7

Biostatistics

Doctor 2018 | Medicine | JU

◉ Sheet ○ Slides

Done By

Noor Adnan

Contributed in the Scientific Correction

Hala ALdaradka

Contributed in the Grammatical Correction

Hala ALdaradka

Doctor

Hamza Alduraidi

-> **T-test**, which is one of the parametric tests, is very similar in its steps to Chi-square, but there are some differences:

Chi- square	T-test
The variables are nominal	The dependent variables are continuous
So, we are comparing proportions **Example**: the proportion of women who use sunblock and have cancer ***versus*** the proportion of women who don't use sunblock and have cancer	So, we are calculating the means of the dependent variables
Calculations are done on the **sample** level	It gives prediction on the **population** level
Is one type only	Has two types: independent sample T-test and dependent sample T-test ***independent sample (between groups) T-test:** **when there are two separate groups** (example athletes and non-athletes). ***dependent sample (within group/ paired/ repeated measure) T-test:** **the same group have been tested twice** (example before and after certain intervention) **and each individual has a pair of data.**

Before using T-test, make sure that the **assumptions** have come true, which include that the dependent variable is continuous, the groups are randomly drawn from normally distributed populations and the variances between the two groups on the population level are similar. <do not worry a lot about this step in the exam, the question will state that the assumptions have come true 😊>

-> We start T-test by **stating the hypothesis**

Ha (alternative hypothesis): is the researcher hypothesis and it states that **μ1 ≠ μ2**, because we are assuming that the dependent variable that we are calculating its mean is different among the two groups we are studying.

Ho (null hypothesis): is the opposite of Ha and it states that **μ1 = μ2,** because it assumes that there is no difference between the means of the two groups. <that's what we are trying to reject>

Example: we are wondering if the mean heart rate among athletes (**a**) is different than the mean heart rate among non-athletes (**n**) in Jordanian men population. => Ha: μ**a** ≠ μ**n** Ho: μ**a** = μ**n**

-> Next, we must choose alpha, it is usually at **0.05**.

-> Calculate **T** (T $_{calc}$) using equations.

-> Compare it with the critical value of **T** (T $_{cv}$) using T table.

-> **If T $_{calc}$ > T $_{cv}$ --- REJECT the null hypothesis.**

If T $_{calc}$ < T $_{cv}$ ---- KEEP the null hypothesis = we can't find a difference between the means of the two groups.

Remember that the equations used to calculate T and degrees of freedom differ between the dependent and independent sample T-test.

The independent sample t- test

t calc = $$t_{\bar{X}_1 - \bar{X}_2} = \frac{\bar{X}_1 - \bar{X}_2}{SE_{diff}}$$

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

***SD² is the variance. N is the number of individuals in the group.**

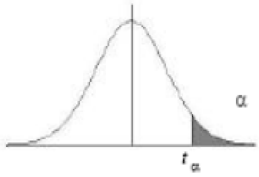
After calculating t, refer to t table on the right. The columns are alpha value which is **usually** 0.05. the rows are the degrees of freedom (df).

df= N1 + N2 – 2

example: if t $_{calc}$= 6.5, alpha= 0.05, Na= 16, N**n**=15 -> df= 29 then t $_{cv}$ = 1.7

t $_{calc}$ > t $_{cv}$ **REJECT the null hypothesis**

Table 4: Percentage Points of the *t* distribution



df	α 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

our interpretation is that: **We are 95% confident that the mean heart rate among athlete population is different than the mean heart rate among non-athlete population in Jordan (alpha= 0.05, t $_{calc}$ = 6.5)**.

The dependent sample t- test

Example: we have 12 students and we are wondering if their exam scores differ before (b) and after (a) the lecture.

Ha: μ**b** ≠ μ**a**   Ho: μ**b** = μ**a**   (alpha=0.05)

Calculate t:

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

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

D is the difference. In this example there are 12 students each with before and after score. We establish a table with three columns, the first for before, the second for after and the last for the difference (D / delta) between the two scores. Sum all the Ds, divide it by 12, this is $\overline{D}$

SD_D: standard deviation of the difference column.

N_{pairs}: number of individuals. (in our example 12)

After calculating t, refer to the previous t table, you need to calculate df: **df= N (pairs) – 1**

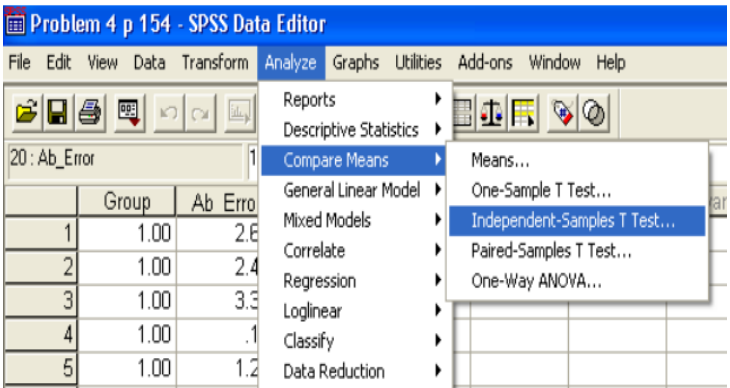
In our example: df= 12 – 1 = 11, t_{cv} = 1.8 let's assume the t_{calc} = 1.1 => $t_{calc} < t_{cv}$ **KEEP the null hypothesis**, the lecture wasn't successful in significantly improving students' scores in the exam.

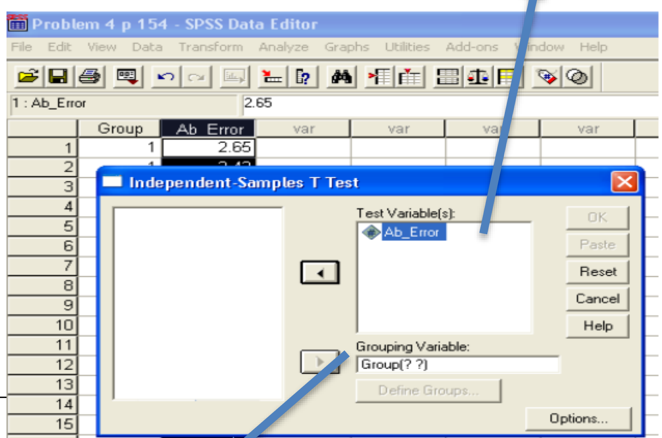
our interpretation is that: **We are 95% confident that the mean exam score among the Jordanian student population before is the same as the mean exam score among the Jordanian student population after the lecture. (alpha= 0.05, t_{calc} = 1.1)**

-> if alpha is 0.25 then t_{cv}= 0.697 -> **REJECT the null hypothesis.**

* SPSS *

enter the **dependent** (outcome) **variable**

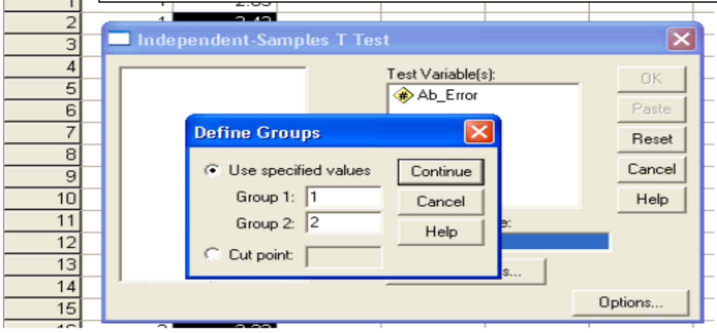




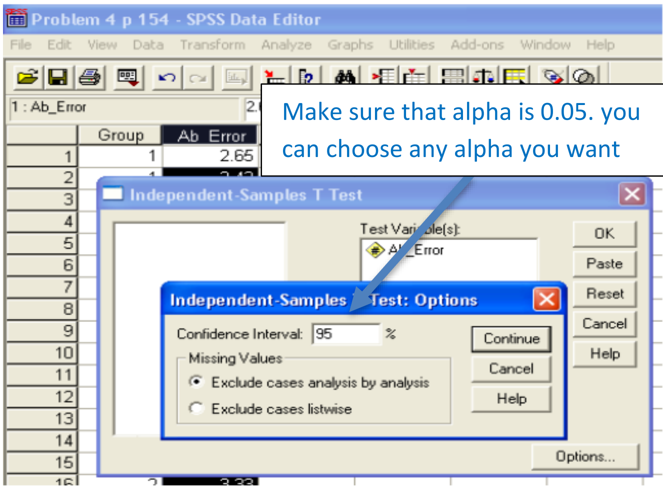
After entering the data, choose **analyze-> compare means -> independent-samples t-test**.

Enter the **independent variable** (it must be nominal)

After choosing define groups, this window will appear. Define each group by choosing a number for it, for example, **1** for athletes, **2** for non-athletes, and so on.



Make sure that alpha is 0.05. you can choose any alpha you want



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

Assuming an equal variance

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Ab_Error	Equal variances assumed	.513	.483	.511	18	.615	.31600	.617
	Equal variances not assumed			.511	17.382	.615	.31600	.617

Assuming an unequal variance (not required)

***the table below is more important than the table above always.**

-> in this example, there is two groups (active, passive), in the first table we have N, mean, SD and Std. error mean for each group. (df= 10 +10 – 2 = 18)

-> in the second table, what we really care about is the three values in the red box,

1) t $_{calc}$ 2) df 3) sig.

here t $_{calc}$= 0.511, df= 18, sig (p value) =0.615

sig >>> alpha -> accept Ho "there is no difference between the two groups".

For dependent (paired) sample t – test, apply the same steps but choose **paired-samples t test** instead of independent-samples t test then enter the two sets of observations (pre-post). After completing all the steps, three tables will appear ->

All the descriptive information is here for each group

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

This is the D table. (not required)

Paired Samples Correlations

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

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Pre - Post	-1.50000	.97183	.30732	-2.19520	-.80480	-4.881	9	.001

t $_{calc}$= 4.881, df= 9, sig= 0.001.

alpha >>>> sig REJECT Ho "the population mean before is different than the mean after on the population level"

***sometimes sig= 0.000 that means reject Ho**

Important notes:

1- independent random samples:

**the samples are randomly and independently selected from normally distributed populations.

** the variances of the population of $\sigma^2{}_1$and $\sigma^2{}_2$ are equal or nearly equal to ensure that the procedures are valid.

2- dependent random samples:

** Ha:$\mu d \neq 0$

** df = n -1

n= number of paired differences.

و ما معنى الوقوع ان كنت تجيد النهوض !!!

Eğer kalmayı biliyorsa düşmek ne ki !!