# Non parametric tests

Elias Zintzaras, M.Sc., Ph.D.

Professor in Biomathematics-Biometry
Department of Biomathematics
School of Medicine
University of Thessaly

Institute for Clinical Research and Health Policy Studies
Tufts University School of Medicine
Boston, MA, USA

Theodoros Mprotsis, MSc, PhD
Teacher & Research Fellow
(http://biomath.med.uth.gr)
University of Thessaly
Email: tmprotsis@uth.gr

Energy consumption between thin and obese women
Mann-Whitney U test
(Independent sample t-test)
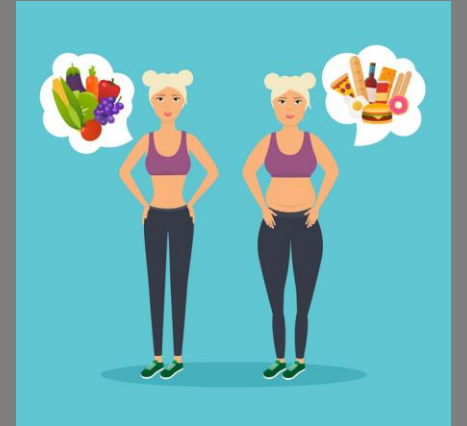
Energy consumption between thin and obese women

Thin ($n = 13$)	Obese ($n = 9$)
6.13	8.79
7.05	9.19
7.48	9.21
7.48	9.68
7.53	9.69
7.58	9.97
7.90	11.51
8.08	11.85
8.09	12.79
8.11	
8.40	
10.15	
10.88	

The energy consumption over 24 hours for a group of thin women and a group of obese women is shown in the table.

Is there a difference in energy consumption between the two groups?

Entering data and defining variables

- Enter the data in the **Data View.**
- One variable represents the **group**, and the other represents the corresponding **energy consumption**
- Then, name the variables accordingly in the **Variable View**

Running the Mann-Whitney U test

To analyze the data, select from the menu

Analyze -> Nonparametric Tests -> Legacy Dialogs -> 2 Independent Samples
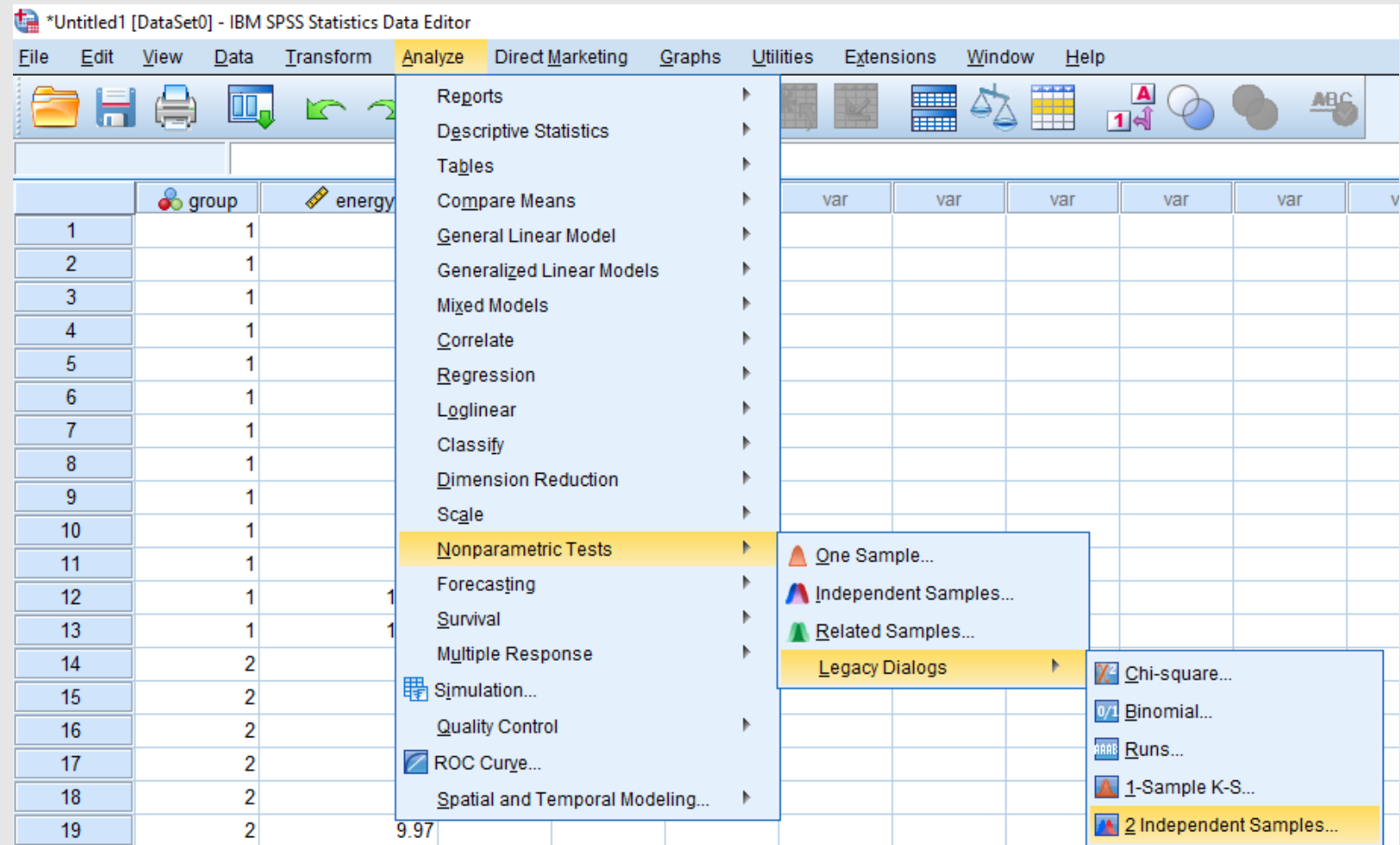
Running the Mann-Whitney U test

In window (1), drag the variable **energy** from the left panel to the **Test Variable List** box, and drag the variable **group** to the **Grouping Variable:** box**.** Click the **Define Groups…** button, specify the two groups (3), and then click **Continue.** Next, click **Options**, select **Descriptive** (2), and click **Continue** and **OK**.

Results and interpretation

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
energy	22	8.9791	1.69750	6.13	12.79
group	22	1.41	.503	1	2

Ranks

	group	N	Mean Rank	Sum of Ranks
energy	Thin	13	7.92	103.00
	Obese	9	16.67	150.00
	Total	22		

Test Statistics[a]

	energy
Mann-Whitney U	12.000
Wilcoxon W	103.000
Z	-3.106
Asymp. Sig. (2-tailed)	.002
Exact Sig. [2*(1-tailed Sig.)]	.001[b]

a. Grouping Variable: group

b. Not corrected for ties.

The mean rank for the **thin** group is **7.92**, and for the **obese** group, it is **16.67**. This difference in mean ranks suggests that the energy consumption is generally lower in the **thin** group compared to the **obese** group.

The **Asymptotic Sig. (2-tailed)** value of is less than **0.05**. This indicates that there is a statistically significant difference in energy consumption between the two groups **(U=12, p=0.002<0.05)**.

For large samples (where at least one of the two groups has n>20) Given that the absolute Z value (**|−3.106|**) is much larger than the typical critical value of 1.96 (for a two-tailed test at the 0.05 significance level), it indicates that the difference between the two groups is statistically significant.

# Comparison of Newborn Weights Between Non-Smoking and Smoking Mothers

The weights of newborn children, born to 15 non-smoking women and 14 smoking women are as follows:

Sample1 Non-smoking	Sample 2 Smoking
3.99	3.18
3.79	2.84
3.60	2.90
3.73	3.27
3.21	3.85
3.60	3.52
4.08	3.23
3.61	2.76
3.83	3.60
3.31	3.75
4.13	3.59
3.26	3.63
3.54	2.38
3.51	2.34
2.71	

Is there a difference in the weights of the children between the two groups?

Comparing Relief Times of Two Analgesic Drugs Wilcoxon test

Paired sample t-test
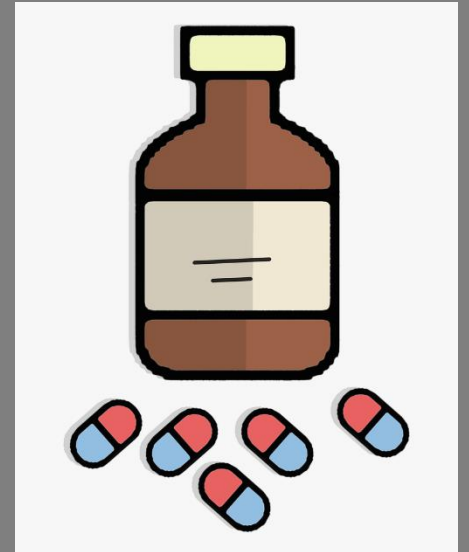
Comparing Relief Times of Two Analgesic Drugs

Suppose we want to compare the relief times of two analgesic drugs, A and B. Additionally, suppose that 8 patients were given drug A first and then drug B.

The following relief times were observed:

*Untitled1 [DataSet0] - IBM SPSS Statistics Da		
File Edit View Data Transform		
	drug_a	drug_b
1	7.50	3.80
2	3.60	2.40
3	8.90	4.70
4	6.40	5.90
5	9.50	2.10
6	7.40	3.20
7	10.10	4.50
8	4.70	3.80

To analyze the data, select from the menu

Analyze -> Nonparametric Tests -> Legacy Dialogs -> K Independent Samples

Running the Wilcoxon test

In window (1), drag the variables **drag_a** and **drag_b** from the left panel to the **Test Pairs** box, one by one**.** Click the **Options…** button, select **Descriptive** (2), and click **Continue** and **OK**.

Results and interpretation

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
drug_a	8	7.2625	2.28469	3.60	10.10
drug_b	8	3.8000	1.24900	2.10	5.90

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
drug_b - drug_a	Negative Ranks	8[a]	4.50	36.00
	Positive Ranks	0[b]	.00	.00
	Ties	0[c]		
	Total	8		

a. drug_b < drug_a

b. drug_b > drug_a

c. drug_b = drug_a

Test Statistics[a]

	drug_b - drug_a
Z	-2.524[b]
Asymp. Sig. (2-tailed)	.012

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

- All 8 patients experienced a lower relief time with drug B than with drug A, which means drug B was less effective for all patients compared to drug A (negative ranks).

- No patients experienced a higher relief time with drug B compared to drug A (positive ranks).

- There were no ties; every patient's relief time differed between drug A and drug B

Based on the **Wilcoxon Signed Ranks Test** results, we can conclude that there **is a statistically significant difference** in the relief times between the two analgesic drugs, A and B. Specifically, drug A provides significantly longer relief times than drug B (p<0.05). Therefore, drug A appears to be more effective than drug B in this sample of patients.

Effectiveness of a Hypnotic Drug

In a clinical study to test the effectiveness of a hypnotic drug, sleep duration (in hours) was observed in 10 patients on one night after drug administration and on another night after placebo administration. The results were as follows:

Patient ID	Drug	Placebo
1	6.1	5.2
2	7.0	7.9
3	8.2	3.9
4	7.6	4.7
5	6.5	5.3
6	8.4	5.4
7	6.9	4.2
8	6.7	6.1
9	7.4	3.8
10	5.8	6.3

Evaluating the Efficacy of Three Treatments for Knee Pain
Kruskal-Wallis

One-way Anova

Evaluating the Efficacy of Three Treatments for Knee Pain

Thirty people with similar knee pain are recruited by a researcher who aims to determine whether three different medications have varying effects on knee pain. The participants are then randomly assigned to one of three groups and given one of the treatments (Drug 1, Drug 2, or Drug 3).

The data is shown in the image.

	drug	pain	var
1	1	78	
2	1	65	
3	1	63	
4	1	44	
5	1	50	
6	1	78	
7	1	70	
8	1	61	
9	1	50	
10	1	44	
11	2	71	
12	2	66	
13	2	56	
14	2	40	
15	2	55	
16	2	31	
17	2	45	
18	2	66	
19	2	47	
20	2	42	
21	3	57	
22	3	88	
23	3	58	
24	3	78	
25	3	65	
26	3	61	
27	3	62	
28	3	44	
29	3	48	
30	3	77	

To analyze the data, select from the menu

Analyze -> Nonparametric Tests -> Legacy Dialogs -> K Independent Samples

Running the Kruskal-Wallis test

In window (1), drag the variable **pain** from the left panel to the **Test Variable List:** box, and the variable **drug** from the left panel to the **Grouping Variable:** box**.** Click the **Options…** button, select **Descriptive** (2), and click **Continue**. Then, click the **Define Range…** button, enter 1 into the **Minimum:** field and 3 into the **Maximum:** field. Click **Continue** and **OK**.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
pain	30	58.6667	13.73978	31.00	88.00
drug	30	2.00	.830	1	3

Kruskal-Wallis Test

Ranks

	drug	N	Mean Rank
pain	1	10	16.70
	2	10	11.60
	3	10	18.20
	Total	30	

Test Statistics[a,b]

	pain
Kruskal-Wallis H	3.097
df	2
Asymp. Sig.	.213

a. Kruskal Wallis Test

b. Grouping Variable: drug

The **Ranks** table shows the average ranks of the pain scores for each of the three drug groups:
- Drug 1 has a mean rank of 16.70.
- Drug 2 has a mean rank of 11.60.
- Drug 3 has a mean rank of 18.20.

The Kruskal-Wallis test p-value of 0.213 is greater than the significance level of 0.05. This indicates that there is **no statistically significant difference** in the median pain scores among the three drug groups. In other words, based on this test, there is not enough evidence to suggest that the three drugs have different effects on knee pain.

Comparing the Effects of Different Exercise Types on Depression Levels

A clinical study was conducted to evaluate the impact of three different workout routines on depression levels. Eighteen patients diagnosed with depression were randomly assigned to one of three exercise groups:

- **Aerobic Exercise Group** (e.g., jogging)
- **Resistance Training Group** (e.g., weight lifting)
- **Yoga Group**

Each group followed their assigned exercise program for 8 weeks. At the end of the study, depression levels were measured using the Beck Depression Inventory (BDI), with lower scores indicating lower levels of depression.

The data is shown in the table.

Participant ID	Exercise Group	Depression Score (BDI)
1	Aerobic	12
2	Aerobic	15
3	Aerobic	10
4	Aerobic	14
5	Aerobic	11
6	Aerobic	13
7	Resistance	8
8	Resistance	9
9	Resistance	7
10	Resistance	10
11	Resistance	6
12	Resistance	8
13	Yoga	14
14	Yoga	16
15	Yoga	12
16	Yoga	18
17	Yoga	17
18	Yoga	13