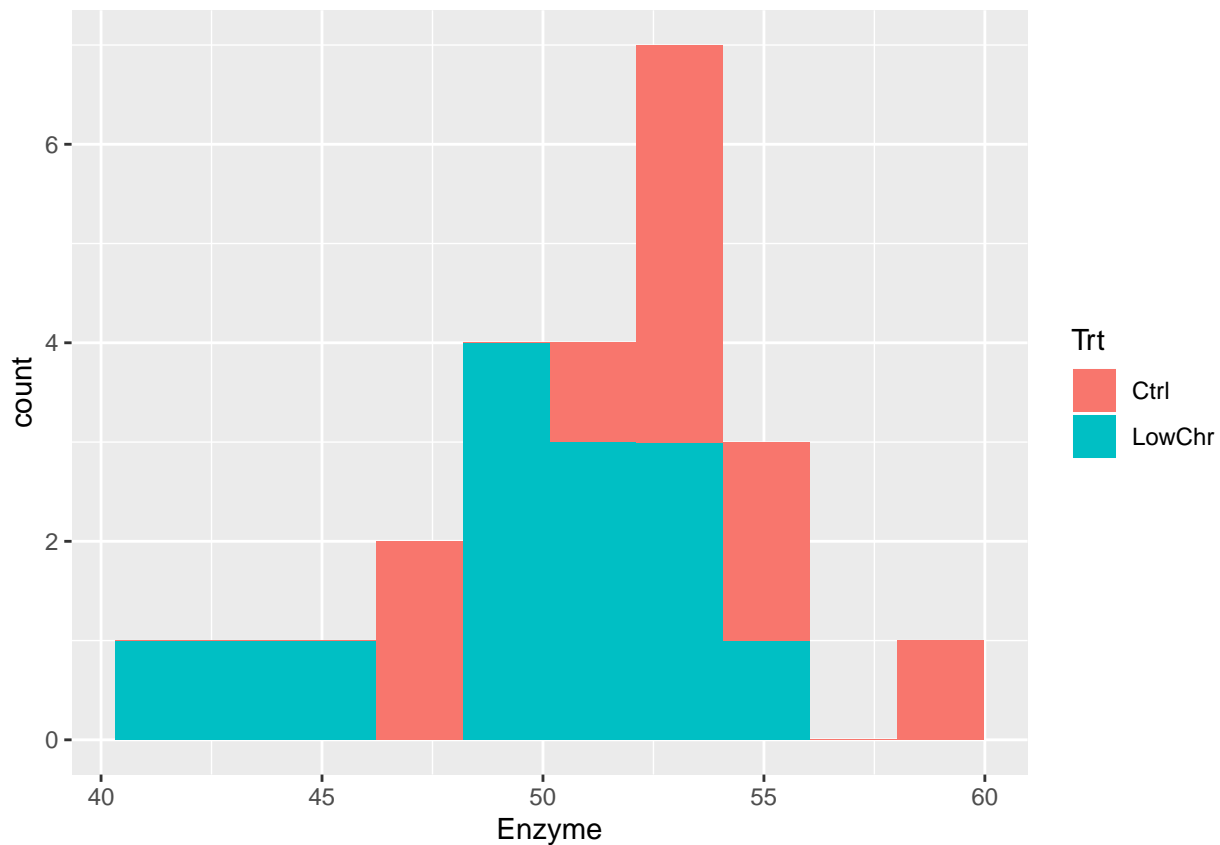
Part 1A

Construct side-by-side boxplots of the data.

```
rat_data %>%
  ggplot(aes(y = Enzyme, x = Trt)) +
  geom_boxplot()
```



```
rat_data %>%  
  ggplot(aes(Enzyme, fill = Trt)) +  
  geom_histogram(bins = 10)
```



Part 1B

Use the F-test to test for equality of variances. Give the null hypothesis, test statistic, p-value and conclusion. (4 pts)

Null hypothesis

Equal variances; true ratio of variances is equal to 1.

F-test for equality of variances

```
rat_ftest <- broom::tidy(var.test(Enzyme ~ Trt, data = rat_data))
```

```
## Multiple parameters; naming those columns num.df, denom.df
```

```
rat_ftest
```

```
## # A tibble: 1 x 9
```

```
##   estimate `num df` `denom df` statistic p.value conf.low conf.high method
##   <dbl>     <int>     <int>     <dbl>   <dbl>    <dbl>    <dbl> <chr>
## 1    0.790         9        13    0.790  0.737    0.238    3.03 F tes~
## # ... with 1 more variable: alternative <chr>
```

Test statistic

The test statistic for the F-test for equality of variances is 0.7897775.

P-value

The p-value for the F-test for equality of variances is 0.7373033.

Conclusion

Fail to reject null hypothesis. Because the above p-value is above $\alpha = 0.05$, there is no evidence to suggest that the true ratio of variances is not equal to 1.

Part 1C

Use Levene's test (with center="median") to test for equality of variances. Give the p-value and conclusion.

Levene's Test

```
rat_levene <- broom::tidy(car::leveneTest(Enzyme ~ Trt, data = rat_data,
                                         center = "median"))

## Warning in leveneTest.default(y = y, group = group, ...): group coerced to
## factor.

rat_levene

## # A tibble: 2 x 4
##   term      df statistic p.value
##   <chr> <int>    <dbl>   <dbl>
## 1 group      1     0.176   0.679
## 2 ""        22      NA      NA
```

P-Value

The p-value for the Levene Test for equality of variances with median as center is 0.6788657, NA.

Conclusion

Fail to reject null hypothesis. Because the above p-value is above $\alpha = 0.05$, there is no evidence to suggest that the true ratio of variances is not equal to 1.

Part 1D

Based on your conclusions from the two previous questions, would the pooled variance t-test or Welch-Satterthwaite t-test be preferred?

Pooled t-test

Part 1E

Regardless of your answer to the previous question, run a two-sample t-test assuming equal variances. Give the null hypothesis, test statistic, p-value and conclusion. (4 pts)

Null hypothesis

$\mu_1 = \mu_2$. No difference between sample means.

Pooled two-sample t-test

```
rat_ttest <- broom::tidy(t.test(Enzyme ~ Trt,
                               data = rat_data,
                               var.equal = TRUE))

rat_ttest

## # A tibble: 1 x 9
##   estimate1 estimate2 statistic p.value parameter conf.low conf.high method
##   <dbl>      <dbl>      <dbl>  <dbl>      <dbl>      <dbl>      <dbl> <chr>
## 1      52.9      49.5       2.17  0.0410         22      0.151       6.59 " Two~
## # ... with 1 more variable: alternative <chr>
```

Test statistic

The test statistic for the pooled two-sample t-test of means is 2.1708862.

P-value

The p-value for the pooled two-sample t-test of means is 0.041005.

Conclusion

Reject null hypothesis: $0.041005 < \alpha = 0.05$ level. There is evidence to suggest that there is a difference between sample means, such that the mean level of enzyme in Control rats is higher than those in the Treatment group.

Part 1F

Rerun the analysis as a one-way ANOVA. Give the ANOVA table in your assignment. Compare your results to the previous question and notice that the p-value is the same and $F = t^2$.

Question 2

Read Problem 8.32 which concerns corn yield. The data is available as "CornYield.csv".

Part 2A

Construct a bar plot showing means and SEs for each variety. (4 pts)

Part 2B

Carry out a one-way ANOVA analysis to determine whether there is evidence of differences (using alpha 0.05) in the mean yield for the different varieties. State the null hypothesis, give the test statistic, p-value and conclusion. (4 pts)

Part 2C

Run (unadjusted) pairwise comparisons of means. Give the estimated difference and p-value for each comparison. (4 pts)

Part 2D

Calculate the $LSD(0.05)$ value. Recall that this is the 95% ME for pairwise comparisons of means.

Part 2E

Construct an (unadjusted) “cld” display including the mean for each variety and assigning number groups (or underlining) varieties that are not “significantly” different. (4 pts)

Part 2F

Summarize your findings from parts C and E.

Part 2G

G. Use the `plot()` function to generate the diagnostic plots from the model from part B. You do not have to include the graphs in your assignment, but discuss the plots of (1) Residuals vs Fitted values and (2) qqplot of residuals and whether assumptions appear to be satisfied based on each plot. (4 pts)