Lab 05- T-tests

YOUR NAME HERE

XX/XX/20XX

Table of Contents

# Chunk 1: Load libraries and datasets

## 1.1) Load tidyverse ## 1.2) read in two datasets “WormSperm.csv” and “BlackbirdTesto.csv” ## 1.3) What are the structures of each of these two datasets?

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ ggplot2 3.3.5 ✓ purrr 0.3.4  
## ✓ tibble 3.1.5 ✓ dplyr 1.0.7  
## ✓ tidyr 1.1.4 ✓ stringr 1.4.0  
## ✓ readr 2.0.2 ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.1.2

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

## The following object is masked from 'package:purrr':  
##   
## some

## Rows: 913 Columns: 2

## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (1): Sex  
## dbl (1): SpermSize

##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

## Rows: 26 Columns: 3

## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (1): Treatment  
## dbl (2): blackbird, Antibody.production

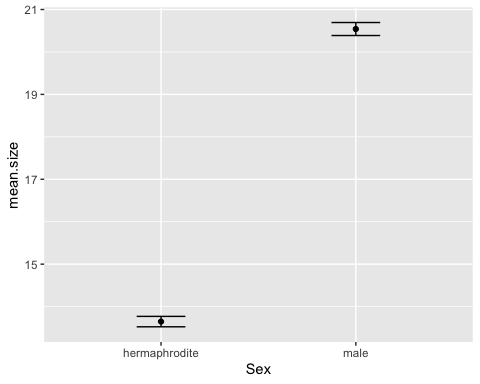
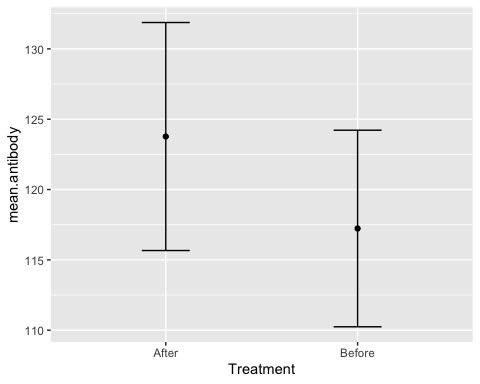
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

## spec\_tbl\_df [913 × 2] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ Sex : chr [1:913] "hermaphrodite" "hermaphrodite" "hermaphrodite" "hermaphrodite" ...  
## $ SpermSize: num [1:913] 10 10 10 10 10 10 10 10 10 10 ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. Sex = col\_character(),  
## .. SpermSize = col\_double()  
## .. )  
## - attr(\*, "problems")=<externalptr>

# Chunk 2: Define Hypotheses and Plot the data!

2.1) Using pipes, summarize the bird and sperm datasets so that you have a mean and standard error per group (treatment in birds and sex in sperm). 2.2) Identify a null and alternative hypotheses testing for differences in means in each of these datasets 2.3) Plot summarized data as a point plot with error bars representing the standard error

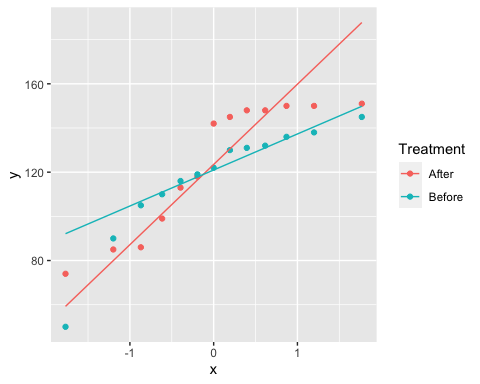
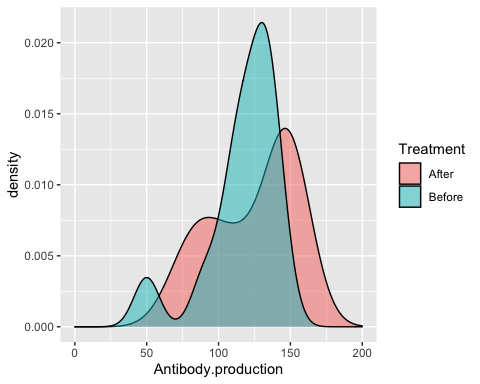
## # A tibble: 2 × 3  
## Treatment mean.antibody se.antibody  
## <chr> <dbl> <dbl>  
## 1 After 124. 8.10  
## 2 Before 117. 6.99



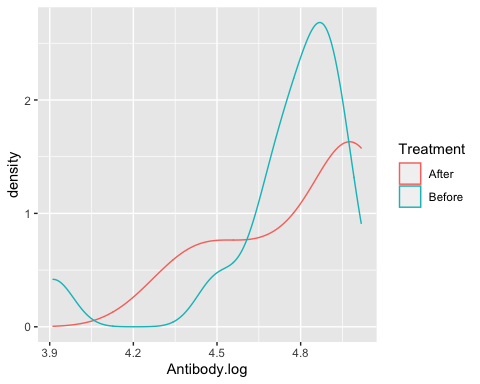
# Chunk 3: Check statistical assumptions

3.1) Using a shapiro test, a density plot and a qq plot check the two datasets for normality 3.1b) If any dataset is not normally distributed attempt a transformation to normalize the data. Does the transformation improve the normality of the distribution? 3.2) Using both a Levene Test and a Flinger-Kileen test, check if the datasets meet the assumption of homeogeneity of variances 3.3) What about the assumption of independence, are all observations in each of the datasets independent of each other?

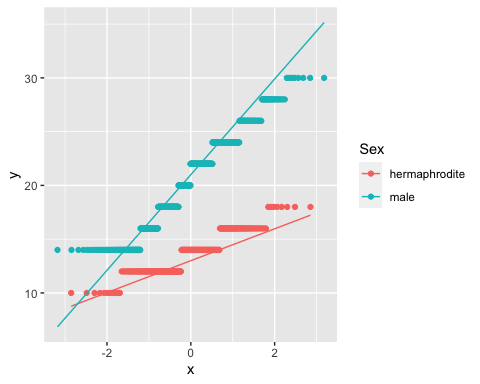
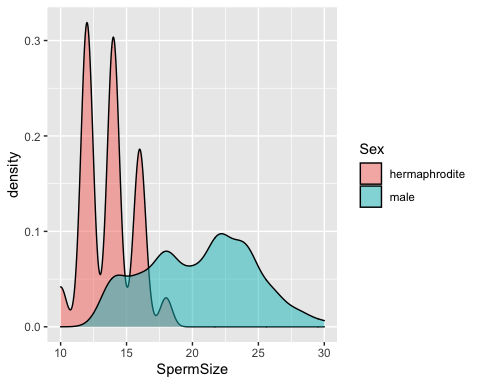
##   
## Shapiro-Wilk normality test  
##   
## data: bbirds$Antibody.production  
## W = 0.91336, p-value = 0.0315



##   
## Shapiro-Wilk normality test  
##   
## data: bbirds$Antibody.log  
## W = 0.84557, p-value = 0.001169



##   
## Shapiro-Wilk normality test  
##   
## data: sperm$SpermSize  
## W = 0.95064, p-value < 2.2e-16



##   
## Shapiro-Wilk normality test  
##   
## data: sperm$SpermSize  
## W = 0.95064, p-value < 2.2e-16

## Warning in leveneTest.default(bbirds$Antibody.production, group =  
## bbirds$Treatment): bbirds$Treatment coerced to factor.

## Levene's Test for Homogeneity of Variance (center = median)  
## Df F value Pr(>F)  
## group 1 0.7604 0.3918  
## 24

## Warning in leveneTest.default(sperm$SpermSize, group = sperm$Sex): sperm$Sex  
## coerced to factor.

## Levene's Test for Homogeneity of Variance (center = median)  
## Df F value Pr(>F)   
## group 1 120.51 < 2.2e-16 \*\*\*  
## 911   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Fligner-Killeen test of homogeneity of variances  
##   
## data: Antibody.production by Treatment  
## Fligner-Killeen:med chi-squared = 0.1087, df = 1, p-value = 0.7416

##   
## Fligner-Killeen test of homogeneity of variances  
##   
## data: SpermSize by Sex  
## Fligner-Killeen:med chi-squared = 108.86, df = 1, p-value < 2.2e-16

# Chunk 4: Compelte an unpaired t-test

4.1) Complete a t.test for both the transformed and untranfromed Antibody production predicted by Treatment in the bbirds dataset. Set the var.equal=TRUE

4.2) Complete a t.test for both the transformed and untranfromed SpermSize predicted by Sex in the sperm dataset. Set the var.equal=FALSE

4.3) Interpret the results in your own words.

##   
## Welch Two Sample t-test  
##   
## data: Antibody.production by Treatment  
## t = 0.61107, df = 23.491, p-value = 0.547  
## alternative hypothesis: true difference in means between group After and group Before is not equal to 0  
## 95 percent confidence interval:  
## -15.57069 28.64761  
## sample estimates:  
## mean in group After mean in group Before   
## 123.7692 117.2308

##   
## Two Sample t-test  
##   
## data: Antibody.log by Treatment  
## t = 0.52013, df = 24, p-value = 0.6077  
## alternative hypothesis: true difference in means between group After and group Before is not equal to 0  
## 95 percent confidence interval:  
## -0.1631528 0.2730924  
## sample estimates:  
## mean in group After mean in group Before   
## 4.789215 4.734245

##   
## Welch Two Sample t-test  
##   
## data: SpermSize by Sex  
## t = -35.124, df = 828.43, p-value < 2.2e-16  
## alternative hypothesis: true difference in means between group hermaphrodite and group male is not equal to 0  
## 95 percent confidence interval:  
## -7.279072 -6.508588  
## sample estimates:  
## mean in group hermaphrodite mean in group male   
## 13.64655 20.54038

##   
## Welch Two Sample t-test  
##   
## data: log.size by Sex  
## t = -33.64, df = 587.07, p-value < 2.2e-16  
## alternative hypothesis: true difference in means between group hermaphrodite and group male is not equal to 0  
## 95 percent confidence interval:  
## -0.4217956 -0.3752614  
## sample estimates:  
## mean in group hermaphrodite mean in group male   
## 2.604159 3.002688

# Chunk 5: Complete a paired t-test and plot a paired point plot

5.1) Complete a paired t test for the bbirds dataset. use paired=TRUE

5.2) Interpret the results and discuss Why is this the appropriate test for these data?

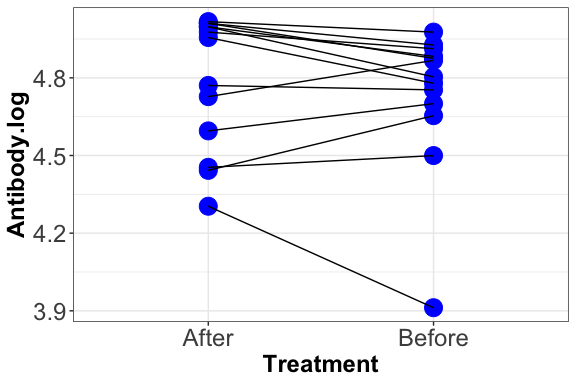
5.3) Plot the data as paired point plot

## spec\_tbl\_df [26 × 4] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ blackbird : num [1:26] 1 2 3 4 5 6 7 8 9 10 ...  
## $ Antibody.production: num [1:26] 105 50 136 90 122 132 131 119 145 130 ...  
## $ Treatment : chr [1:26] "Before" "Before" "Before" "Before" ...  
## $ Antibody.log : num [1:26] 4.65 3.91 4.91 4.5 4.8 ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. blackbird = col\_double(),  
## .. Antibody.production = col\_double(),  
## .. Treatment = col\_character()  
## .. )  
## - attr(\*, "problems")=<externalptr>

##   
## Paired t-test  
##   
## data: Antibody.production by Treatment  
## t = 1.5079, df = 12, p-value = 0.1575  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -2.90934 15.98626  
## sample estimates:  
## mean of the differences   
## 6.538462

##   
## Paired t-test  
##   
## data: Antibody.log by Treatment  
## t = 1.2435, df = 12, p-value = 0.2374  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.04134676 0.15128638  
## sample estimates:  
## mean of the differences   
## 0.05496981

## Warning: Ignoring unknown parameters: levels

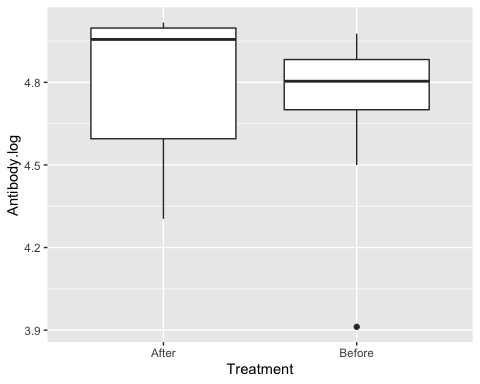


# Chunk 6: Non-parametric alternative

6.1) Complete a wilcoxon rank sums test of Antibody.log predicted by treatment with equal variances and paired data. Interpret the results in your own words.

6.2) Using the sperm data, complte a wilcoxon rank sums test of log.sperm size predicted by sex with unequal variances. Interpret the results in your own words.

##   
## Wilcoxon signed rank exact test  
##   
## data: Antibody.log by Treatment  
## V = 61, p-value = 0.3054  
## alternative hypothesis: true location shift is not equal to 0



##   
## Wilcoxon rank sum test with continuity correction  
##   
## data: log.size by Sex  
## W = 10297, p-value < 2.2e-16  
## alternative hypothesis: true location shift is not equal to 0