

# PS\_Independent and Dependent t-test-for-TAs

Pablo E. Gutiérrez-Fonseca

2023-08-11

R practice.

Install packages.

```
#install.packages("car")  
library(car)
```

## Loading required package: carData

```
library(ggplot2)
```

Load the water pollution data into R.

```
##   Birds   Species Gender Pre_weight Pre_weight.1  
## 1     1 Hylocichla     M         46          43  
## 2     2 Hylocichla     M         44          42  
## 3     3 Hylocichla     M         48          41  
## 4     4 Hylocichla     M         42          49  
## 5     5 Hylocichla     M         51          43  
## 6     6 Hylocichla     M         51          42
```

Part I.

2. Analyze data distribution. Is your pre-weight data normally distributed?

```
shapiro.test(df_bird$Pre_weight)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  df_bird$Pre_weight  
## W = 0.95557, p-value = 0.2725
```

3. Test to see if you meet the assumption of equal variance for this test. Based on the results of this test for unequal variance, which test should you use?

```
# Perform Levene's test  
result <- leveneTest(Pre_weight~Gender, data=df_bird)
```

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

```
print(result)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 1  0.0061 0.9381
##      26
```

```
p_value <- result$Pr[1]
# Significance level
alpha <- 0.05

# Interpret the result
if (p_value < alpha) {
  cat("Reject null hypothesis: Variances are significantly different.\n")
} else {
  cat("Fail to reject null hypothesis: Variances are not significantly different.\n")
}
```

```
## Fail to reject null hypothesis: Variances are not significantly different.
```

#### 4. Run this analysis. What is the test statistic (obtained) value for the test you selected?

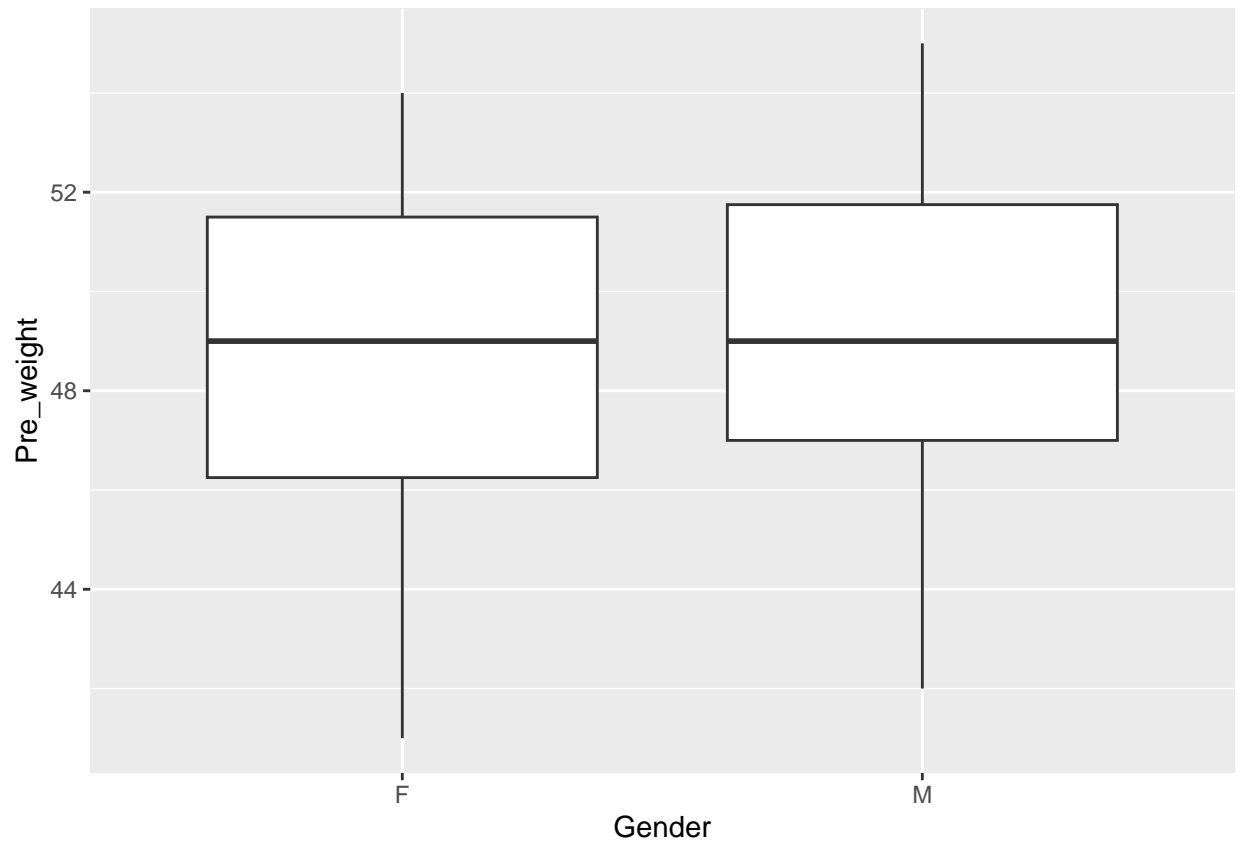
Use alternative to determine your hypothesis. Default= "two.sided".

alternative = c("two.sided", "less", "greater"),

```
# Perform t-test
t_test_result <- t.test(Pre_weight~Gender, data=df_bird,
                        alternative = "less")
print(t_test_result)
```

```
##
## Welch Two Sample t-test
##
## data: Pre_weight by Gender
## t = -0.78654, df = 25.944, p-value = 0.2193
## alternative hypothesis: true difference in means between group F and group M is less than 0
## 95 percent confidence interval:
##      -Inf 1.419121
## sample estimates:
## mean in group F mean in group M
##      48.07143      49.28571
```

```
# Create a boxplot plot using ggplot2
ggplot(df_bird, aes(x = Gender, y = Pre_weight)) +
  geom_boxplot() # Add jitter for better visualization
```



## Part II.

Analyze data distribution. Is your pre-weight data normally distributed?

```
shapiro.test(df_bird$Pre_weight)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  df_bird$Pre_weight
## W = 0.95557, p-value = 0.2725
```

Test to see if you meet the assumption of equal variance for this test. Based on the results of this test for unequal variance, which test should you use?

```
# Perform Levene's test
result <- leveneTest(Pre_weight~Species, data=df_bird)
```

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

```
print(result)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 1  1.2128 0.2809
##      26
```

```

p_value <- result$Pr[1]
# Significance level
alpha <- 0.05

# Interpret the result
if (p_value < alpha) {
  cat("Reject null hypothesis: Variances are significantly different.\n")
} else {
  cat("Fail to reject null hypothesis: Variances are not significantly different.\n")
}

```

## Fail to reject null hypothesis: Variances are not significantly different.

## 8. Run this analysis. What is the test statistic (obtained) value for the test you selected?

Use alternative to determine your hypothesis. Default= "two.sided".

alternative = c("two.sided", "less", "greater"),

```

# Perform t-test
t_test_result <- t.test(Pre_weight~Species, data=df_bird,
                        alternative = "two.sided")
print(t_test_result)

```

```

##
## Welch Two Sample t-test
##
## data: Pre_weight by Species
## t = -2.891, df = 24.273, p-value = 0.007972
## alternative hypothesis: true difference in means between group Hylocichla and group Progne is not equal to 0
## 95 percent confidence interval:
## -6.731487 -1.125656
## sample estimates:
## mean in group Hylocichla      mean in group Progne
##           46.71429           50.64286

```

```

# Create a boxplot plot using ggplot2
ggplot(df_bird, aes(x = Species, y = Pre_weight)) +
  geom_boxplot() # Add jitter for better visualization

```

