People's Democratic Republic of Algeria
University Med Khider of Biskra
Faculty of SNVSTU

Protocol 03- Hypothesis testing

Dr. Ben Gherbal Hanane

Data Analysis in Biosciences — Level: L3 biology
Email: hanane.bengherbal@univ-biskra.dz

There are several types of parametric tests designed to test certain hypotheses related to one or more parameters of a specific probability distribution. These tests are mainly grouped into two categories: conformity tests and homogeneity tests. We illustrate the steps in SPSS to perform a conformity (Goodness of fit) test for a mean and a homogeneity test between two means, using the following examples:

Example 1 (Conformity test of observed and theoretical mean). *To study a batch of tablets, 10 tablets were randomly selected and weighed. The weights (in grams) were:*

0.81	*0.84*	*0.83*	*0.80*	*0.85*	*0.86*	*0.85*	*0.83*	*0.84*	*0.80*

Is the observed mean compatible with the value $0.83g$ *at a 0.02 significance level?*
Hypotheses:

$$H_0 : \mu = \mu_0 \quad vs. \quad H_1 : \mu \neq \mu_0$$

SPSS Procedure:

1. Launch SPSS and enter data via **Variable View** and **Data View**.
2. From the top menu: **Analyze** → **Compare Means** → **One-Sample T Test**.
3. Select the variable `weight`, set the test value to 0.83.
4. Under `Options`, set confidence level to 98% ($1 - \alpha = 0.98$).
5. Click **Continue**, then **OK** to get results.

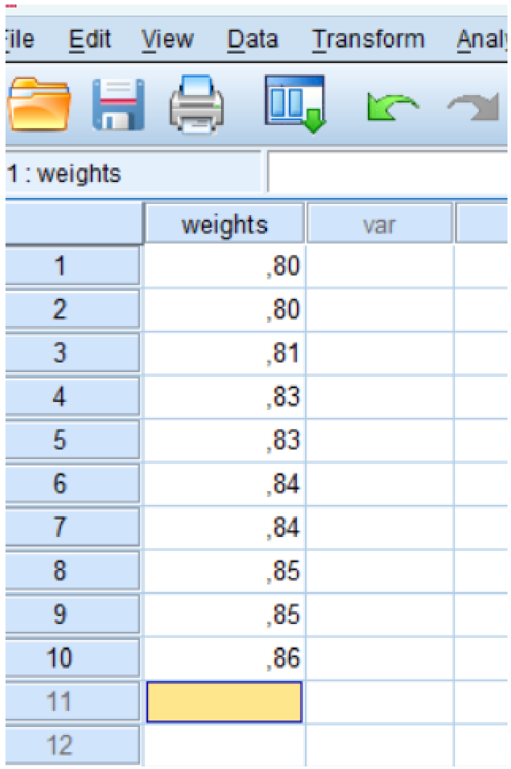

Figure 1: Data entry in SPSS

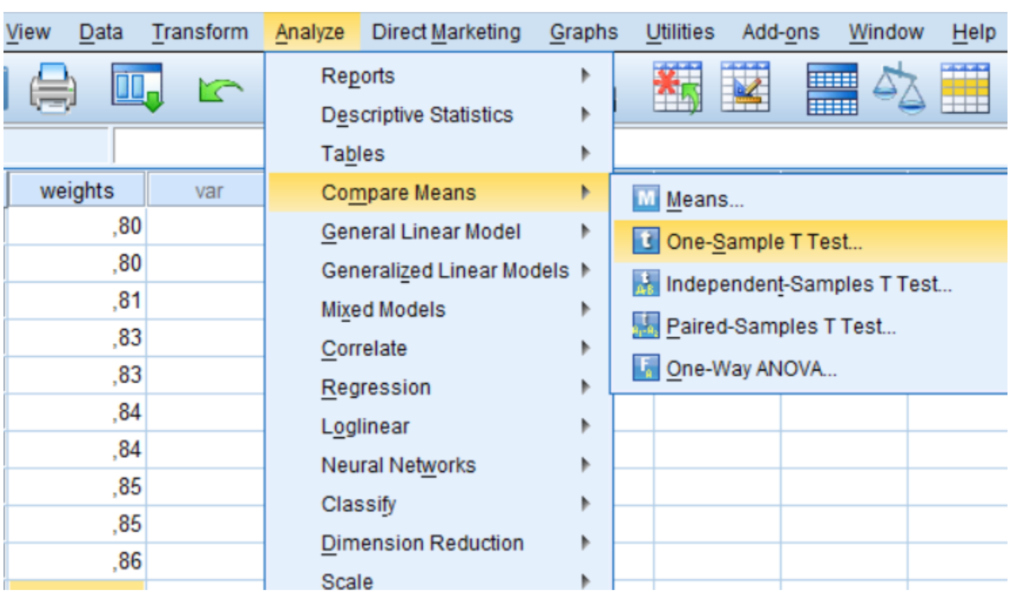

Figure 2: Performing one-sample t-test in SPSS

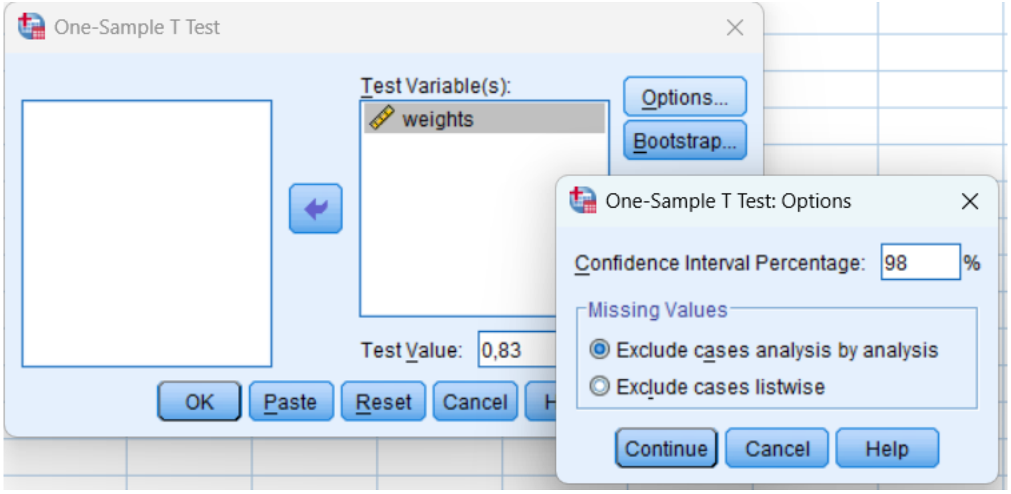

Figure 3: Performing one-sample t-test in SPSS

Interpretation: For a two-tailed test:

$$\begin{cases} \text{Do not reject } H_0 \text{ if } \alpha < \text{p-value} \\ \text{Reject } H_0 \text{ if } \alpha > \text{p-value} \end{cases}$$

In our case, $\alpha = 0.02 <$ p-value $= 0.885$, so we do not reject H_0.

T-Test

[DataSet0]

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
weights	10	,8310	,02132	,00674

One-Sample Test

	Test Value = 0.83					
					98% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
weights	,148	9	,885	,00100	-,0180	,0200

Figure 4: Results window of the one-sample t-test in SPSS

Remark 1. *In the case of a one-tailed test (i.e., when $H_1 : \mu < \mu_0$ or $H_1 : \mu > \mu_0$), the decision to reject or not reject the null hypothesis H_0 is made as follows:*

- *Do **not** reject H_0 if $2\alpha <$ significance (i.e., the mean is not significantly different from μ_0);*
- ***Reject** H_0 if $2\alpha >$ significance (i.e., the mean is significantly greater than μ_0 if $H_1 : \mu > \mu_0$, or significantly less than μ_0 if $H_1 : \mu < \mu_0$).*

Example 2 (Comparing calcium levels between two samples). *The following table presents calcium levels in patients suffering from chronic renal failure :*

Patient	*Sample 1*	*Sample 2*
1	*50*	*80*
2	*25*	*130*
3	*120*	*70*
4	*26*	*120*
5	*48*	*70*
6	*113*	*100*
7	*150*	—

Can we conclude, at a 5% significance level, that the means are significantly different?
Hypotheses:

$$H_0 : \mu_1 = \mu_2 \quad vs. \quad H_1 : \mu_1 \neq \mu_2$$

SPSS Procedure: The data from the two samples must be entered as a single variable (i.e., as one combined sample), and a second variable should be added to indicate the sample number from which each observation (X_i) comes. Thus, for the data in our example, we obtain the following:

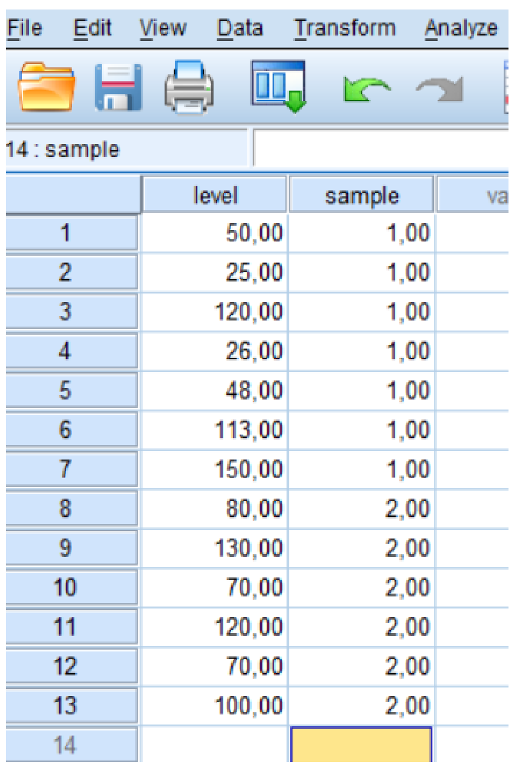

Figure 5: SPSS data format for independent samples t-test

1. Enter data as a single variable with a group identifier.
2. Use `Analyze` → `Compare Means` → `Independent-Samples T Test`.

3. Place the variable `sample` in the box labeled *"Grouping criterion (numeric qualitative)"*.

4. Select the variable `sample` and click on the **Define Groups** button, a small new window will appear.

5. In the *Group* boxes, enter the numbers of the two samples you wish to compare, in our case, "1" for sample 1 and "2" for sample 2.

6. Once the groups are defined, click on **Continue**, then click on **OK** to display the test results.

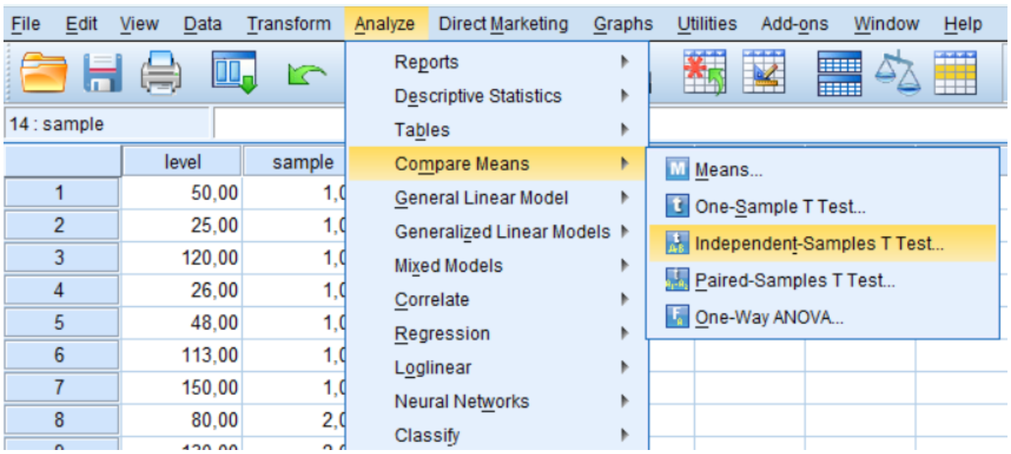

Figure 6: Selecting independent samples t-test in SPSS

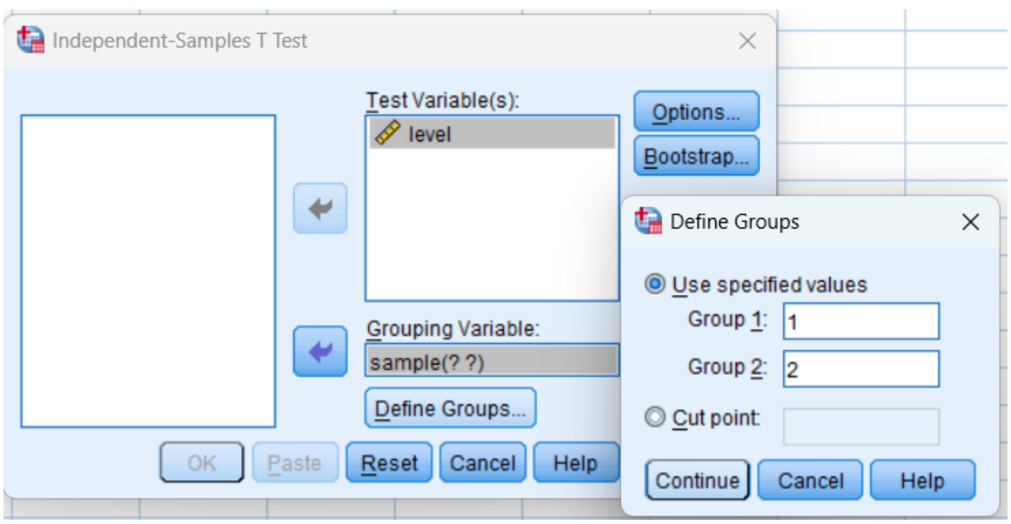

Figure 7: Running the test in SPSS

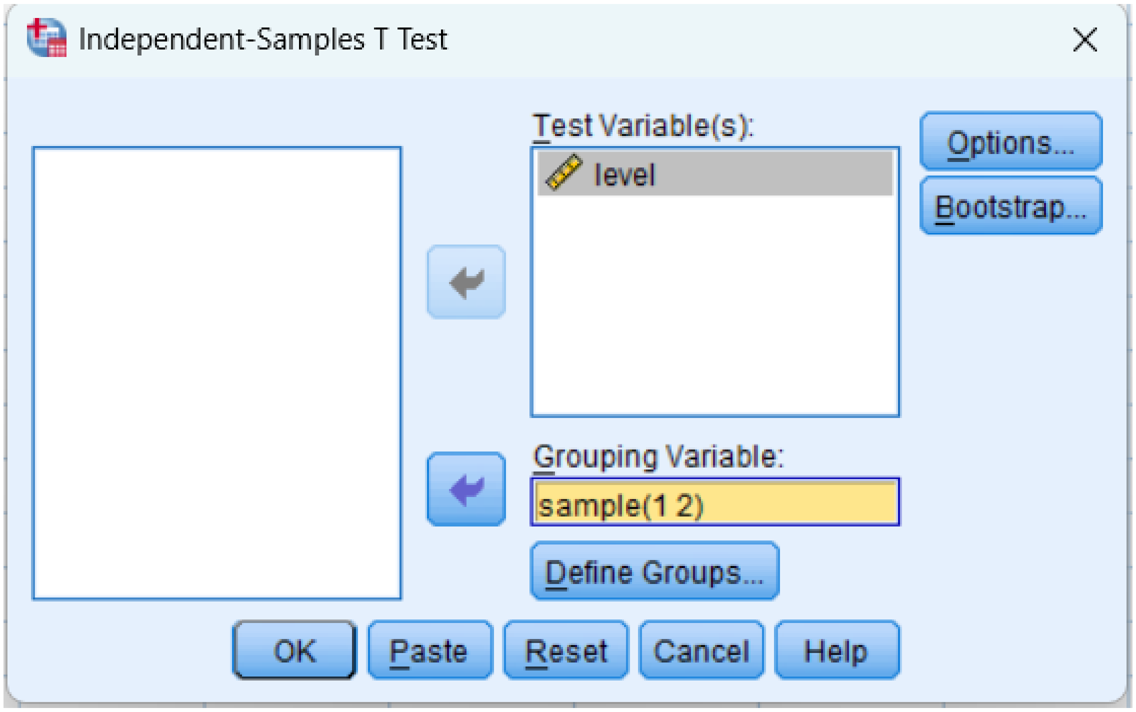

Figure 8: Result window comparing two means in SPSS

Interpretation:
If $\alpha <$ sig1: use sig2 (equal variances)

- Do not reject H_0 if $\alpha <$ sig2
- Reject H_0 if $\alpha \geq$ sig2

If $\alpha \geq$ sig1: use sig3 (unequal variances)

- Do not reject H_0 if $\alpha <$ sig3
- Reject H_0 if $\alpha \geq$ sig3

From the results: $\alpha = 0.05 \geq 0.014$ and $0.05 < 0.406$, then no significant difference (accept H_0)

Group Statistics

	sample	N	Mean	Std. Deviation	Std. Error Mean
level	1,00	7	76,0000	50,56679	19,11245
	2,00	6	95,0000	25,88436	10,56724

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
level	Equal variances assumed	8,470	1 ,014	-,828	11	2 ,425	-19,00000	22,93393	-69,47725	31,47725
	Equal variances not assumed			-,870	9,198	3 ,406	-19,00000	21,83924	-68,24244	30,24244

Figure 9: Result window comparing two means in SPSS

Remark 2. *In the context of a one-tailed test, and for a significance level α, the decision to reject or not reject the null hypothesis H_0 is made according to the following scheme:*

- *If $\alpha < sig1$, then the two variances are not significantly different (i.e., they are equal). In this case, to make the decision concerning the two means, one must use the significance value $sig2$:*

  1. *Do not reject H_0 if $2\alpha < sig2$ (the two means are not significantly different).*
  2. *Reject H_0 if $2\alpha \geq sig2$ ($H_1 : \mu_1 < \mu_2$ or $H_1 : \mu_1 > \mu_2$).*

- *If $\alpha \geq sig1$, then the two variances are significantly different. In this case, to make the decision concerning the two means, one must use the significance value $sig3$:*

  1. *Do not reject H_0 if $2\alpha < sig3$ (the two means are not significantly different).*
  2. *Reject H_0 if $2\alpha \geq sig3$ ($H_1 : \mu_1 < \mu_2$ or $H_1 : \mu_1 > \mu_2$).*

Exercises (Solve the exercises manually, then confirm the results using SPSS)

Example 3. *The weight of grapes produced per vine was measured for 10 vines randomly selected from a vineyard. The obtained results, expressed in kilograms, are as follows:*

2.4	3.4	3.6	4.1	4.3	4.7	5.4	5.9	6.5	6.9

Assuming $\mathcal{N}(\mu, \sigma)$, test whether mean < 8 at 5% level.

Example 4. *In order to compare two types of trees with respect to their heights, we collected height measurements for several trees, which are presented in the following table.*

Tree 1	23.3	24	24.3	24.5	25	25.9
Tree 2	21.1	21.1	22.1	22.4	23.3	—

Assuming $\sigma_1 = \sigma_2$, test at 1% if Tree 2 mean is lower.

Example 5. *We want to compare two teaching methods for reading. For this purpose, two groups of children were formed: the first group was taught using the so-called classical method, and the second group using the modern method. The scores obtained at the end of the school year are as follows:*

Group 1	15	18	17	18	17	19	16	17	14	19
Group 2	20	16	19	16	18	17	16	20	21	17

1. *Is the classical method's mean compatible with 6 at 1% risk?*
2. *Assuming $\sigma_1 = \sigma_2$, test at 5% whether modern method yields better results.*