Parametric Methods in R

Samantha Estrada PhD

ORS Research Design & Data Analysis Lab Consultant
University of Texas at Tyler

1/19/2024

Parametric vs. Non-Parametric

- Parametric methods. Parametric methods usually involved data expressed in absolute numbers (as opposed to ranks for example).
  - Meet the assumption of normality.
  - Parametric statistics are based on assumptions about the distribution of population from which the sample was taken.

- Non-parametric statistics are not based on assumptions, that is, the data can be collected from a sample that does not follow a specific distribution.
  - Upcoming workshop!

| INDEPENDENT VARIABLE			DEPENDENT VARIABLE (CONTINUOUS)	
			Non-Parametric	Parametric
	C A T E G O R I C A L	1 Variable 2 Categories	Mann-Whitney U Test	Independent Samples t-test
		1 Variable 2 Categories	Wilcoxon Signed Rank Test	Paired t-test
		1 Variable > 2 Categories	Kruskal-Wallis H Test	One-Way ANOVA
		1 Variable > 2 Categories	Friedman Test	Repeated Measures
		1 Variable	Spearman rho	Pearson's r

Why R ?

- Free and open source software.

  - *Free as in freedom.*

  - *Open source.*

- Download: https://www.r-project.org/

- Recommended: https://posit.co/download/rstudio-desktop/

file:///Users/sestrada/Library/CloudStorage/OneDrive-UniversityofTexasatTyler/Service/ORS/Parametric/Non Parametric in R/Parametric-in-R.html#(1)

3/22

Independent vs. Dependent Samples

- When testing the difference between two means, we must first determine if the samples are independent or dependent.

- Independent samples

  - *Two samples are independent when the subjects selected for the first sample in no way influence how the subjects are selected for the second sample.*

  - *There is no relationship between the subjects in each sample*

  - *Undergraduates vs Graduates*

  - *Residents of Texas vs New Mexico*

Independent Samples
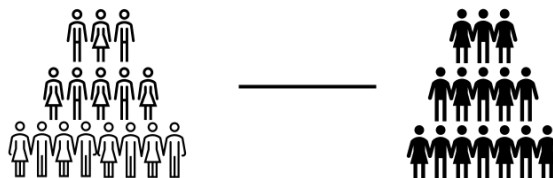
Dependent Samples

- Dependent samples

  - *Each subject in one sample has a corresponding matched subject in the other sample*

  - *E.g., before & after, pre-test vs. post-test*

  - *Statistics knowledge*
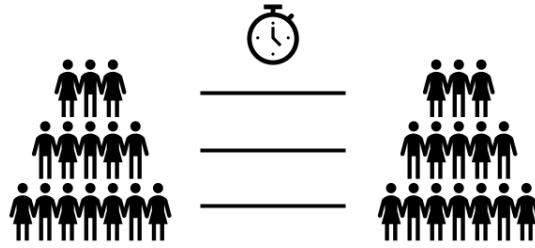
Assumptions

Normality

It is assumed that the data are normally distributed. Specifically, we assume that both groups are normally distributed. We check this with a test called Shapiro Wilk

Shapiro Wilk
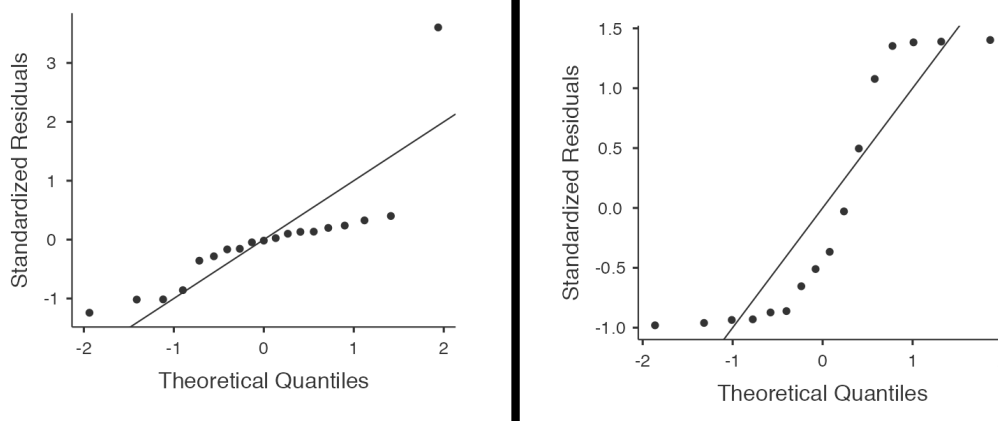
- Shapiro-Wilk is a test of normality (there are many others!).
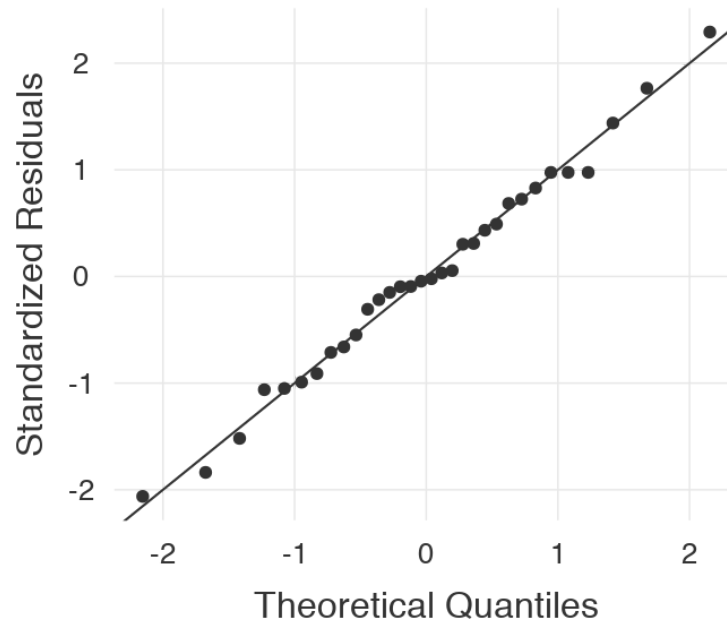
The hypotheses:

- Ho: The data are normally distributed

- Ha: The data are not normally distributed

QQ Plots

Examples of QQ-Plots when the assumption of normality is violated

Example of QQ-Plot when the normality assumption is met

Assumption-shumptions!

⚠️

- What happens if I don't meet the normality assumption?

- In the case where you don't have continuous data, you have ranked data, you can use the Mann-Whitney Test.

Independence

It is assumed that the observations are independently sampled. Firstly, we assume that the observations within each sample are independent of one another (exactly the same as for the one-sample test). However, we also assume that there are no cross-sample dependencies.

- Some researchers use the intra-class correlation coefficient (ICC) to determine independence.

Equality of variance

- Homogeneity of variance (also called "homoscedasticity") or simply equality of variance.

- One of the assumptions of parametric statistical tests is that the samples you are comparing should have fairly homogeneous variances.

Levene's Test for Equality of Variances

You can test this assumption using the Levene test. You want it to NOT be significant, if it is not significant you should use Welch's t-test (a non-parametric test).

- The hypotheses for Levene's Test

Ho: The variances are equal
Ha: The variances are not equal

Assumption-shumptions!

⚠️

- What happens if I don't meet the assumption of equal variance?

- Welch's test!

Resources for Assumptions

- Nimon (2012)

Indepedent Samples t-test

Mann-Whitney U

The Mann-Whitney U test is used to compare whether there is a difference in the dependent variable for two independent groups (Navarro & Foxcroft, 2019)

Hypotheses

- Ho: The Median (Mdn) from the first population is NOT different from the Median (Mdn) of the second population

- Ha: The Median (Mdn) from the first population is different from the Median (Mdn) of the second population

Assumptions

It doesn't have any! How great is that?!

Example

We will compare a non-smoker group to a group of smokers on their lung capacity.

- 👥 Smoke: Smoking status

- 📏 LungCapacity

Performing Mann-Whitney U in `jamovi`

Let's start with Descriptive information and visualizing our data.

```
summary(NonParametricData$LungCap)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.507   6.150   8.000   7.863   9.800  14.675
```

```
table(NonParametricData$Smoke)
```

```
##
##  no yes
## 648  77
```

```
library(dplyr)
group_by(NonParametricData,Smoke) %>%
  summarise(
    count = n(),
    median = median(LungCap, na.rm = TRUE),
    IQR = IQR(LungCap, na.rm = TRUE))
```

```
## # A tibble: 2 × 4
##   Smoke count median   IQR
##   <chr> <int>  <dbl> <dbl>
## 1 no      648    7.9  3.73
## 2 yes      77   8.65  2.78
```

Let's visually examine our data:
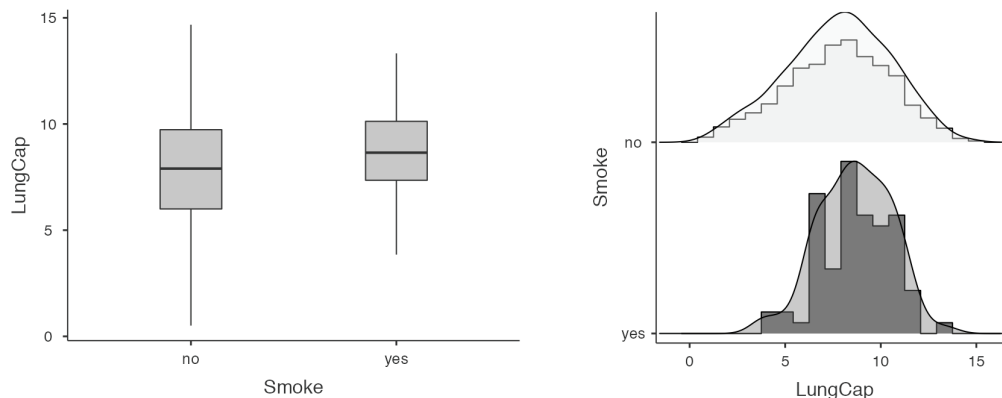
The code to perform Shapiro-Wilk is as follows

```
shapiro.test(continuousvariable)
```

Now check the normality assumption using the Shapiro-Wilk
procedure in R:

```
shapiro.test(NonParametricData[,1])
```

```
##
##  Shapiro-Wilk normality test
```

```
##
## data:  NonParametricData[, 1]
## W = 0.99305, p-value = 0.001886
```

- Recall a p-value < than α means we reject the null hypothesis.

- The null hypothesis here is that our data is normally distributed. We reject this based on the results of our test. We do not meet the assumption.

- Next we should conduct a Mann-Whitney U test.

The code to perform the procedure:

```
# DV: dependent variable
# IV: Independent variable

wilcox.test(DV~IV, data = dataset, exact = FALSE)
```

Data Analysis

```
wilcox.test(LungCap~ Smoke, data = NonParametricData, exact = FALSE)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  LungCap by Smoke
## W = 20128, p-value = 0.005538
## alternative hypothesis: true location shift is not equal to 0
```

Interpretation

Mann-Whitney U test showed that there was a significant difference ($\mathrm{W} = 20{,}128$, $\mathrm{p} = 0.006$) in lung capacity between the group of smokers ($\mathrm{Mdn} = 8.65$) compared to the non-smoker group ($\mathrm{Mdn} = 7.90$).

Welch t-test

Welch's t-test assumes that both groups of data are normally distributed, but it does not assume that those two populations have the same variance.

Hypotheses

- Ho: The means are equal

- Ha: The means are not equal

Example

A teacher wants to compare the exam scores of 12 students who used an exam prep booklet to prepare for some exam vs. 12 students who did not (*How to Perform Welch's t-Test in r*, 2020)

- 📏 Booklet: Exam score

- 👥Booklet Group: What type of prep booklet they used. Dichotomous variable where 1 = Used Booklet and 2 = Did not use Booklet

Descriptives

Let's start with Descriptive information and visualizing our data.

```
summary(NonParametricData$Booklet)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   66.00   82.00   87.50   84.38   90.00   95.00
```

```
table(NonParametricData$Booklet.Group)
```

```
##
##  1  2
## 12 12
```

```
library(dplyr)
group_by(NonParametricData,Booklet.Group) %>%
  summarise(
    count = n(),
    median = median(Booklet, na.rm = TRUE),
    IQR = IQR(Booklet, na.rm = TRUE))
```

```
## # A tibble: 2 × 4
##   Booklet.Group count median   IQR
##           <int> <int>  <dbl> <dbl>
## 1             1    12   88.5  3.75
## 2             2    12   83.5  16.8
```

Let's check our assumption of equality of variance:

The code for Levene's Test

The function `leveneTest()` in the `car` package can be used.

```
# DV: dependent variable
# IV: Independent variable
library(car)
leveneTest(DV ~ IV, data = dataset)
```

Now let's perform Levene's test on our variables:

```
library(car)
# Levene's test
leveneTest(Booklet~as.factor(Booklet.Group), data = NonParametricData)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value   Pr(>F)
## group  1  9.2487 0.005993 **
##       22
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As we can see in the `Assumptions` output, we are violating the equal variances assumption ($F(1, 22) = 9.25, p < .001$)

Data Analysis

```
t.test(Booklet~Booklet.Group, data=NonParametricData)
```
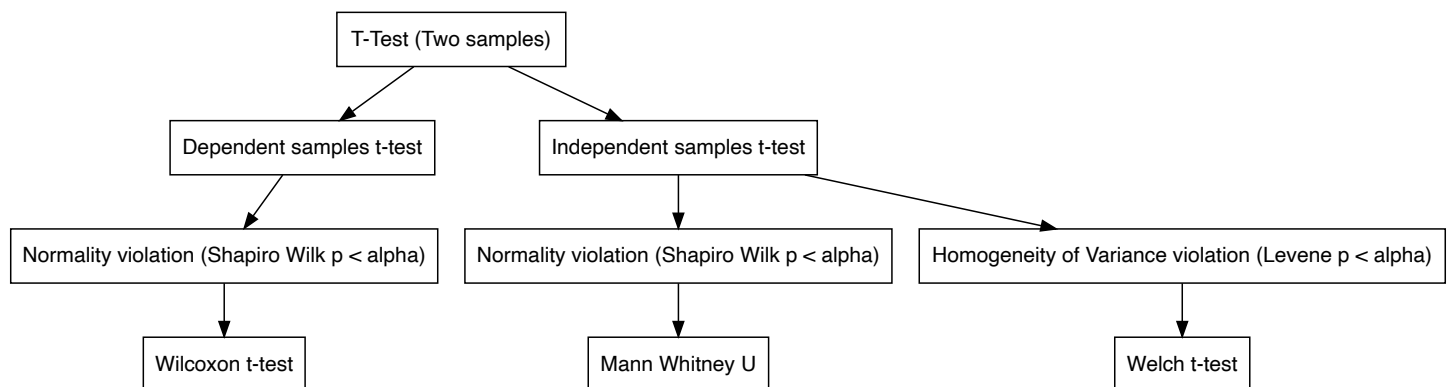
```
##
##  Welch Two Sample t-test
##
## data:  Booklet by Booklet.Group
## t = 2.2361, df = 14.354, p-value = 0.04171
## alternative hypothesis: true difference in means between group 1 and group
2 is not equal to 0
## 95 percent confidence interval:
##   0.3048395 13.8618272
## sample estimates:
```

```
## mean in group 1 mean in group 2
##          87.91667        80.83333
```

Interpretation

The assumption of equal variances was checked using Levene's Test (F (1,22) = 16.6, p < .001 and we found the assumption was not met. Due to the violation of the equal variance assumption a Welch's T-Test was performed. We found there was a statistically significant difference ($Welch's\ t$ (14.4) = 2.24, p = 0.042). Next… report the means and/or mean difference.

Decision Chart

```
                        ┌─────────────────────┐
                        │  T-Test (Two samples) │
                        └─────────────────────┘
                         ╱                       ╲
          ┌──────────────────────┐      ┌─────────────────────────┐
          │ Dependent samples t-test │      │ Independent samples t-test │
          └──────────────────────┘      └─────────────────────────┘
                    │                      │                        ╲
┌──────────────────────────────────┐  ┌──────────────────────────────────┐  ┌─────────────────────────────────────────────┐
│ Normality violation (Shapiro Wilk p < alpha) │  │ Normality violation (Shapiro Wilk p < alpha) │  │ Homogeneity of Variance violation (Levene p < alpha) │
└──────────────────────────────────┘  └──────────────────────────────────┘  └─────────────────────────────────────────────┘
                    │                      │                                          │
          ┌──────────────────┐      ┌──────────────────┐                    ┌──────────────────┐
          │  Wilcoxon t-test  │      │  Mann Whitney U   │                    │   Welch t-test   │
          └──────────────────┘      └──────────────────┘                    └──────────────────┘
```

ANOVA

Kruskal-Wallis ANOVA

- Kruskal-Wallis test used for comparing the differences between two or more groups (similar to ANOVA).

- It is an extension to the Mann-Whitney U Test (in the same way a ANOVA is an extension of a t-test), which is used for comparing two groups. It compares the mean ranks (medians) of groups.

- Kruskal-Wallis test does not assume any specific distribution (such as normal distribution of samples). Thus is useful if you have a normality violation.

- Kruskal-Wallis test is not as efficient as the F test, so it might require more dramatic differences for the null hypothesis to be rejected.

Example

For example, suppose weights of poplar trees are different based on treatments (Control, Treatment 1, Treatment 2). But the weight samples are not normally distributed. The research question is to test whether the poplar tree weights are different under the four treatments.

- 📏 `Continuous DV:` `Weight` of the samples.

- 👥 `Nominal IV:` Three different types of fertilzing:

  - *0 = Control*

  - *1 = Treatment 1*

  - *2 = Treatment 2*

Hypotheses

- Ho: Populations medians are equal

- Ha: At least, one population median is different from other populations

Descriptives

```
# Descriptives
library(dplyr)
group_by(NonParametricData,Plant.Growth.Group) %>%
  summarise(
    count = n(),
    median = median(Weight, na.rm = TRUE),
    IQR = IQR(Weight, na.rm = TRUE))
```

```
## # A tibble: 3 × 4
##   Plant.Growth.Group count median    IQR
##                <int> <int>  <dbl>  <dbl>
## 1                  0    10   5.15  0.743
## 2                  1    10   4.55  0.662
## 3                  2    10   5.44  0.467
```

Assumptions

```
shapiro.test(NonParametricData$Weight)
```

Assume the p-value for this test was $p < \alpha = 0.05$, indicating we violate the assumption of normality.

Data Analysis

Next, we will try Kruskal-Wallis, to address the lack of normality in our data.

The code to perform Kruskall-Wallis is as follows:

```
kruskal.test(DV ~ IV, data = data)
```

Now let's conduct the analysis:

```
kruskal.test(Weight ~ Plant.Growth.Group, data = NonParametricData)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  Weight by Plant.Growth.Group
## Kruskal-Wallis chi-squared = 7.9882, df = 2, p-value = 0.01842
```

Post Hoc

The code to conduct the post hoc test for Kruskall-Wallis ANOVA:

```
pairwise.wilcox.test(DV, IV, p.adjust.method = "bonferroni")
```

Let's conduct our post hoc:

```
pairwise.wilcox.test(NonParametricData$Weight, NonParametricData$Plant.Growth.Group,
                     p.adjust.method = "bonferroni")
```

```
##
##  Pairwise comparisons using Wilcoxon rank sum test with continuity
correction
##
## data:  NonParametricData$Weight and NonParametricData$Plant.Growth.Group
##
##   0     1
## 1 0.596 -
## 2 0.189 0.027
##
## P value adjustment method: bonferroni
```

What we see in our output:

- When we compare the Control group (labeled 0) to Treatment 1 (labeled 1) there is no statistical significance 0.596 our p-value is greater than α = 0.05

- When we compare the Control group (labeled 0) to Treatment 2 (labeled 2) there is no statistical significance 0.189 our p-value is greater than α = 0.05

- Finally! When we compare the Treatment 1 group (labeled 1) to Treatment 2 (labeled 2) there IS statistical significance 0.027 our p-value is less than α = 0.05

Interpretation

A Kruskal-Wallis test was conducted to examine the differences between three different treatment used for plant grow. The results indicated that there were statistically significant differences χ^2 (2) = 7.99, p = 0.018. Next, you describe your post hoc analyses

Welch's ANOVA

We used Welch's ANOVA when we have a violation of homogeneity of variance.

Hypotheses

Same as ANOVA.

Example

Three reading instructions are given to 15 subjects; then a reading test is given where the number of words per minute is recorded for each subject. The question is to test whether the three instructions makes any difference to the reading score

🔶 `Continuous` DV: Number of words.
🔵 `Nominal` IV: Three different instructions

Descriptives

```
# Descriptives
library(dplyr)
group_by(NonParametricData,Method) %>%
  summarise(
    count = n(),
    median = median(Words, na.rm = TRUE),
    IQR = IQR(Words, na.rm = TRUE))
```

```
## # A tibble: 3 × 4
##   Method count median   IQR
##   <chr>  <int>  <int> <dbl>
## 1 A          7    820   215
```

```
## 2 B            7     480     65
## 3 C            7     600     80
```

Data Analysis

In R the Welch's test is done by defaul. I will show a newer package of conducting t-test.

The code to conduct a `t-test` using the `rstatix` package.

```
oneway.test(DV ~ IV, data = data, var.equal = FALSE)
```

```
library(misty)
oneway.test(Words ~ Method, data = NonParametricData, var.equal = FALSE)
```

```
##
##  One-way analysis of means (not assuming equal variances)
##
## data:  Words and Method
## F = 14.867, num df = 2.000, denom df = 11.184, p-value = 0.0007079
```

Post Hoc

```
library(rstatix)
NonParametricData %>% games_howell_test(Words ~ Method)
```

```
## # A tibble: 3 × 8
##    .y.   group1 group2 estimate conf.low conf.high p.adj p.adj.signif
## * <chr> <chr>  <chr>     <dbl>    <dbl>     <dbl> <dbl> <chr>
## 1 Words A      B         -280     -424.     -136.  0.001 ***
## 2 Words A      C         -220     -365.     -75.3 0.006 **
## 3 Words B      C          60.0    -15.3     135.  0.126 ns
```

Interpretation

Due to the violation of homogeneity of variance, we conducted a Welch's ANOVA a non-parametric alternative to examine the differences in the number of words recorded by the different methods A, B, C. The results indicated that there were statistically significant differences $F(2, 11.18) = 14.87$, $p = 0.001$. Continue to the post hoc interpretation…

Decision Chart

```
                              ┌──────────┐
                              │  ANOVA   │
                              └──────────┘
                         ╱                    ╲
                        ╱                      ╲
    ┌────────────────────────────────────┐   ┌────────────────────────────────────────────┐
    │ Normality violation (Shapiro Wilk  │   │ Homogeneity of Variance violation (Levene  │
    │ p < alpha)                         │   │ p < alpha)                                 │
    └────────────────────────────────────┘   └────────────────────────────────────────────┘
                    │                                         │
                    ▼                                         ▼
           ┌──────────────────┐                      ┌──────────────────┐
           │  Kruskal Wallis  │                      │   Welch ANOVA    │
           └──────────────────┘                      └──────────────────┘
```
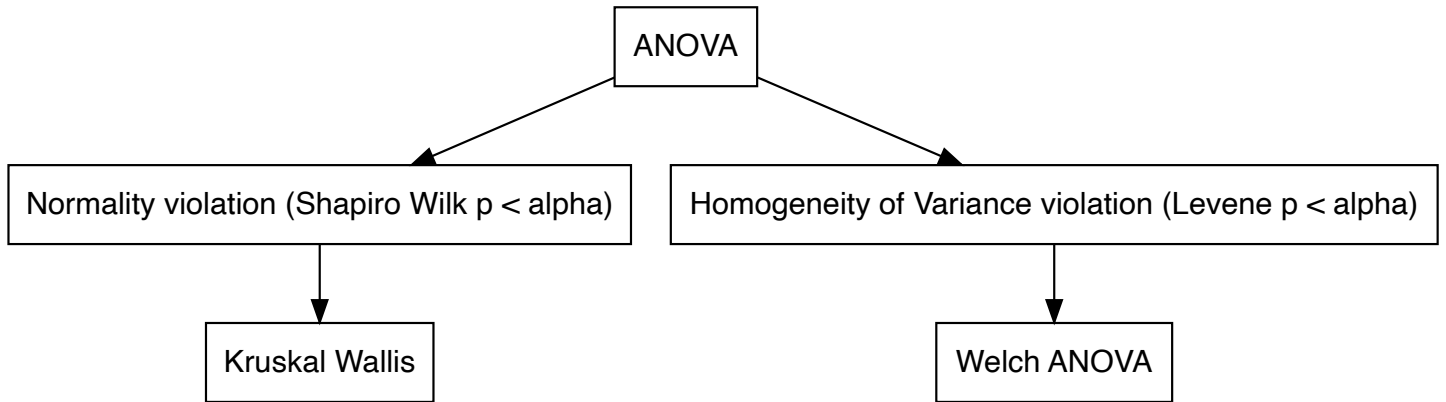
Resources

- Research Design & Data Analysis Lab:

  https://www.uttyler.edu/research/ors-research-design-data-analysis-lab/

- Schedule a consultant appointment with me:

  https://www.uttyler.edu/research/ors-research-design-data-analysis-lab/ors-research-design-data-analysis-lab-consultants/

- Check out Lab Resources (including recording of this webinar):

  https://www.uttyler.edu/research/ors-research-design-data-analysis-lab/resources/

References

How to perform welch's t-test in r. (2020). https://www.statology.org/welch-t-test-in-r/

Navarro, D., & Foxcroft, D. (2019). *Learning statistics with jamovi: A tutorial for psychology students and other beginners (version 0.70).*

Nimon, K. F. (2012). Statistical assumptions of substantive analyses across the general linear model: A mini-review. *Frontiers in Psychology, 3,* 322.