Non-Parametric Methods in SPSS

Samantha Estrada PhD
ORSSP Data Analysis Lab Consultant

Parametric vs. Non-Parametric

- 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.

Parametric	Non-Parametric
Independent Samples T-Test	Mann-Whitney U
Independent Samples T-Test	Welch's T-Test
Dependent Samples T-Test	Wilcoxon (not covered)
ANOVA	Kruskall Wallis ANOVA
ANOVA	Welch's ANOVA

Assumptions

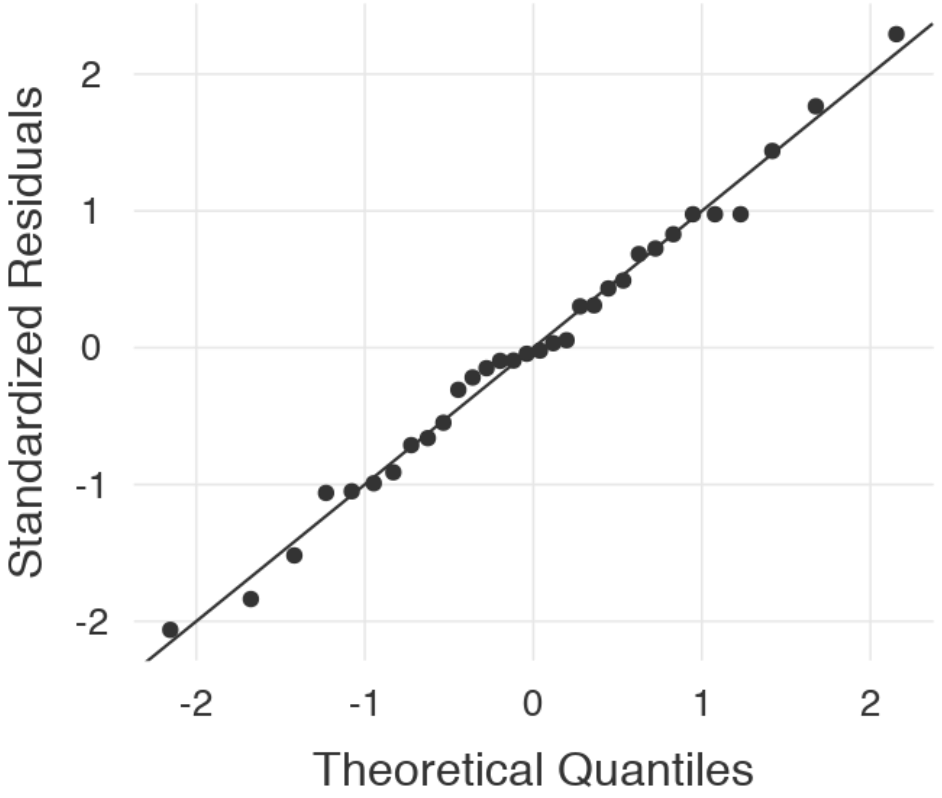
Normality

It is assumed that the data are normally distributed. Specifically, we assume that both groups are normally distributed.

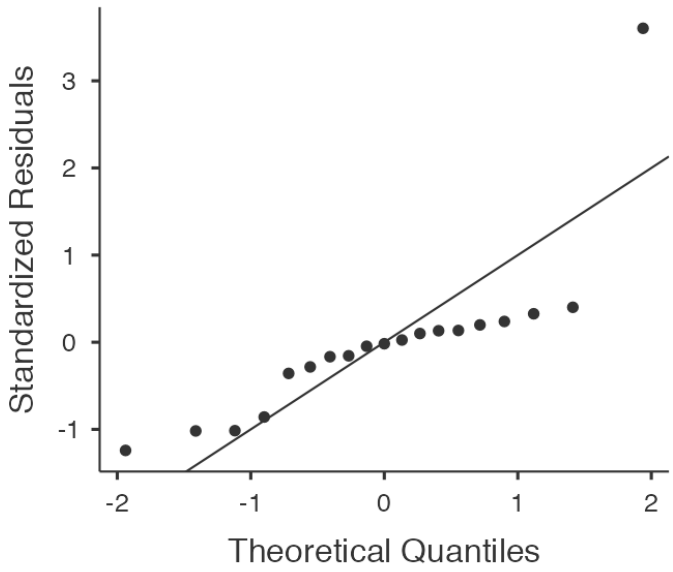
QQ Plots

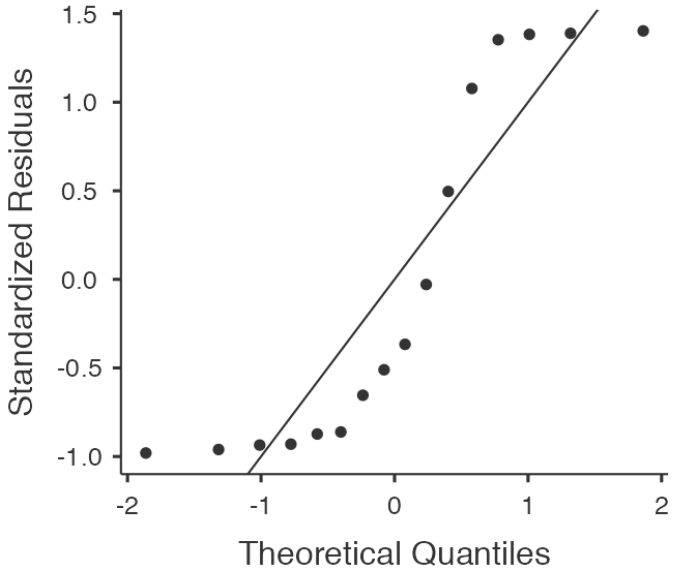
Examples of QQ-Plots when the assumption of normality is violated

- Shapiro-Wilk Test of Normality



Example of QQ-Plot when the normality assumption is met


Standardized Residuals
3
2
1
0
-1
-2
-1
0
1
2
Theoretical Quantiles


Standardized Residuals
1.5
1.0
0.5
0.0
-0.5
-1.0
-2
-1
0
1
2
Theoretical Quantiles

Assumptions

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).

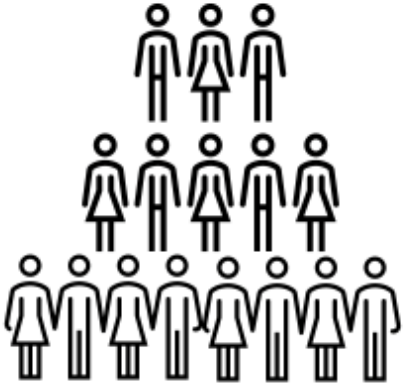
- The hypotheses for Levene’s Test

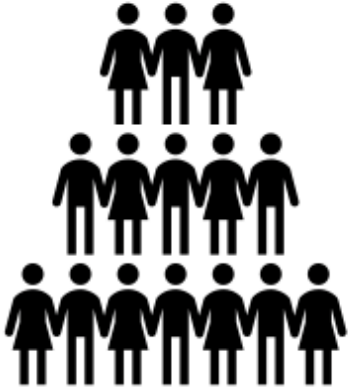
Ho: The variances are equal
Ha: The variances are not equal

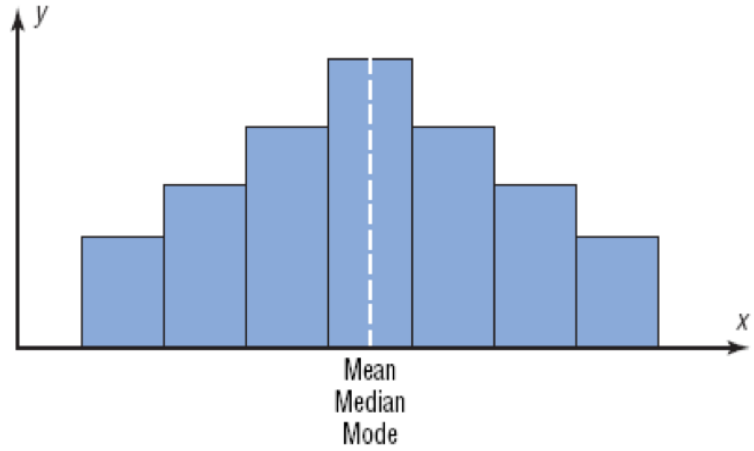
Resources for Assumptions

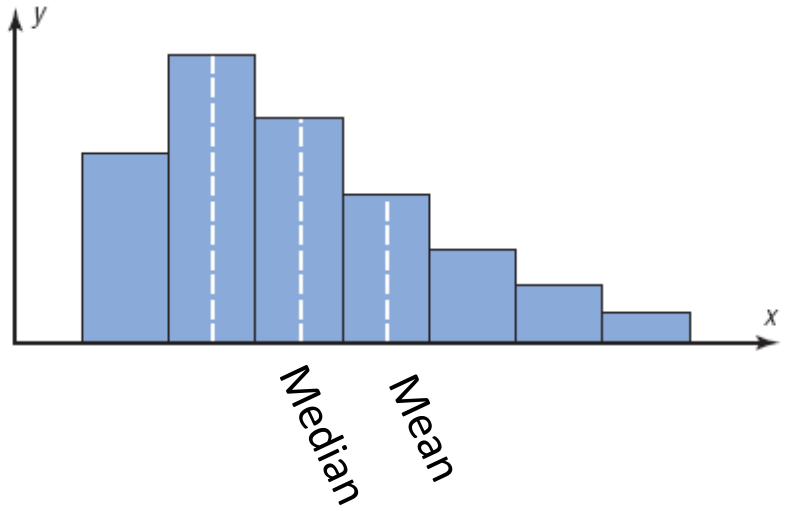
- Nimon (2012)

What is a t-test?






y
x
Mean
Median
Mode


y
x
Median
Mean

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

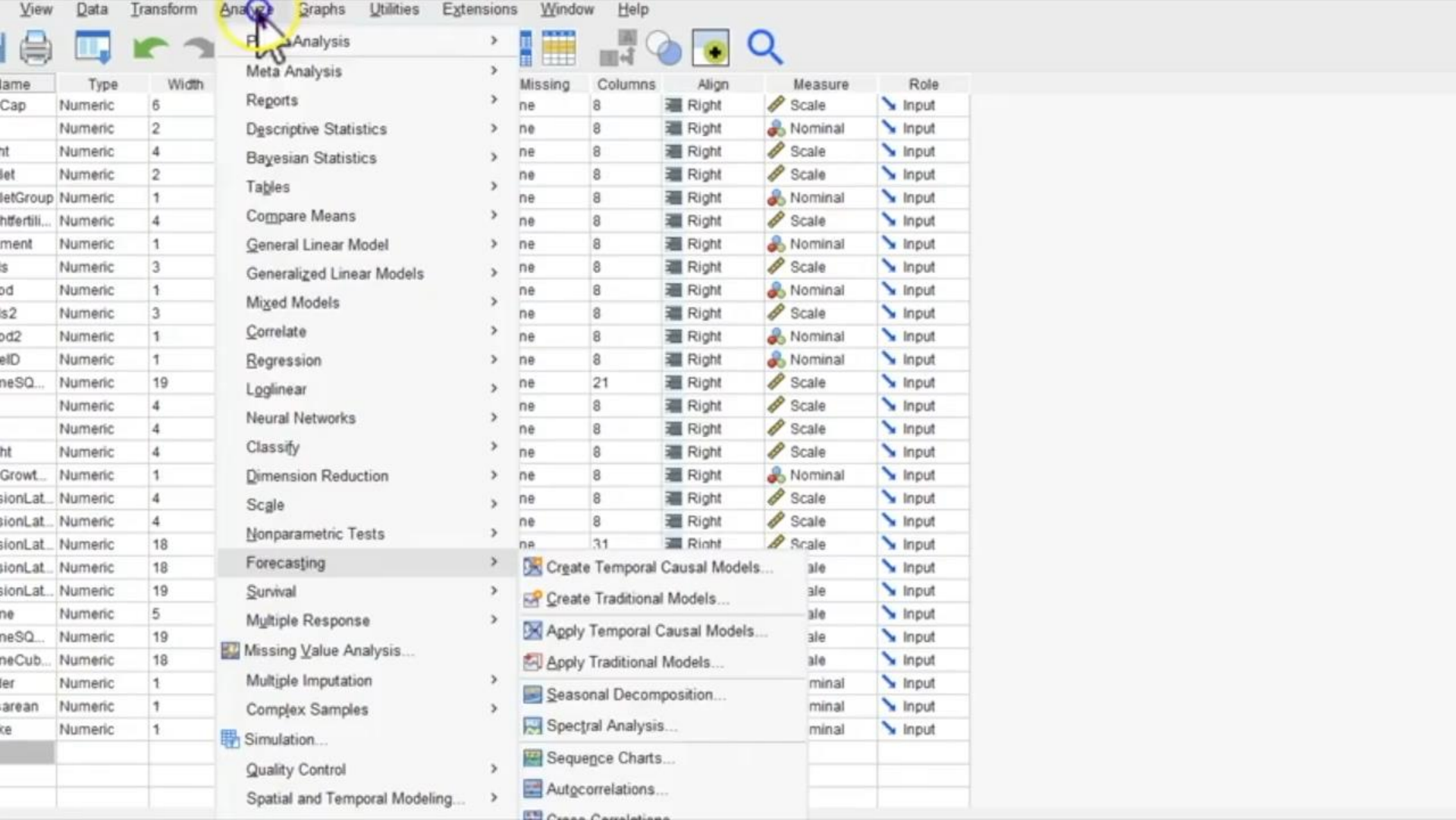
Assumptions

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

Mann-Whitney U

Example

- We will compare a non-smoker group to a group of smokers on their lung capacity.
  - Smoke: Smoking status
  - LungCapacity: Lung Capacity


View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Create Temporal Causal Models...
Create Traditional Models...
Apply Temporal Causal Models...
Apply Traditional Models...
Seasonal Decomposition...
Spectral Analysis...
Sequence Charts...
Autocorrelations...
Name
Type
Width
Missing
Columns
Align
Measure
Role

Key Output

Nonparametric Tests

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.[a,b]	Decision
1	The distribution of LungCap is the same across categories of Smoke.	Independent-Samples Mann-Whitney U Test	.006	Reject the null hypothesis.

a. The significance level is .050.

b. Asymptotic significance is displayed.

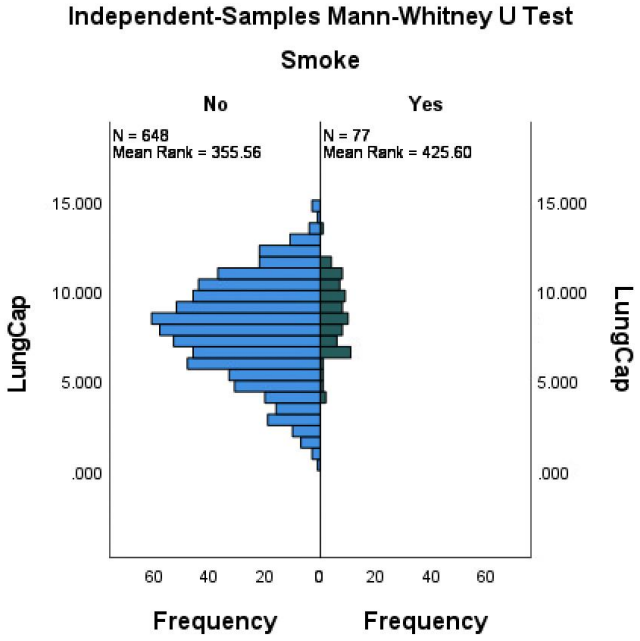
Key Output

Independent-Samples Mann-Whitney U Test

LungCap across Smoke

Independent-Samples Mann-Whitney U Test Summary

Total N	725
Mann-Whitney U	29768.000
Wilcoxon W	32771.000
Test Statistic	29768.000
Standard Error	1737.426
Standardized Test Statistic	2.774
Asymptotic Sig.(2-sided test)	.006



Interpretation

Mann-Whitney U test showed that there was a significant difference (W = 29,768, p = 0.006) in lung capacity between the group of smokers (Mdn = 8.65) compared to the non-smoker group (Mdn = 7.90).

Welch’s t-test

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.

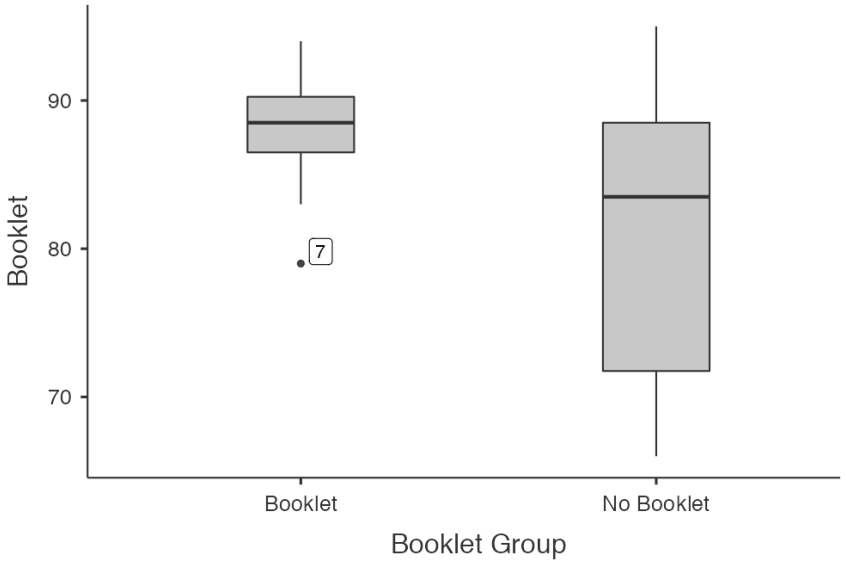
Hypotheses

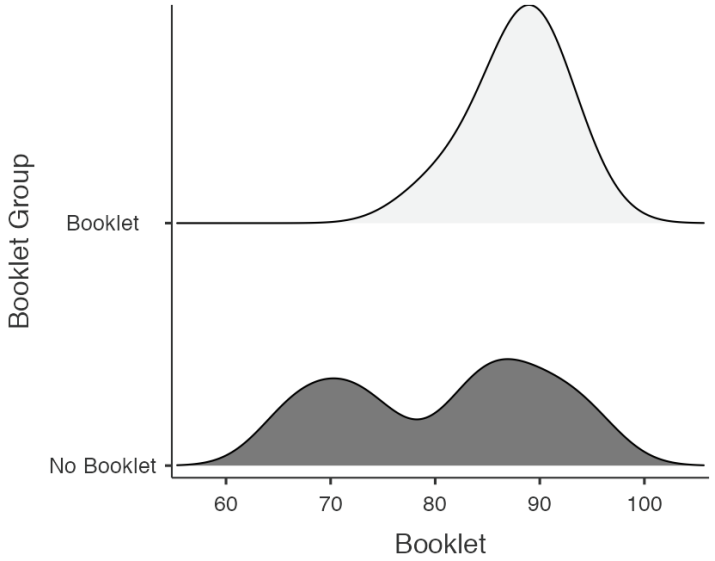
- Ho: The means are equal
- Ha: The means are not equal

Data Example: Welch's t-test

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 Group: Booklet and No Booklet
- Booklet Exam Score


Booklet
90
80
70
7
Booklet
No Booklet
Booklet Group


Booklet Group
Booklet
No Booklet
60
70
80
90
100
Booklet

Welch’s t-test

Independent Samples

View Data Transform Analyze Graphs Utilities Extensions Window Help

lame	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
Cap	Numeric	6	3		None	None	8	Right	Scale	Input
	Numeric	2	0		None	None	8	Right	Nominal	Input
ht	Numeric	4	1		None	None	8	Right	Scale	Input
let	Numeric	2	0		None	None	8	Right	Scale	Input
letGroup	Numeric	1	0		None	None	8	Right	Nominal	Input
htfertili...	Numeric	4	2		None	None	8	Right	Scale	Input
ment	Numeric	1	0		None	None	8	Right	Nominal	Input
ls	Numeric	3	0		None	None	8	Right	Scale	Input
od	Numeric	1	0		None	None	8	Right	Nominal	Input
ls2	Numeric	3	0		None	None	8	Right	Scale	Input
od2	Numeric	1	0		None	None	8	Right	Nominal	Input
elD	Numeric	1	0		None	None	8	Right	Nominal	Input
neSQ...	Numeric	19	15		None	None	21	Right	Scale	Input
	Numeric	4	1		None	None	8	Right	Scale	Input
	Numeric	4	1		None	None	8	Right	Scale	Input
ht	Numeric	4	2		None	None	8	Right	Scale	Input
Growt...	Numeric	1	0		None	None	8	Right	Nominal	Input
sionLat...	Numeric	4	0		None	None	8	Right	Scale	Input
sionLat...	Numeric	4	0		None	None	8	Right	Scale	Input
sionLat...	Numeric	18	16		None	None	31	Right	Scale	Input
sionLat...	Numeric	18	15		None	None	30	Right	Scale	Input
sionLat...	Numeric	19	16		None	None	31	Right	Scale	Input
ne	Numeric	5	0		None	None	8	Right	Scale	Input
neSQ...	Numeric	19	15		None	None	22	Right	Scale	Input
neCub...	Numeric	18	15		None	None	22	Right	Scale	Input
ler	Numeric	1	0		{1, Female}...	None	9	Right	Nominal	Input
sarean	Numeric	1	0		{1, No}...	None	12	Right	Nominal	Input
ke	Numeric	1	0		{1, No}...	None	8	Right	Nominal	Input

Output

Group Statistics

	BookletGroup	N	Mean	Std. Deviation	Std. Error Mean
Booklet	1	12	87.92	4.033	1.164
	2	12	80.83	10.205	2.946

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Booklet	Equal variances assumed	16.603	<.001	2.236	22	.018	.036	7.083	3.168	.514	13.653
	Equal variances not assumed			2.236	14.354	.021	.042	7.083	3.168	.305	13.862

Interpretation

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Booklet	Equal variances assumed	16.603	<.001	2.236	22	.018	.036	7.083	3.168	.514	13.653
	Equal variances not assumed			2.236	14.354	.021	.042	7.083	3.168	.305	13.862

- As we can see in the `Equal Variance assumed` output, we are violating the equal variances assumption ($F(1, 22) = 2.236$, $p < .001$)

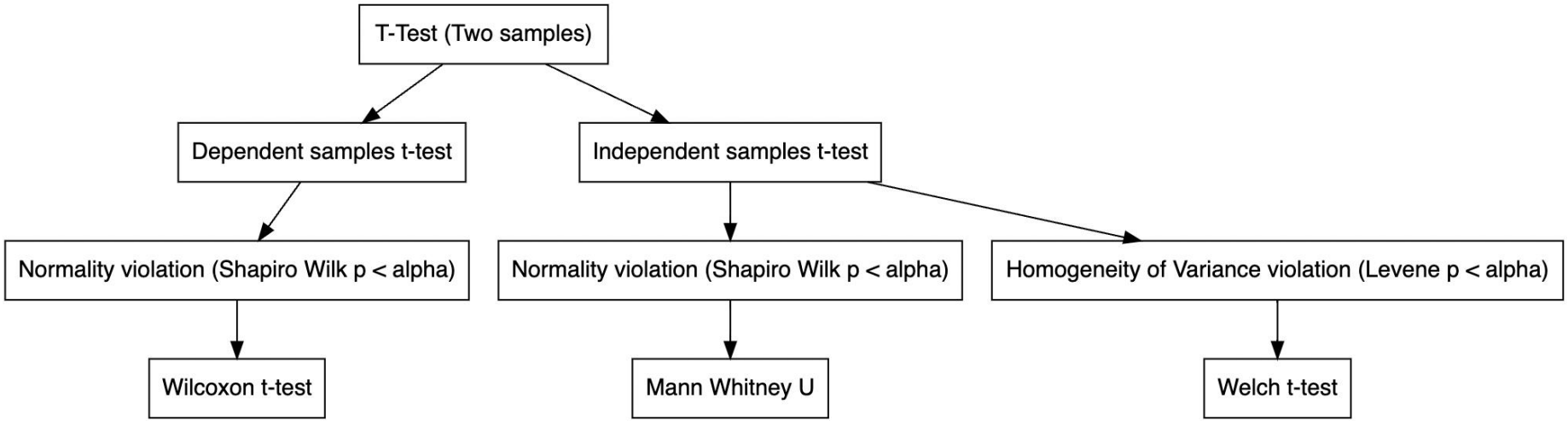
Interpretation

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Booklet	Equal variances assumed	16.603	<.001	2.236	22	.018	.036	7.083	3.168	.514	13.653
	Equal variances not assumed			2.236	14.354	.021	.042	7.083	3.168	.305	13.862

The assumption of equal variances was checked using Levene's Test (F (1,22) = 2.236, 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) = 2.236, p = 0.042). Next… report the means and/or mean difference.

Decision Chart



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.

Data Example: Kruskal-Wallis

For example, suppose weights of poplar trees are different based on treatments (none treatment, fertilizer, irrigation, or fertilizer and irrigation). 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.

Hypotheses

- Ho: Populations medians are equal
- Ha: At least, one population median is different from other populations

View Data Transform Analyze Graphs Utilities Extensions Window Help

lizer .02 Visible: 35

Weightfertilizer	Treatment	VAR00010
.15	0	
.02	0	
.16	0	
.37	0	
.22	0	
1.34	1	
.14	1	
.02	1	
.06	1	
.03	1	
.23	2	
.04	2	
.34	2	
.16	2	
.05	2	
2.03	3	
.27	3	
.92	3	
1.07	3	
2.38	3	
.	.	
.	.	
.	.	
.	.	
.	.	
.	.	
.	.	

NPar Tests

Kruskal-Wallis Test

Ranks

	Treatment	N	Mean Rank
Weightfertilizer	None	5	9.00
	Fertilizer	5	7.50
	Irrigation	5	8.50
	Fertilizer and Irrigation	5	17.00
	Total	20	

Test Statistics[a,b]

	Weightfertilizer
Kruskal-Wallis H	8.233
df	3
Asymp. Sig.	.041

a. Kruskal Wallis Test

b. Grouping Variable: Treatment

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 (3) = 8.23, p = .041.

Welch’s ANOVA

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

Hypotheses

Same as ANOVA.

Data Example: Welch's ANOVA

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

View Data Transform Analyze Graphs Utilities Extensions Window Help

2 Visible: 34

AR0001 1	Words	Method	VAR00012
.		1	
.	850	1	
.	820	1	
.	640	1	
.	920	1	
.	480	2	
.	460	2	
.	500	2	
.	570	2	
.	580	2	
.	500	3	
.	550	3	
.	480	3	
.	600	3	
.	610	3	
.	640	1	
.	920	1	
.	480	2	
.	460	2	
.	600	3	
.	610	3	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	

Robust Tests of Equality of Means

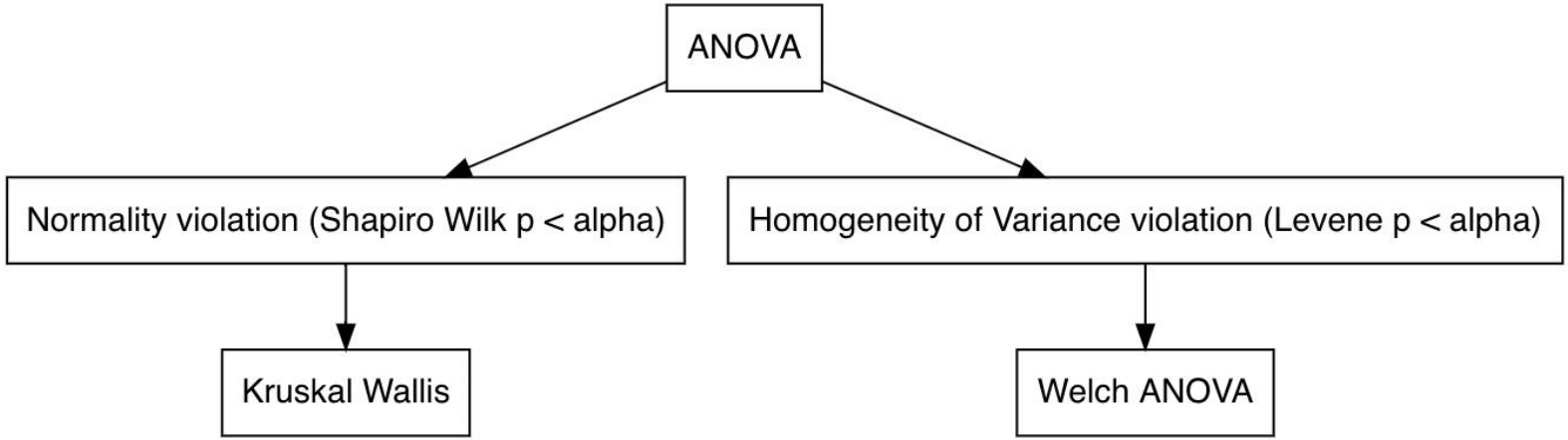
Words

	Statistic[a]	df1	df2	Sig.
Welch	14.867	2	11.184	<.001

a. Asymptotically F distributed.

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, 12) = 14.867, p <.001.


ANOVA
Normality violation (Shapiro Wilk p < alpha)
Homogeneity of Variance violation (Levene p < alpha)
Kruskal Wallis
Welch ANOVA

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.