

Non-parametric methods: The Sign Test in SPSS

In this document we introduce nonparametric tests. Especially we focus on the 'Sign test' that can be used as an alternative to matched pairs t-test.

Consider 'Schizophrenia' example where volume of left hippocampus region in brain was measured for 15 pairs of twins. One twin was affected by schizophrenia while other was not. We have to input these data manually since there is no data file for this example. So open a new file in SPSS and click on 'Variable View' button, here we introduce two variables:

File	Edit	View Data	Transform	<u>A</u> nalyze I	Direct <u>M</u> arket	ing <u>G</u> raphs <u>U</u>	tilities Add-	ons Window	/ <u>H</u> elp				
1			📮 🗠 r	ы 📕	=	i ip Aa		ž 🚍 4	à 🔳 🛯	M 📀	A		
		Name	Type	Width	Decimals	Labe	l .	Values	Missing	Columns	Align	Measure	Role
1		Unaffect	Numeric	8	2	Not Affected by s	chizophrenia	None	None	8	🔳 Right	🛷 Scale	🔪 Input
2		Affected	Numeric	8	2	Affected by schize	ophrenia	None	None	8	🔳 Right	🛷 Scale	🔪 Input

Next in the 'Data View' section we enter the observations:

<u>F</u> ile	Edit	View	Data	1	ransforr	
6			1	Ļ	5	
5:						
		Unaff	ect	Aff	ected	
	1		1.94		1.27	
	2		1.44		1.63	
	3		1.56		1.47	
	4		1.58	1.39		
	5		2.06	1.93		
	6		1.66		1.26	
	7		1.75		1.71	
	8		1.77		1.67	
	9		1.78	1.28		
•	10		1.92	1.85		
· ·	11		1.25		1.02	
•	12		1.93		1.34	
	13		2.04		2.02	
	14		1.62		1.59	
	15		2.08		1.97	

The goal is test if the means of these two variables ('Unaffect' and 'Affected') are the same or not, equivalently whether the difference between two means is zero or not. Since this is a matched pair example, we go to Analyze > Compare Means > Paired-Sample T Test

<u>A</u> nalyze	Direct Marketing	<u>G</u> raphs	Utilities	Add- <u>o</u> ns	Window	<u>H</u> elp				
Repo	rts	•	×.							
Desc	riptive Statistics									
Ta <u>b</u> le	s	+								
Com	oare Means	- F.	Means	B						
<u>G</u> ene	ral Linear Model	•	One-Sample T Test							
Gene	ralized Linear Model	Is 🕨		endent-Sam	ples T Test.					
Mixed	Mixed Models		Pairer	-Samples 1	Test					
Corre	late				TOJE.					
Regre	ession		Une-v	vay ANUVA.	-					

Now using arrow move 'Unaffect' to first column and 'Affected' to second (we use only pair 1):

ţ	ł		Paired-	Samples T T	est		×
	Not Affected by schiz Affected by schizoph		Paired Variables: Pair Variable1 Variable2 1 & Not Affec & Affected 2				Options) Bootstrap
		*				¢	
	(ок	Paste (Reset Car	icel Help	÷	

Click **OK** and the following table is produced:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
1	Not Affected by schizophrenia	1.7587	15	.24242	.06259
	Affected by schizophrenia	1.5600	15	.30126	.07778

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Not Affected by schizophrenia & Affected by schizophrenia	15	.635	.011

Pair

				Paired Samples	s Test					
	Paired Differences									
				Std. Error	95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Mean	Lower	Upper	t	df	Sig. (2-tailed)	
Pair 1	Not Affected by schizophrenia - Affected by schizophrenia	.19867	.23829	.06153	.06670	.33063	3.229	14	.006	

The two sided p-value is 0.006 and therefore we reject null hypothesis and conclude that means are not the same. However we should not use paired t-test without checking some conditions. In order to use t-test, the pairs have to be independent and distribution of differences approximately normal. We assume that the first condition is satisfied by design but to check second condition we have to analyze the differences. Let's first compute variable of differences by going to **Transform** > **Compute Variable**, call new variable 'Diff' and enter a simple expression:

ta		Compute Variable
Target Variable:		Numeric Expression:
Diff	=	Unaffect - Affected

Click **OK**, the new column is computed:

Unaffect	Affected	Diff
1.94	1.27	.67
1.44	1.63	19
1.56	1.47	.09
1.58	1.39	.19
2.06	1.93	.13
1.66	1.26	.40
1.75	1.71	.04
1.77	1.67	.10
1.78	1.28	.50
1.92	1.85	.07
1.25	1.02	.23
1.93	1.34	.59
2.04	2.02	.02
1.62	1.59	.03
2.08	1.97	.11

Once we have variable of differences we can easily make a boxplot of this variable (Graph > Chart Builder > Boxplot):

GGraph



We observe that the boxplot is skewed to the right. Also we can check normality using quantilequantile plot. Go to Analyze > Descriptive Statistics > Q-Q Plots

<u>A</u> nalyze	Direct <u>M</u> arketing	Graphs	Utilities	Add-ons	1
Repo	irts	•	*		Π
Desc	riptive Statistics	•	123 Frequ	encies	
Table	s	•	Descr	iptives	
Com	pare Means		A Explor	e	
<u>G</u> ene	ral Linear Model	•	Cross	tabs	
Gene	ralized Linear Mode	ls 🕨	TURE	Analysis	
Mi <u>x</u> eo	Models	•	- Ratio	, analyere	
Corre	late	*			
Regr	ession	+	P-PP	ots	
Logli	near	- F	2-Q P	lots	

Move 'Diff' to the 'Variables' section:

6	Q-Q Plots	×
Not Affected by schiz Affected by schizoph	Variables:	Test Distribution
	Transform Natural log transform Standardize values Difference: 1 Seasonally difference: 1 Current Periodicity: None OK Paste Reset Cancel	Proportion Estimation Formula @ giom's O Randt O Typkey's O Van der Waerden's Rank Assigned to Ties @ Mean O High O Low O Break ties arbitrarity Help

Click **OK** and two plots are created, we need only the first one:



We see that points depart from the straight line and therefor we cannot assume that the distribution of differences is nearly normal and therefore t-test may not be appropriate.

We can however compare these two groups using the sign test (exact binomial) rather than t test. To use sign test we do not have to assume normally of differences. For the sign test we only need to know number of positive and negative differences, the magnitudes of differences are not used at all. There is a built-in function in SPSS for the sign test, go to **Analyze** > **Nonparametric Tests** > **Legacy Dialogs** > **Binomial**:

Analyze	Direct Marketing	Graphs	Utilities	Add- <u>o</u> ns	Window	Help				
Repo	rts	+	×.		-				AB	
Desc	riptive Statistics						1.4			
Table	s	•								
Com	oare Means		var	var	var		var	var	Vé	
Gene	ral Linear Model	•								
Gene	ralized Linear Mode	ls ▶								
Mixed	Models	•								
Corre	late									
Regr	ession	•								
Loglin	near	•								
Neura	al Net <u>w</u> orks									
Class	ify									
Dime	nsion Reduction									
Scale										
Nonp	arametric Tests	- F	🛕 One S	ample						
Forec	asting		/ Indep	endent Sam	oles					
Surviv	al		Relate	d Samples						
Multip	le Response		Lega	v Dialogs		E C	EE ou			
💯 Missir	ng Value Analysis		20900	, s.a.ogo			2m-squar	e		
Multie	le Imputation					<u>0/1</u> E	inomial			

Move 'Diff' to 'Test Variable List' section, also in 'Define Dichotomy' option select 'Cut point' and enter 0, since we need number of positive and negative differences:

(Binomial Test	×
Not Affected by schizop Affected by schizophre	Test Variable List	Exact Options
Oefine Dichotomy Get from data CCL point OK Pase	T <u>e</u> st Proportion: 0.50)

To finish click $\mathbf{OK},$ here are the results:

NPar Tests

Binomial Test								
		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2- tailed)		
Diff	Group 1	<= 0	1	.07	.50	.001		
	Group 2	> 0	14	.93				
	Total		15	1.00				

We see that we have 14 positive and 1 negative differences. The two sided p-value for this test is 0.001, we reject null hypothesis and conclude that two groups have different means.