Biostatistics

Doctor 2017 | Medicine | JU

Number >>	11
Doctor	hamza
Done By	shehab
CorrectedBy	...

In the last two lectures we have discussed the following:

1 – the difference between parametric and non-parametric interferential tests

2 – Chi-square as a non-parametric test

In this lecture we will discuss **T-test** as an example of the parametric test

T-tset

It is the most common parametric test used.

In order to use it, 3 conditions must be satisfied:

A – the dependent variable must be continuous

B – the sample must be normally distributed

C – the independent variable has only two groups

Note: if the independent variable has 3 or more groups we must use **One way anova**

What is t test?

It is a useful way for **comparing mean** values of two sets of numbers. The comparison will provide you with a statistic for evaluating whether the difference between two means is statistically significant or not **on the population level.**

Types of T-test

1 **– the independent t-test:**

It is used to compare two different groups (there is no over lapping between them)

Ex: males and females, dentists and physician, ill and healthy

2 **– the dependent t-test or Paired:**

It is used when there is only one group that has been tested twice.

Ex: the health state of individual before and after taking a specific drug.

The process of biostatistics or hypothesis testing is divided into 6 main steps:

1 – **setting the value of alpha**

Alpha is the acceptable margin of committing a Type One Error.

Type one error (rejecting the null hypothesis when it is true)

Usually Alpha is 0.05 but **not always**.

2 – **write the alternative and the null hypothesis**

In t-test: the alternative hypothesis always states that there is a difference in the mean between two groups and always the null hypothesis states that there is no difference.

3**- calculation process**:

in t-test we must calculate **the t value** using specific equation. The equation is different among the two types of t-test. This t value that we calculate is called the **calculated t value or T calc.**

4 **– find the critical value of the statistic:**

After calculating T calc, we must compare it with a critical value. This critical value is called **critical value of T.** we get this value from T-table

The column represents alpha

The row represents (degree of freedom). This value is calculated as follows:

In the independent test:

DF=(number of individual in group 1 + number of individual in group 2) - 2

In the paired t-test:

DF= number of pairs (number of individuals) - 1

	α					
df	0.250	0.100	0.050	0.025	0.010	0.005
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
⋮						
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
40	0.681	1.303	1.684	2.021	2.423	2.704
60	0.679	1.296	1.671	2.000	2.390	2.660
120	0.677	1.289	1.658	1.980	2.358	2.617
∞	0.674	1.282	1.645	1.960	2.326	2.576

5- State the decision rule

In t-test, we compare T calc with the critical value of T and the follows applies:

If the Tcalc was **larger** than The critical value of T →there is a difference → **reject** the null hypothesis

If the Tcalc was **smaller** than The critical value of T → there is **no** difference → **accept** the null hypothesis.

6- write the result

The independent t-test

the picture to the write shows the formula that we use to calculate t value.

Note:

SD is the standard deviation

SD **square** is the variance

N1 is the number of individuals in group 1

Formula is: $t_{\bar{X}_1-\bar{X}_2} = \frac{\bar{X}_1 - \bar{X}_2}{SE_{diff}}$

Terms in the numerator are the sample means.

Term in the denominator is the standard error of the difference between means.

The formula for the standard error of the difference in means:

$$SE_{diff} = \sqrt{\frac{SD_1^2}{N_1} + \frac{SD_2^2}{N_2}}$$

Paired t test

this is the formula that we use to calculate t value.

Note**: D represent the difference between the final and initial observation of the individual**

And **D bar represents the Mean of the D.**

note: n pairs is the number of individuals. Each individual represents a pair because he/she is tested twice.

Formulas:

$$t_{\bar{X}_D} = \frac{\bar{D}}{SE_{diff}}$$

is the difference in means over a standard error.

$$SE_{diff} = \frac{SD_D}{\sqrt{n_{pairs}}}$$

The standard error is found by finding the difference between each pair of observations. The standard deviation of these difference is SD_D. Divide SD_D by sqrt (number of pairs) to get SE_{diff}.

What is the proper way of writing the conclusion?

Let's assume that you conduct a study to find out if there is a difference in height between males and females in Jordan. After collecting, organizing and summarizing data, you used an independent t-test method to find the answer of your study. You set an alpha of 0.03 and after doing the calculation you got a t-value of 3.6 and the critical value of T was 1.67. what is the conclusion?

The conclusion is written as follows

(you are **97% confident** that there is a statically significant difference in height (the dependent variable) between males and females (independent variable) on the population level in Jordan (t clac=3.6, df=)

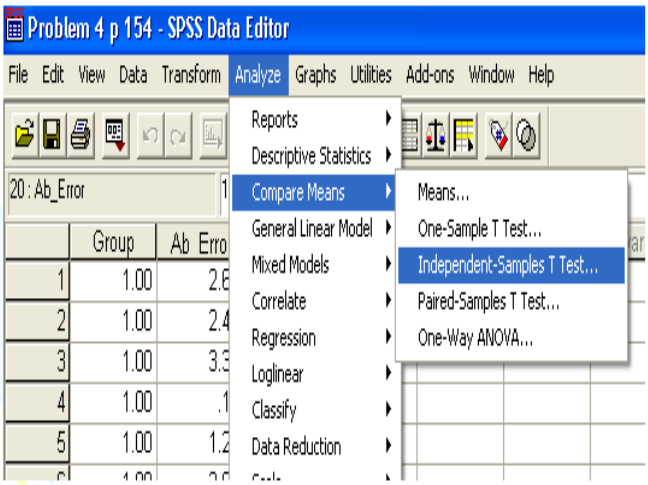
Note: please refer to the slides and try to take a look on other examples.

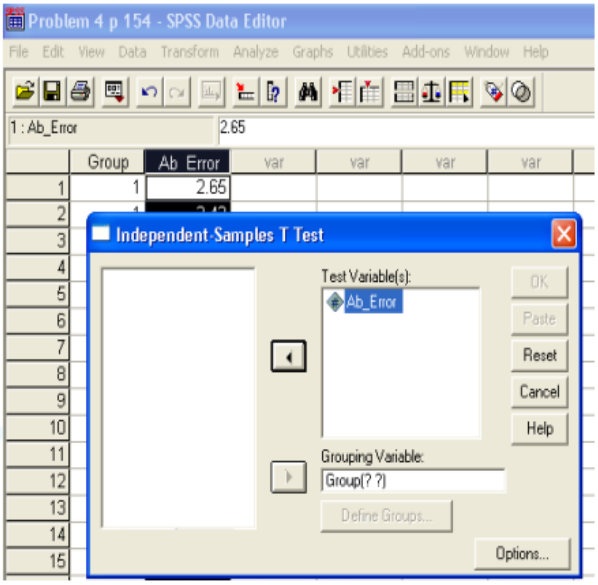
SPSS

It is a computer program that is used to calculate t-value.

How to use the program:

After entering the data follow the steps that are shown in the right picture.



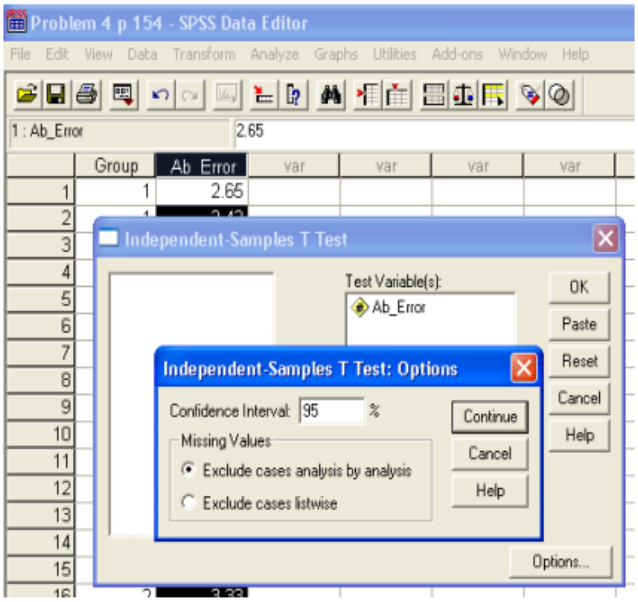


Test variable is the dependent variable

Grouping variable is the independent variable

Set the value of alpha.

Note that the program askes you to put the confidence interval. Simply if your alpha is 0.05, put 95%



Independent t-Test: Output

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Ab_Error	Active	10	2.2820	1.24438	.39351
	Passive	10	1.9660	1.50606	.47626

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Ab_Error	Equal variances assumed	.513	.483	.511	18	.615	.31600	.61780	-.98194	1.61394
	Equal variances not assumed			.511	17.382	.615	.31600	.61780	-.98526	1.61726

Are the groups different?

t(18) = .511, p = .615

NO DIFFERENCE

Very important to know these two things:

1 – during calculation, it is okay to get a negative value of T, treat it as positive and do the comparison.

2 – SPSS does the comparison by itself and gives another value called **P value (represented in the table as sig.).** we use this value to make our decision.

If the p value was larger than alpha → there is no difference

If p value was smaller than alpha → there is a difference.

Dependent or Paired t-Test: Output

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre	4.7000	10	2.11082	.66750
	Post	6.2000	10	2.85968	.90431

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pre & Post	10	.968	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Pre - Post	-1.50000	.97183	.30732	-2.19520	-.80480	-4.881	9	.00

Is there a difference between pre & post?

t(9) = -4.881, p = .001