T test

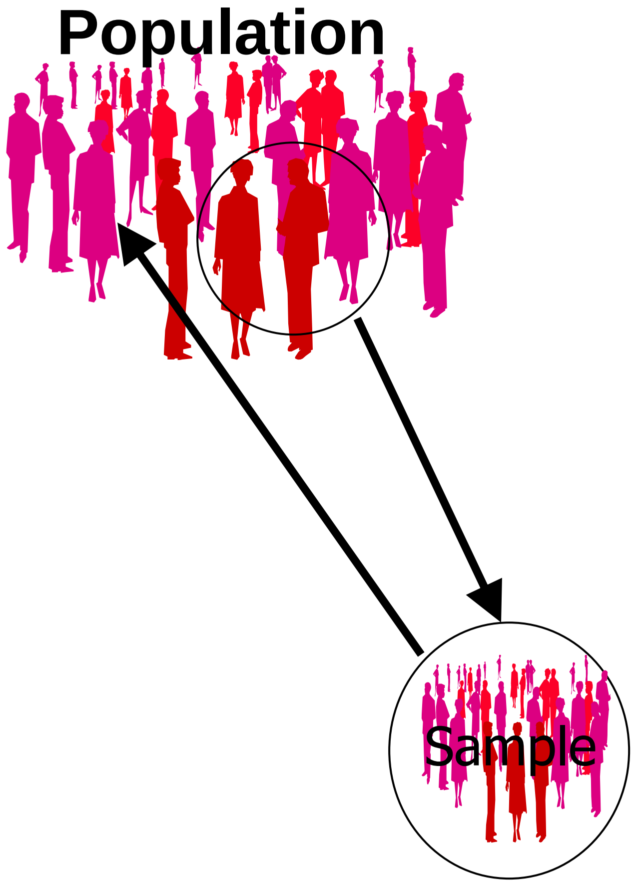
Prof. Dr Sami Abdo Radman

Types of *T test : (Students t test*

- **Two sample t-test (independent sample t test)**
- ***Paired T test (dependent sample t-test***
- ***One sample T test***

Inferential Statistics

- Two ways to generalize from samples to populations
  - **Estimation of parameters (Confidence Interval, CI)**
  - **Hypothesis testing (Test of significance, p value)**
- Purpose
  - Make decisions about population characteristics



Two sample t-test (independent sample t test)

- Hypothesis testing (Tests of statistical significance)
- Difference in mean

- Variables:

✓Continues variable
✓Grouping variable

- A significance test uses data from a sample to show the likelihood that a hypothesis about a population is true

eg

- T test answers Questions like:
- Is there difference in mean cholesterol level between smokers and non smokers?

✓Test variable: cholesterol level **(continuous variable)**

✓Groups : smokers , non-smokers **(categorical)**

- Hypothesis: there is difference in cholesterol level between the two groups

(μ 1$\neq$ μ 2

- Null hypothesis : there is no difference in cholesterol level between the two groups (μ1 = μ 2)

eg

- Is there difference in mean hemoglobin level between urban and rural children?

- Test variable: hemoglobin level continuous
- Groups : urban and rural children categorical

- Hypothesis: there is difference in hemoglobin level between the two groups

($\mu 1 \neq \mu 2$

- Null hypothesis : there is no difference in hemoglobin level between the two groups ($\mu 1 = \mu 2$)

Eg

- does a new treatment reduce blood pressure more than an existing treatment?

- The null hypothesis: mean blood pressure is the same in the

two treatment groups (no difference)

- The alternative hypothesis is that mean blood pressure is different in the two treatment groups (there is difference)

Statistical test :

- Statistical test :

is there a real difference or the difference is due to chance ??

- Statistical test

we can decide to reject or accept the null hypothesis → decide to accept or reject the hypothesis

- P value < 0.05 → reject the null → accept the hypothesis
- P value > 0.05 → accept the null → reject the hypothesis

- P value < 0.05 → reject the null → accept the hypothesis →
There is statistically significant difference

- P value ≥0.05 → accept the null → reject the hypothesis →
- There is no statistically significant difference

- P value is the probability that the null hypothesis is true
- P value indicate wether the difference is a reall difference or due to chance

Steps in doing a statistical test

- 1. Specify the hypothesis of interest as a null and alternative hypothesis.
- 2.Collect data and enter data to software (eg. SPSS)
- 2. Decide what statistical test is appropriate.
- 3. Use the test to calculate the P value.
- 4. accept/reject the null → accept/reject the hypothesis
- Write conclusion

Two sample t-test (independent sample t test)

- Compare mean between two independent groups
- eg:
- Compare mean SBP between two independent groups (males and females)

- Hypothesis: there is difference in SBP between the two groups

(μ 1$\neq$ μ 2

- Null hypothesis : there is no difference in SBP between the two groups (μ1 = μ 2)

Assumptions

- Dependent variable is continuous
- Two groups are independent
- Normal distribution of SBP
- Homogeneity of variance (variances are equal)

Assumptions: normality

- The dependent variable must be continuous and normally distributed

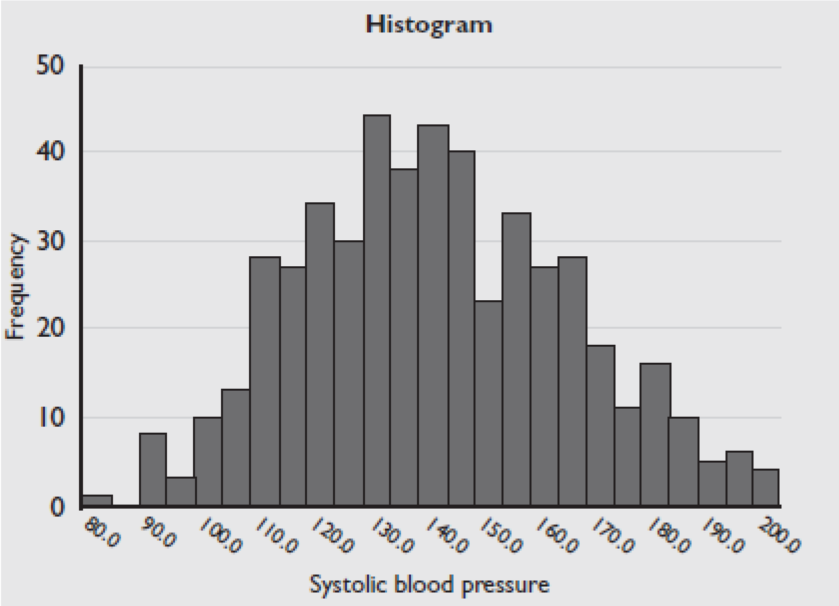


Table IX. Normality assumption checks.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Systolic BP before	.048	167	.200	.991	167	.388

Should be > 0.05

If p value > 0.05→ normally distributed

If p value <0.05→not normally distributed

Assumptions.

Homogeneity of variance ***(The population variances are equal).*** The Levene's Test for equality of variances

- The Levene's Test for equality of variances ;
- The Null hypothesis is: Equal Variances assumed (no differences between variances)
- the hypothesis is : there is difference between variance (variances are not equal)

Table XIV. 2 Sample T test.

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	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	131.183	.000	27.638	498	.000	44.5280	1.61111	41.36258	47.69342
Equal variances not assumed			27.638	353.465	.000	44.5280	1.61111	41.35943	47.69657

- Assumptions.

  Homogeneity of variance ***(The population variances are equal).*** The Levene's Test for equality of variances

- We want p value to be not significant
- If p value not sig. (>0.05) → variances are equal (variances are homogenous)
- If p is sig. (<0.05) → variances are not equal
- The Sig value (given in the 3rd column) shows that the p<0.05→ variances are not equal → assumption is violated

Table XIV. 2 Sample T test.

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
Equal variances assumed	131.183	.000	27.638	498	.000	44.5280	1.61111	41.36258	47.69342
Equal variances not assumed			27.638	353.465	.000	44.5280	1.61111	41.35943	47.69657

- If the p value in the third column is significant → there is differences in variance → the assumption is violated
- So we should go the other option → our p value will be in the last row

Table XIV. 2 Sample T test.

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	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	131.183	.000	27.638	498	.000	44.5280	1.61111	41.36258	47.69342
Equal variances not assumed			27.638	353.465	.000	44.5280	1.61111	41.35943	47.69657

Example

- Depression score was measured in males (n=17) and females (n=29).
- Is there difference in mean depression score between males and females?

- Hypothesis: there is difference
- Null: there is no difference

- We will conduct independent t test:

depression	gender
3.00	1.00
3.00	2.00
5.00	1.00
2.00	1.00
3.00	1.00
4.00	1.00
7.00	2.00
2.00	1.00
4.00	2.00
7.00	1.00
5.00	2.00
3.00	1.00
4.00	1.00
8.00	2.00
7.00	2.00

Normality

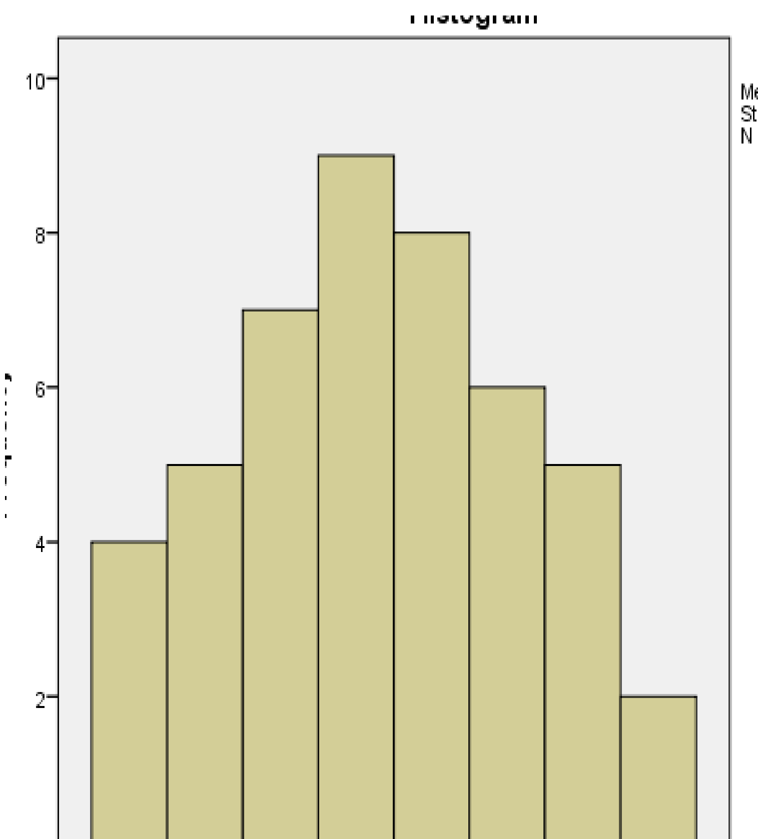
P value > 0.05→ depression score is normally distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
depression	.106	46	.291*	.958	46	.095

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction



Depression score among males and females

Levene's Test for Equality of Variances Equal variances were assumed (p=0.094)

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
depression	Equal variances assumed	2.933	.094	-2.052	44	.046	-1.14620	.55855	-2.27189	-.02051
	Equal variances not assumed			-1.944	31.037	.061	-1.14620	.58963	-2.34869	.05630

Depression score among males and females

Group Statistics

	gender	N	Mean	Std. Deviation	Std. Error Mean
depression	male	19	4.6316	2.19116	.50269
	female	27	5.7778	1.60128	.30817

Is there difference in mean depression score between males and females

Group Statistics

	gender	N	Mean	Std. Deviation	Std. Error Mean
depression	male	19	4.6316	2.19116	.50269
	female	27	5.7778	1.60128	.30817

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
depression	Equal variances assumed	2.933	.094	-2.052	44	.046	-1.14620	.55855	-2.27189	-.02051
	Equal variances not assumed			-1.944	31.037	.061	-1.14620	.58963	-2.34869	.05630

)

p value = 0.046

Conclusion

- P value is significant (p=0.046)→ reject the null hypothesis → accept the hypothesis
- There is a statistically significant difference in mean depression score between males (4.6±2.19) and females (5.77±1.6), (p=0.046)
- the mean difference is -1.136 with a 95%CI -2.27 , -.020
-

Group Statistics

	gender	N	Mean	Std. Deviation	Std. Error Mean
depression	male	19	4.6316	2.19116	.50269
	female	27	5.7778	1.60128	.30817

- What about the 95%CI of the difference?
- Mean difference is -1.146 with a 95%CI -2.27 , -0.020 (not include zero)

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
depression	Equal variances assumed	2.933	.094	-2.052	44	.046	-1.14620	.55855	-2.27189	-.02051
	Equal variances not assumed			-1.944	31.037	.061	-1.14620	.58963	-2.34869	.05630

Table

Association between socio-demographic factors and depression

	Depression		
	Mean	**SD**	**P value**
Gender			
Male	4.63	2.19	
Female	5.77	1.6	0.046

- What does degree of freedom (df) tell us

- df = sample size – 2 (number of groups)
- df= (n1 + n2) – 2 or : (n1-1) + (n2-1)
- If df = 44 → sample size = 46

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
depression	Equal variances assumed	2.933	.094	-2.052	44	.046	-1.14620	.55855	-2.27189	-.02051
	Equal variances not assumed			-1.944	31.037	.061	-1.14620	.58963	-2.34869	.05630

- What is t statistic :
- It is the statistic of t test and it is used with degree of freedom to compute p value

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	95% Confidence Interval of the Difference	
									Lower	Upper
depression	Equal variances assumed	2.933	.094	-2.052	44	.046	-1.14620	.55855	-2.27189	-.02051
	Equal variances not assumed			-1.944	31.037	.061	-1.14620	.58963	-2.34869	.05630

$$t = \frac{\textit{difference in means}}{SE(\textit{difference})} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SD_p^2}{n_1} + \frac{SD_p^2}{n_2}}}$$

where $\bar{X}_1, \bar{X}_2$ are the means, SD_p is the pooled standard deviation calculated from the group SDs, SD_1 and SD_2 (see following equation), and n_1,n_2 are the totals in the two groups.

$$SD_p = \sqrt{\frac{(n_1 - 1)\,SD_1^2 + (n_2 - 1)\,SD_2^2}{n_1 + n_2 - 2}}$$

t follows a Student's t distribution with $n_1 + n_2 - 2$ degrees of freedom. P values are obtained from tabulated values of the t distribution or a computer program.

- ***Paired T test***

Paired T test

- It analyses mean difference in a paired sample
- Used to compare mean before and after :
- for example In a group of patients, SBP was measured today and then after three weeks
- There is one sample (one group) and two means (before and after)

- Hypothesis
- The mean change is not equal zero (there is change) (there is difference) (mean before ≠ mean after)

- Null hypothesis The mean change or difference is zero (there is no difference) (there is no change) (mean before = mean after)

Assumption for the Paired T test:

- ***Assumption for the Paired T test:***

- Groups are dependent
- The data are continuous and normally distributed

- *The differences between the before & after is normally distributed*

- *Equal variances*

- Eg weight before intervention and three months after intervention among 218 women

	weight1	weight2
1	72.00	42.00
2	44.00	30.00
3	40.00	28.00
4	58.00	42.00
5	46.00	14.00
6	40.00	24.00
7	64.00	30.00
8	60.00	58.00
9	46.00	46.00
10	42.00	40.00
11	58.00	46.00
12	50.00	38.00
13	48.00	44.00
14	58.00	46.00
15	48.00	40.00
16	52.00	36.00
17	60.00	56.00
18	60.00	52.00
19	46.00	36.00
20	50.00	32.00
21	38.00	32.00
22	46.00	34.00
23	64.00	36.00

- Eg weight before intervention and three months after intervention among 218 women

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	weight before	54.789	218	10.9890	.7443
	Weight after	50.71	218	9.606	.651

Paired Samples Test

	Paired Differences							
				95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1 weight before - Weight after	4.0803	6.5706	.4450	3.2032	4.9574	9.169	217	.000

- P value = 0.000 ($p < 0.001$) → significant → reject the null → accept the hypothesis
- There a statistically significant difference in weight before and after the intervention with a mean difference = 4.08 and 95%CI = 3.20 to 4.95

- df=217 = (n-1)

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	weight before - Weight after	4.0803	6.5706	.4450	3.2032	4.9574	9.169	217	.000

- Before conducting the test check the normality of the differences

Test statistic

Doing a paired t test

$$t = \frac{\text{mean difference}}{SE(\text{mean difference})} = \frac{\bar{d}}{\sqrt{\frac{SD^2}{n}}}$$

where if $x_{i1} - x_{i2} = d_i$ then the mean of the difference d_i is $\bar{d}$, SD^2 is the standard deviation of the differences, n is the sample size.

t follows a t distribution with $n - 1$ degrees of freedom.

95% CI for the mean difference

$$\text{mean difference} \pm t_{n-1} SE(\text{mean difference})$$

$$= \bar{d} - t_{n-1}\sqrt{\frac{SD^2}{n}} \text{ to } \bar{d} + t_{n-1}\sqrt{\frac{SD^2}{n}}$$

where t_{n-1} is the 2-tailed 5% point of the t distribution with $n - 1$ degrees of freedom which is obtained from tables or a statistical program.

1 sample T test

- to compare mean of a group with a reference mean (may be the population mean)

- ***You have data for one group only***

- ***Assumption : Data are normally distributed.***

- It answers questions like :
- Is the hemoglobin level of children in a refugee camp is different from that of the children general population (the reference normal value)

- Is weight of babies in a conflict area is different from the reference weight of the population (babies)

- Is diastolic BP of health care workers is different from t that of the general population

- We have one sample → obtain mean

- We have the reference value (mean) of the population (from text books, previous research, governmental data, etc)

- Compare the two means

- Eg compare if the weight of a 66 female students is different from the mean of the population which is 50 kg?

- Data was collected from the 280 female students
- Hypothesis: there is difference ($\mu \neq 50$)
- Null: they are equal ($\mu = 50$)

	weight
1	72.00
2	62.00
3	64.00
4	60.00
5	38.00
6	68.00
7	40.00
8	36.00
9	52.00
10	54.00
11	52.00
12	44.00
13	38.00
14	44.00
15	42.00
16	52.00
17	60.00
18	60.00
19	46.00

- Eg compare if the weight of a 66 female students is different from the mean of the population which is 50 kg

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
weight	66	51.0303	8.98712	1.10624

One-Sample Test

	Test Value = 50					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
weight	.931	65	.355	1.03030	-1.1790	3.2396

- Mean population = 50
- Mean sample= 51.03

- P value =0.335→ no significant difference → NO difference

- Mean difference= 1.03 (95%CI -1.179 , 3.239).

- 95% CI of the difference include ZERO → not SIG → NO difference

- Accept the null →Mean of the group = population mean

- Test statistic

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

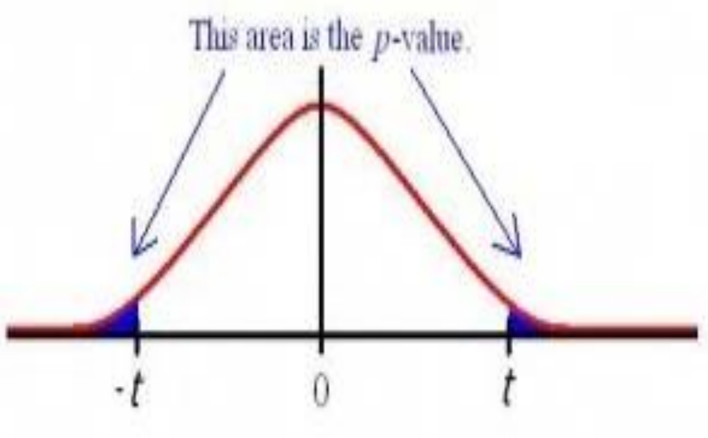
P values

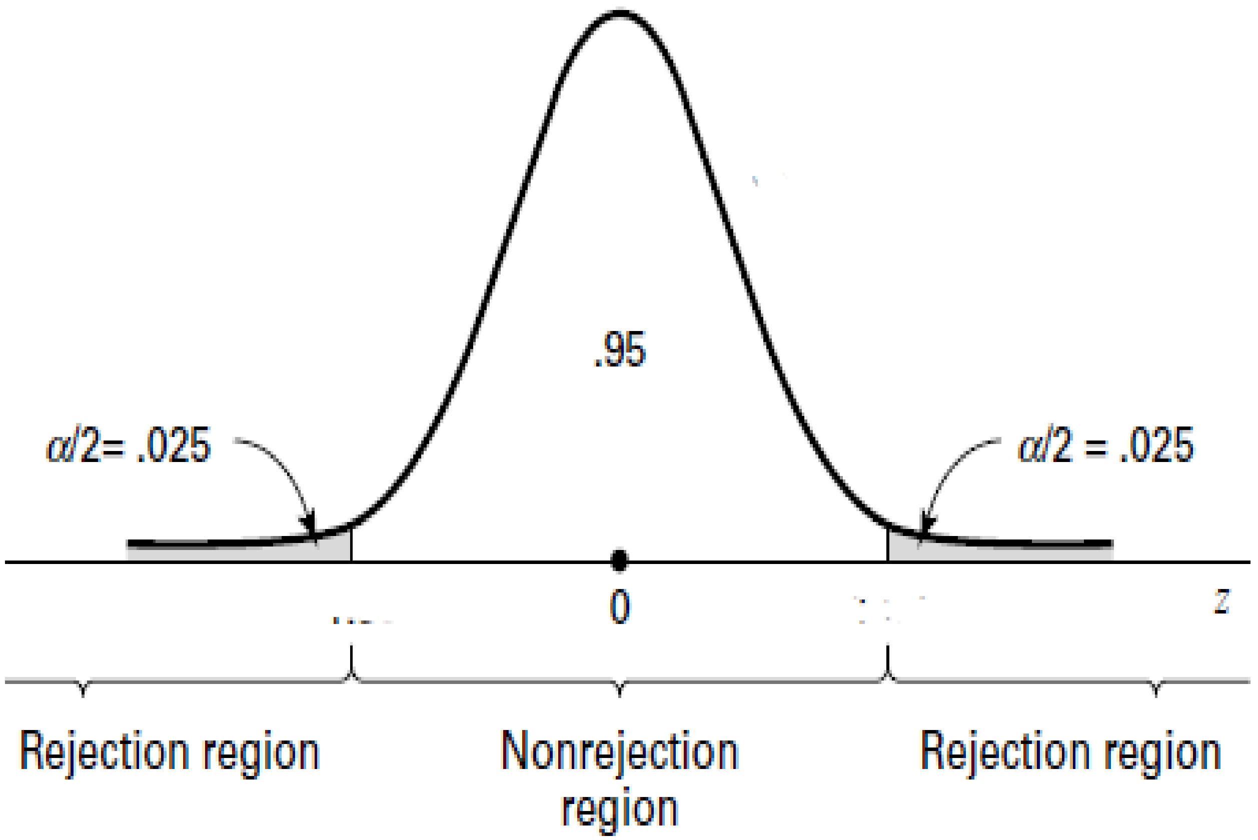
- What is a P value?
- A P value is a probability, and therefore lies between 0 and 1
- It comes from a statistical test that is testing a particular null hypothesis
- It expresses the weight of evidence in favor of or against the stated null hypothesis
- Precise definition: P value is the probability, given that the null hypothesis is true, 0.05 or 5% is commonly used as a cut- off, such that if the observed P is less than this ($P < 0.05$) we consider that there is good evidence that the null hypothesis is not true. This is directly related to the type 1 error
- $P < 0.05$ is described as statistically significant and $P \geq 0.05$ is described as not statistically significant

P values

- Large samples are more likely to show a significant difference (small p value)
- it is possible for data to show a statistically significant result when the size of the effect is **too small** to be clinically important
- **Statistical significance does not mean clinical significance**

Test statistic and rejection region and non rejection region





Test statistic and rejection region and non rejection region

- The figure below shows a t-distribution with **30** degrees of freedom and alpha = 0.05

- Non rejection region
- Rejection region

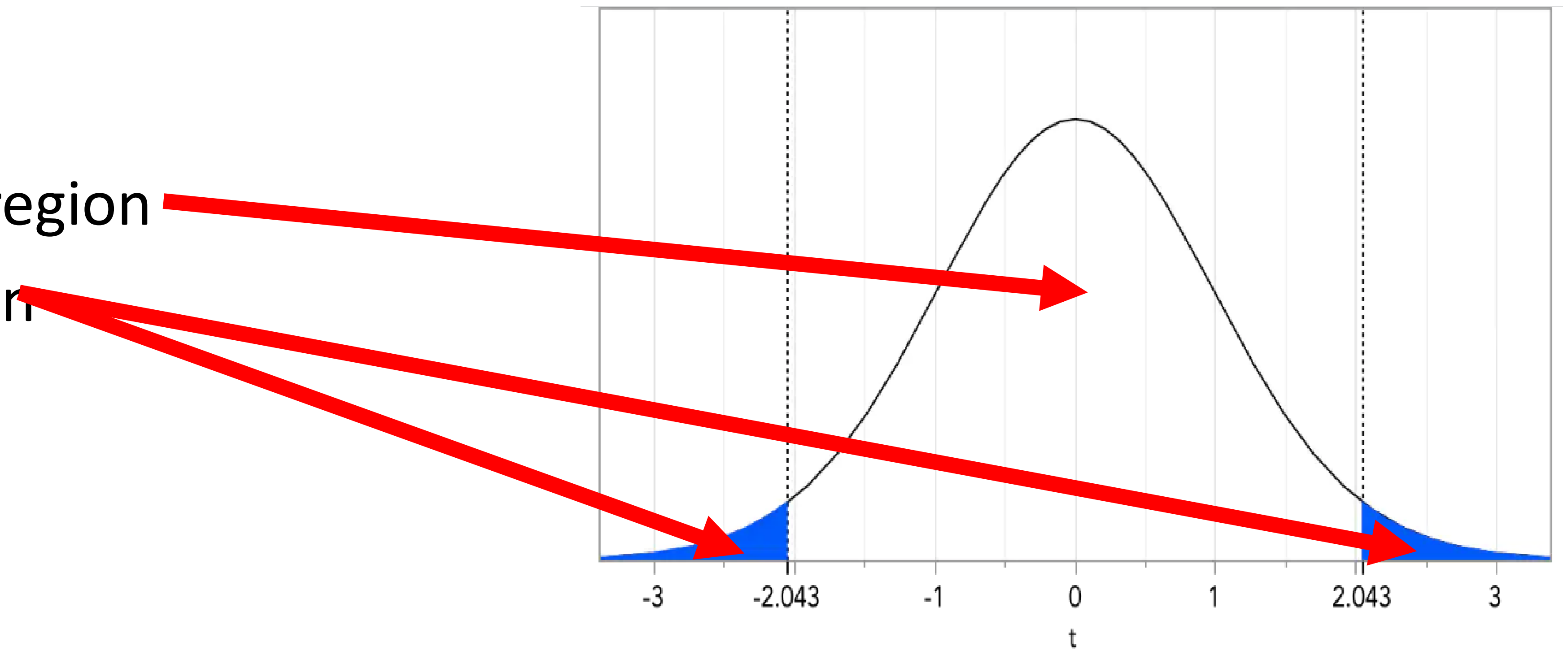


Figure 6: t-distribution with 30 degrees of freedom and α = 0.05

- For example the computed t test was 2.443 and the df =20

- the critical value from the table =2.086

If the computed > critical → Reject the null (p is sig)

In this case our computed t is > Than the critical → p <0.05

two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707

- For example the computed
- t statistic was 1.433 and the
- df =16

- the critical value
- from the table =2.120
- If the computed > critical → Reject the null (p is sig)

- In this case our computed t is< than the critical →p >0.05 (not sig.)

two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	[illegible]
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	[illegible]	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	[illegible]	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	[illegible]	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707