◉ Sheet ○ Slides

Number

13

Done by:

Omar Ashour

corrected by:

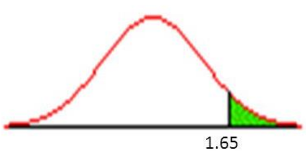
Doctor

Mahmoud Al-Hussami

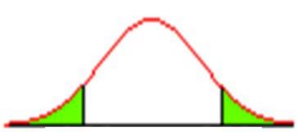
t-test

Once sample data has been gathered through an observational study or experiment, statistical inference allows analysts to assess evidence in favor of some claim about the population from which the sample has been drawn. The methods of inference used to support or reject claims based on sample data are known as **tests of significance**. There are <u>six</u> steps for significance testing:

1. **Set alpha**
   - Recall that **α** has 3 readings
     - The minimum level being 0.05
     - The intermediate level usually 0.01
     - The maximum level set at 0.001
2. **State hypotheses**
   - Null Hypothesis $\boldsymbol{H_0}$ (statistical hypothesis)
     - States that there is no difference between the variables tested.
   - Alternative Hypothesis $\boldsymbol{H_1}$ (research hypothesis)
     - Directional alternative hypotheses are used when anticipating the "direction" of a relationship. It states the researcher's expectation regarding whether one variable is going to be higher or lower than the other variable. This type can be represented as the area under, either the right or left side of the normal distribution curve.



     - Non-directional alternative hypotheses are merely used when predicting the "presence" of a relationship, without anticipating the direction of a relationship. This type can be represented as the area under both sides of the normal distribution curve.



<u>***For example***</u>:

- *Non-directional*: I predict that sleeping in lectures and grades will be significantly **related**.
- *Directional*: I predict that, as sleeping in lectures **increases**, grades will **decrease**.

3. **Calculate the test statistic (*t*)**

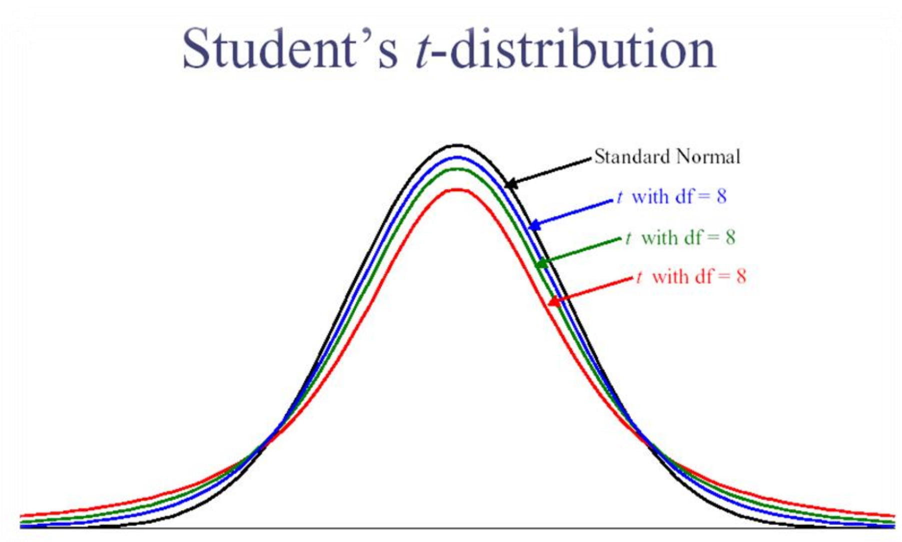
4. **Find the critical value of the statistic from the given *t* table**

5. **State the decision rule**

- If SPSS is used, significance will be determined by comparing the P value to **α**
  - If P > **α** we **fail to reject** the null hypothesis
  - If P < **α** we **reject** the null hypothesis
- If done manually, *t* is calculated and compared to the critical value from the table
  - If *t* > critical value we **reject** the null hypothesis
  - If *t* < critical value we **fail to reject** the null hypothesis

6. **State the conclusion**

The *t*-test is a useful technique for comparing **mean** values of two sets of numbers. It can be used to determine if two sets of data are significantly **different** from each other. The *t*-test is applied when the calculated test statistic (*t)* follows a modified form of the normal distribution known as the *t*-distribution. The *t*-distribution depends on the sample size and varies according to the degrees of freedom of different samples.



Assumptions of the t-test:

1. Dependent variable should be continuous (I/R)

2. The groups should be randomly drawn from normally distributed and independent populations

3. The independent variable should be categorical with two levels

4. Distribution for the two independent variables should be normal

5. Variance should be equal (homogeneity of variance)

6. Small variation is preferred because a large variation would less likely lead to a significant *t*-test; thus, accepting the null hypothesis (fail to reject) resulting in Type II error and less power.

There are 3 types of *t*-tests:

1. The **one**-sample *t*-test is used to compare a single sample with a known standard population value. For example, a test could be conducted to compare the average salary of surgeons within a hospital with a value that was known to represent the national average for surgeons.

Exercise:

Let us assume that you want to buy lightbulbs from a specific retail store in Amman. Now, in order to test out the efficiency of the light bulbs you compare their lifetime with the **standard Jordanian** average life of 1000 hours. You set up a *t*-test to make sure that the lightbulbs are made in Jordan.

> The null hypothesis always states that the means are equal and that there is no difference.

Testing whether light bulbs have a life of 1000 hours

1. Set alpha. $\alpha = .05$
2. State hypotheses.
   - Null hypothesis is H_0: $\mu = 1000$.
   - Alternative hypothesis is H_1: $\mu \neq 1000$.
3. Calculate the test statistic

Calculating the Single Sample t

800
750
940
970
790
980
820
760
1000
860

What is the mean of our sample?

$$\bar{X} = 867$$

What is the standard deviation for our sample of light bulbs?

SD= 96.73

$$SE = \frac{SD}{\sqrt{N}} = \frac{96.73}{\sqrt{10}} = 30.59$$

$$t_{\bar{X}} = \frac{\bar{X} - \mu}{S_{\bar{X}}} = \frac{867 - 1000}{30.59} = -4.35$$

In order to calculate the one-sample (*t*), we should first find the **mean** and **standard deviation** of the sample. Then we use the standard deviation to calculate the **standard error**. After that, we subtract the population mean from the sample mean and then divide it by the standard error.

<u>Notice that the standard deviation here is pretty large; this indicates large variation within the sample which may result in a less powerful test. Also, note that we take the absolute value of the calculated (*t*).</u>

$$t_{\bar{X}} = \frac{\bar{X} - \mu}{S_{\bar{X}}}$$

*Because we are doing this test manually, we calculate the value of **t** and extract the critical value from the table. We also use the decision rules mentioned earlier.*

We calculate the degree of freedom through (n – 1) n being the sample size

(n=10 in the exercise)

then we use the table below to find the critical value:

df	0.20	0.10	0.05	0.02	0.01	0.001
1	3.0777	6.3138	12.7062	31.8205	63.6567	636.6192
2	1.8856	2.9200	4.3027	6.9646	9.9248	31.5991
3	1.6377	2.3534	3.1824	4.5407	5.8409	12.9240
4	1.5332	2.1318	2.7764	3.7469	4.6041	8.6103
5	1.4759	2.0150	2.5706	3.3649	4.0321	6.8688
6	1.4398	1.9432	2.4469	3.1427	3.7074	5.9588
7	1.4149	1.8946	2.3646	2.9980	3.4995	5.4079
8	1.3968	1.8595	2.3060	2.8965	3.3554	5.0413
9	1.3830	1.8331	2.2622	2.8214	3.2498	4.7809
10	1.3722	1.8125	2.2281	2.7638	3.1693	4.5869
11	1.3634	1.7959	2.2010	2.7181	3.1058	4.4370
12	1.3562	1.7823	2.1788	2.6810	3.0545	4.3178
13	1.3502	1.7709	2.1604	2.6503	3.0123	4.2208

4. Determine the critical value. Look up in the table (Munro, p. 451). Looking for alpha = .05, two tails with df = 10-1 = 9. Table says 2.262.

Table A.3 Values for Student's *t* Distribution

	Two-tailed test		
df	.10	.05	.01
1	6.314	12.706	63.657
2	2.920	4.303	9.925
3	2.353	3.182	5.841
4	2.132	2.776	4.604
5	2.015	2.571	4.032
6	1.943	2.447	3.707
7	1.895	2.365	3.499
8	1.860	2.306	3.355
9	1.833	2.262	3.250
10	1.812	2.228	3.169
11	1.796	2.201	3.106
12	1.782	2.179	3.055
13	1.771	2.160	3.012
14	1.761	2.145	2.977
15	1.753	2.131	2.947
16	1.746	2.120	2.921
17	1.740	2.110	2.898
18	1.734	2.101	2.878
19	1.729	2.093	2.861
20	1.725	2.086	2.845

t Values

- Critical value decreases if N is increased. ★
- Critical value increases if alpha is increased. ★
- Differences between the means will not have to be as large to find sig if N is large or alpha is increased.

5. State decision rule. If absolute value of sample is greater than critical value, reject null.

If $|-4.35| > |2.262|$, reject H_0.

Stating the Conclusion

6. State the conclusion. We reject the null hypothesis that the bulbs were drawn from a population in which the average life is 1000 hrs. The difference between our sample mean (867) and the mean of the population (1000) is SO different that it is unlikely that our sample could have been drawn from a population with an average life of 1000 hours.

Obviously, the sample lightbulb lifetimes are not the same as those of the Jordanian standard population. Therefore, you conclude that they are imported from another country.

SPSS Results

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
BULBLIFE	10	867.0000	96.7299	30.5887

One-Sample Test

	Test Value = 1000					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
BULBLIFE	-4.348	9	.002	-133.0000	-202.1964	-63.8036

SPSS uses P values. Here it is calculated it to be 0.02. Now, 0.02 is less than our alpha (0.05) thus, we reject the null hypothesis.

---

2. The **independent**-sample *t*-test is used to compare two groups' scores on the same variable. For example, it could be used to compare the salaries of dentists and physicians to evaluate whether there is a difference in their salaries.

Independent Samples *t*-test

- Used when we have two independent samples, e.g., treatment and control groups.
- Formula is: $t_{\bar{X}_1-\bar{X}_2} = \dfrac{\bar{X}_1 - \bar{X}_2}{SE_{diff}}$
- 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:

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

Exercise:

Suppose we want to study the effect of caffeine on human motor activity. The task is to keep the mouse of a computer centered on a moving dot. The faster the response the better activity. Now, everyone gets a drink; half get caffeine, half get placebo; nobody knows who got what.

Experimental (Caff)	Control (No Caffeine)
12	21
14	18
10	14
8	20
16	11
5	19
3	8
9	12
11	13
	15
N_1=9, M_1=9.778, SD_1=4.1164	N_2=10, M_2=15.1, SD_2=4.2805

The data collected represents the amount of time it took the participants to complete the task.

1. Set alpha. Alpha = .05
2. State Hypotheses.

   Null is H_0: $\mu_1 = \mu_2$.

   Alternative is H_1: $\mu_1 \neq \mu_2$.

Remember: the steps here are only valid when the test is done manually. When using SPSS, the P value is calculated and compared with the cutoff **α** value.

3. Calculate test statistic:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{SE_{diff}} = \frac{9.778 - 15.1}{1.93} = \frac{-5.322}{1.93} = -2.758$$

$$SE_{diff} = \sqrt{\frac{SD_1^2}{N_1} + \frac{SD_2^2}{N_2}} = \sqrt{\frac{(4.1164)^2}{9} + \frac{(4.2805)^2}{10}} = 1.93$$

We calculate the test statistic (*t*) by using the same steps in the previous example but with the independent 2 group equations shown above.

4. Determine the critical value. Alpha is .05, 2 tails, and df = N1+N2-2 or 10+9-2 = 17. The value is 2.11.
5. State decision rule. If |-2.758| > 2.11, then reject the null.
6. Conclusion: Reject the null. the population means are different. Caffeine has an effect on the motor pursuit task.

The degree of freedom for each group is (n-1).
Caff group (9-1) = 8
No Caff (10-1) = 9

The test depends on both groups so the total is 9 + 8 = **17**

Recall that equal variance is required in order to validate the results of the *t*-test. The independent *t*-test measures the difference between the means of 2 groups. When using SPSS, a test called "Levene's Test for Equality of Variances" is used to evaluate the homogeneity between the 2 groups of the independent *t*-test. Levene's test calculates a specific value called *Levene's P*, which is used as follows:

- If Levene's P is greater than the alpha significance level (usually 0.05) then the **variances of both groups are equal.**
- If Levene's P is less than the alpha significance level (usually 0.05) then the **variances of both groups are unequal** and the *t*-test results would not be significant.

Therefore, when using SPSS, Levene's test must be computed in order to generate significant results when testing for significance. Equal variance is necessary because the difference will be a **result of the means of the groups** *(that's what we want)*. When variance is unequal and there is heterogeneity, the difference will be a **result of the variability of the groups**; thus, the *t*-test would be insignificant and invalid.

---

3. The **dependent** *t*-test is used to compare the means of two variables within a single group. It is useful to control individual differences and can result in a more powerful test than independent samples *t*-test. For example, it could be used to see if there is a statistically significant difference between starting salaries and current salaries among the general physicians in a single hospital.

Formulas:

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

t 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}.

<-- These formulas are needed for the calculation of the dependent test.

D denotes the difference of the means of the sample.

The equation below is a combination of both formulas on the left, and can be used to calculate *t* in one step.

$$t_{\overline{X}_D} = \frac{\overline{D}}{SD_D \big/ \sqrt{n_{pairs}}}$$

Exercise:

Let's assume that you want to test the effect of a pain relieving drug on a *single* group of patients. The experimental design requires you to give each patient the drug and a placebo. Data is then collected by measuring the time for which the painkilling effect lasts for each patient.

Person	Painfree (time in sec)	Placebo	Difference
1	60	55	5
2	35	20	15
3	70	60	10
4	50	45	5
5	60	60	0
M	55	48	7
SD	13.23	16.81	5.70

1. Set alpha = .05
2. Null hypothesis: H_0: $\mu_1 = \mu_2$.
   Alternative is H_1: $\mu_1 \neq \mu_2$.
3. Calculate the test statistic:

$$SE_{diff} = \frac{SD}{\sqrt{n_{pairs}}} = \frac{5.70}{\sqrt{5}} = 2.55$$

$$t = \frac{\overline{D}}{SE_{diff}} = \frac{55-48}{2.55} = \frac{7}{2.55} = 2.75$$

4. Determine the critical value of t.
   Alpha =.05, tails=2
   df = N(pairs)-1 =5-1=4.
   Critical value is 2.776
5. Decision rule: is absolute value of sample value larger than critical value?
6. Conclusion. Not (quite) significant.
   Painfree does <u>not</u> have an effect.

Please check the PPT slides for images of the SPSS interface.