

The t-distribution

t test-origin:

- Founder WS Gosset
- Wrote under the pseudonym “Student”
- Mostly worked in tea (t) time
- ? Hence known as Student's *t* test.
- Certainly if $n \leq 30$

Types:

- **One sample**
 compare with population
- **Unpaired**
 compare with control
- **Paired**
 same subjects: pre-post

T-test:**1. Test for single mean**

Whether the sample mean is equal to the predefined population mean ?

2. Test for difference in means

Whether the CD4 level of patients taking treatment A is equal to CD4 level of patients taking treatment B ?

3. Test for paired observation

Whether the treatment conferred any significant benefit ?

Test direction

- One tailed t test
- Two tailed test

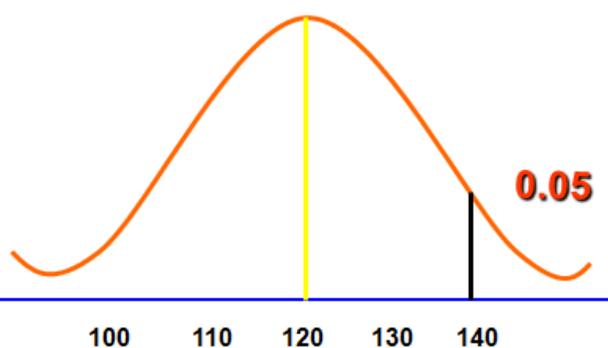
Developing the Pooled-Variance t Test:

- Setting Up the Hypothesis:

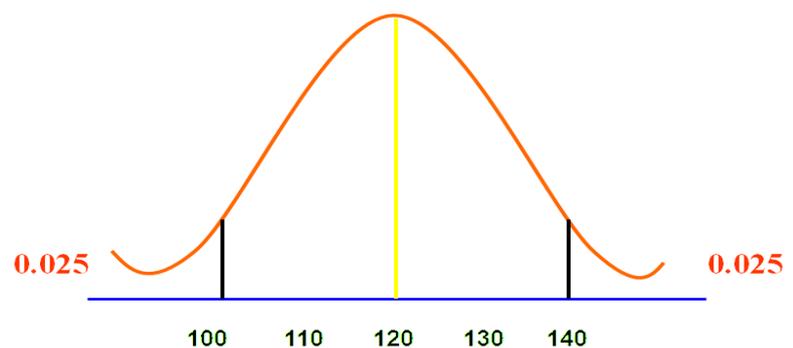
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	OR	$H_0: \mu_1 - \mu_2 = 0$ $H_1: \mu_1 - \mu_2 \neq 0$	Two Tail
$H_0: \mu_1 \leq \mu_2$ $H_1: \mu_1 > \mu_2$	OR	$H_0: \mu_1 - \mu_2 \leq 0$ $H_1: \mu_1 - \mu_2 > 0$	Right Tail
$H_0: \mu_1 \geq \mu_2$ $H_1: \mu_1 < \mu_2$	OR	$H_0: \mu_1 - \mu_2 \geq 0$ $H_1: \mu_1 - \mu_2 < 0$	Left Tail

Example:

Mean systolic BP in nephritis is significantly higher than of normal person

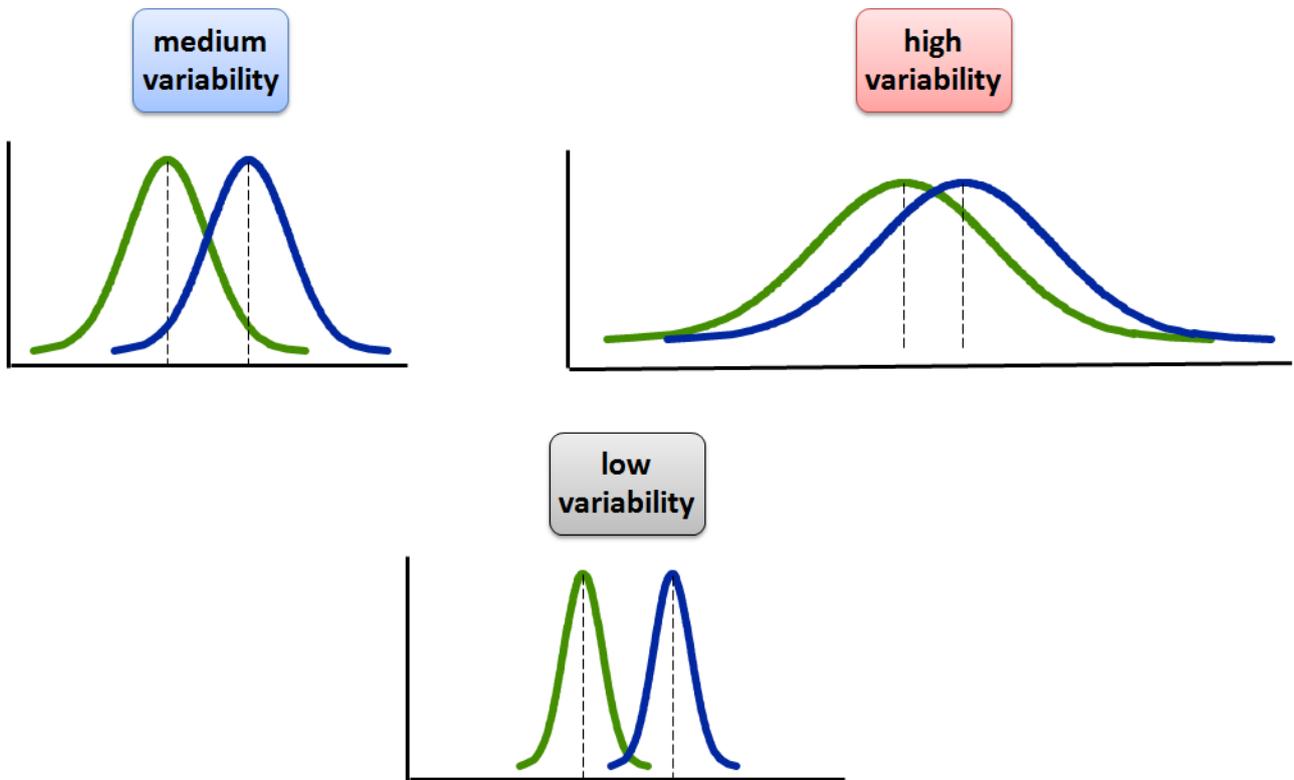


Mean systolic BP in nephritis is significantly different from that of normal person



Statistical Analysis

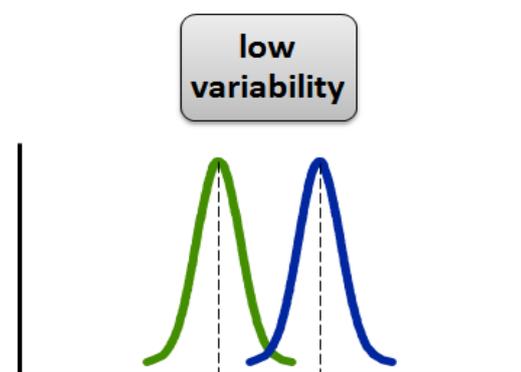
What does difference mean??



The mean difference is the *same* for all three cases

Which one shows the *greatest* difference?

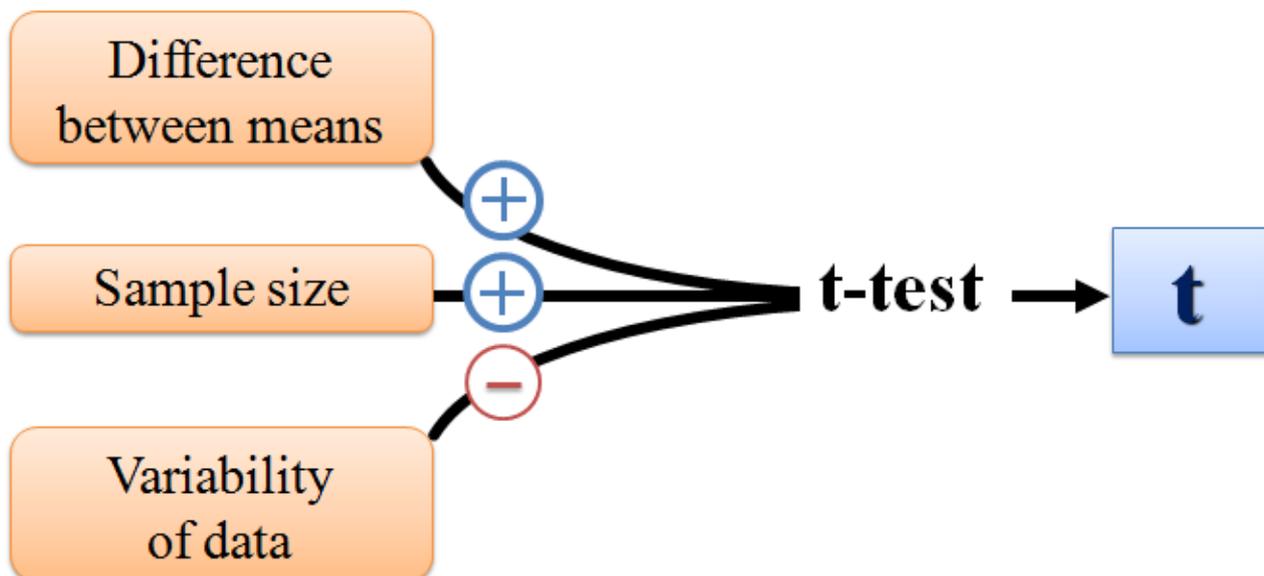
- a statistical difference is a function of the *difference between means* relative to the *variability*
- a small difference between means with large variability could be due to *chance*
- like a *signal-to-noise* ratio



So we estimate

$$\begin{aligned} \frac{\text{signal}}{\text{noise}} &= \frac{\text{difference between group means}}{\text{variability of groups}} \\ &= \frac{\bar{X}_T - \bar{X}_C}{\text{SE}(\bar{X}_T - \bar{X}_C)} \\ &= \text{t-value} \end{aligned}$$

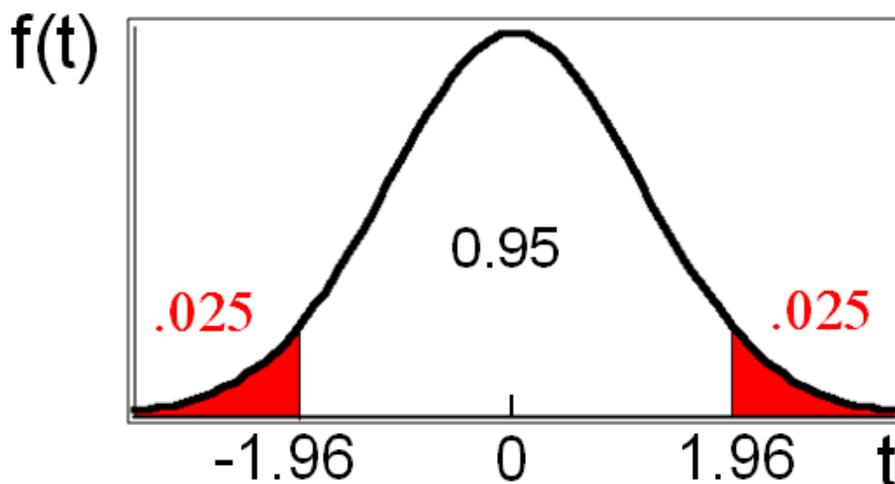
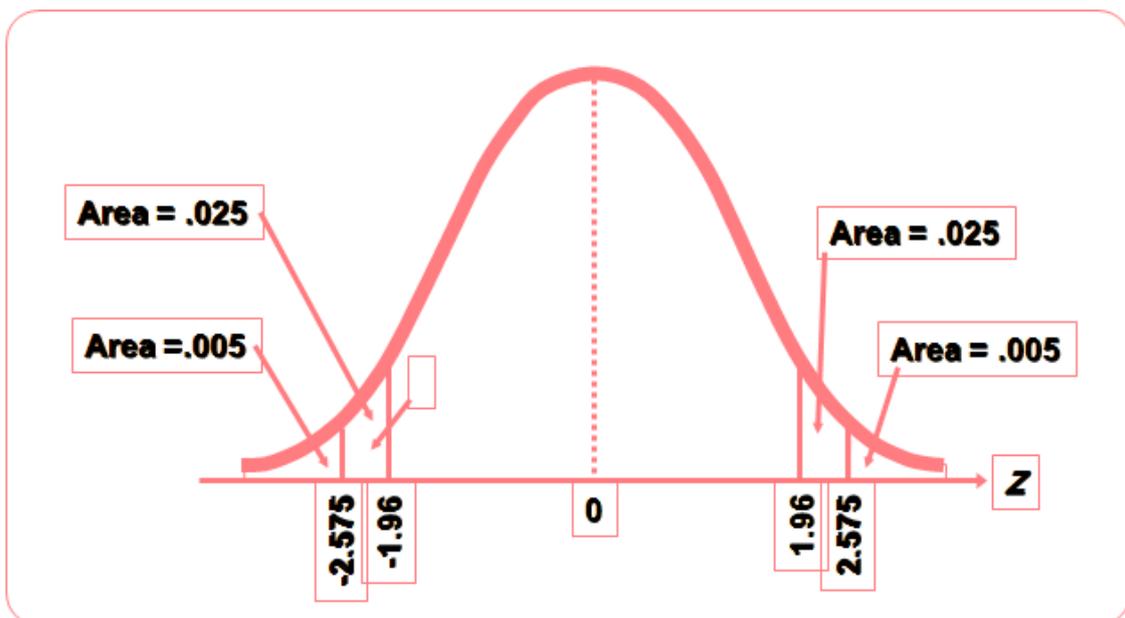
Two sample t - test



Probability – p

- With t we check the probability
- Reject or do not reject Null hypothesis
- You reject if $p < 0.05$ or still less
- Difference between means (groups) is more & more significant if p is less & less

Determining the p-Value:



red area = rejection region for 2-sided test

Assumptions: when you apply t test that you have

- Normal distribution
- Equal variance
- Random sampling

t-Statistic

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

When the sampled population is normally distributed, the t statistic is Student t distributed with n-1 degrees of freedom.

T- test for single mean:

Example.

The following are the weight (mg) of each of 20 rats drawn at random from a large stock. Is it likely that the mean weight of these 20 rats are similar to the mean weight (24 mg) of the whole stock ?

9	18	21	26	14	18	22	27
15	19	22	29	15	19	24	30
		16	20	24	32		

Steps for test for single mean:**1. Questioned to be answered**

Is the Mean weight of the sample of 20 rats is 24 mg?

$N=20$, $\bar{x} = 21.0$ mg, S.D = 5.91, $\mu = 24.0$ mg

2. Null Hypothesis

The mean weight of rats is 24 mg. That is, The sample mean is equal to population mean.

3. Test statistics

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \quad \text{--- } t_{(n-1)} \text{ df}$$

4. Comparison with theoretical value

if tab $t_{(n-1)} < \text{cal } t_{(n-1)}$ reject H_0 ,

if tab $t_{(n-1)} > \text{cal } t_{(n-1)}$ accept H_0 ,

5. Inference

So.

Test statistics : $N=20$, $\bar{x} = 21.0$ mg, S.D = 5.91, $\mu = 24.0$ mg

$$t = \frac{|21.0 - 24|}{5.91 / \sqrt{20}} = -2.30$$

$$t_{\alpha} = t_{.05, 19} = 2.093$$

Accept H_0 if $t < 2.093$

Reject H_0 if $t \geq 2.093$

Inference :

There is no evidence that the sample is taken from the population with mean weight of 24 gm.

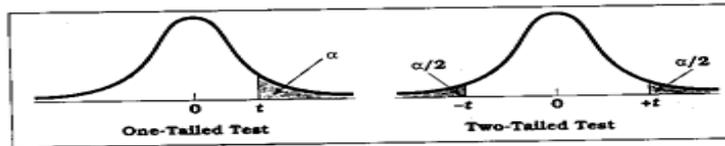


Table D.6 Percentage Points of the t Distribution (Source: The entries in this table were computed by the author.)

df	Level of Significance for One-Tailed Test								
	.25	.20	.15	.10	.05	.025	.01	.005	.0005
	Level of Significance for Two-Tailed Test								
	.50	.40	.30	.20	.10	.05	.02	.01	.001
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	63.662
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	3.496
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.390
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

T-test for difference in means:

Example.

Given below are the 24 hrs total energy expenditure (MJ/day) in groups of lean and obese women. Examine whether the obese women’s mean energy expenditure is significantly higher ?

	<u>Lean</u>		
6.1	7.0	7.5	
7.5	5.5	7.6	
7.9	8.1	8.1	
8.1	8.4	10.2	
	10.9		

	<u>Obese</u>		
8.8	9.2	9.2	
9.7	9.7	10.0	
11.5	11.8	12.8	

Null Hypothesis:

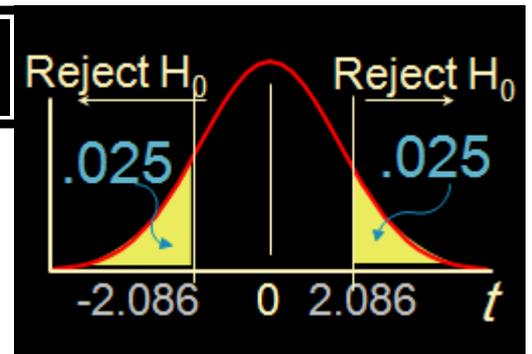
Obese women's mean energy expenditure is equal to the lean women's energy expenditure.

Data summary		
	Lean	Obeses
N	13	9
	8.10	10.30
S	1.38	1.25

Solution:

- $H_0: \mu_1 - \mu_2 = 0$ ($\mu_1 = \mu_2$)
- $H_1: \mu_1 - \mu_2 \neq 0$ ($\mu_1 \neq \mu_2$)
- $\alpha = 0.05$
- $df = 13 + 9 - 2 = 20$

Critical Value(s):



Compute the Test Statistic:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Hypothesized Difference (usually zero when testing for equal means)

$$df = n_1 + n_2 - 2$$

Calculate the Pooled Sample Variances as an Estimate of the Common Populations Variance:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

S_p^2 = Pooled-Variance

n_1 = Size of Sample 1

S_1^2 = Variance of Sample 1

n_2 = Size of Sample 2

S_2^2 = Variance of sample 2

First, estimate the common variance as a weighted average of the two sample variances using the degrees of freedom as weights

$$S_p^2 = \frac{(13-1) \cdot 1.38^2 + (9-1) \cdot 1.25^2}{(13-1) + (9-1)} = 1.765$$

- Calculating the test statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(8.1 - 10.3) - 0}{\sqrt{1.76 \cdot \left(\frac{1}{13} + \frac{1}{9}\right)}} = 3.82$$

tab t 9+13-2 =20 dff = t 0.05,20 =2.086

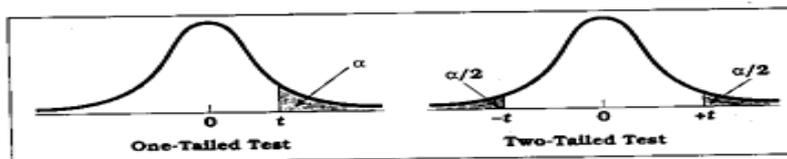


Table D.6 Percentage Points of the t Distribution (Source: The entries in this table were computed by the author.)

df	Level of Significance for One-Tailed Test								
	.25	.20	.15	.10	.05	.025	.01	.005	.0005
	Level of Significance for Two-Tailed Test								
	.50	.40	.30	.20	.10	.05	.02	.01	.001
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	63.662
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	3.496
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.390
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

T-test for difference in means:

Inference : The cal t (3.82) is higher than tab t at 0.05, 20. i.e. 2.086 . This implies that there is a evidence that the mean energy expenditure in obese group is significantly ($p < 0.05$) higher than that of lean group.

Example.

Suppose we want to test the effectiveness of a program designed to increase scores on the quantitative section of the Graduate Record Exam (GRE). We test the program on a group of 8 students. Prior to entering the program, each student takes a practice quantitative GRE; after completing the program, each student takes another practice exam. Based on their performance, was the program effective?

- **Each subject contributes 2 scores: repeated measures design**

Student	Before Program	After Program
1	520	555
2	490	510
3	600	585
4	620	645
5	580	630
6	560	550
7	610	645
8	480	520

- **Can represent each student with a single score: the difference (D) between the scores:**

Student	Before Program	After Program	D
1	520	555	35
2	490	510	20
3	600	585	-15
4	620	645	25
5	580	630	50
6	560	550	-10
7	610	645	35
8	480	520	40

- **Approach:**
test the effectiveness of program by testing significance of D
- **Null hypothesis:**
There is no difference in the scores of before and after program
- **Alternative hypothesis:**
program is effective → scores after program will be higher than scores before program → average D will be greater than zero

$$H_0: \mu_D = 0$$

$$H_1: \mu_D > 0$$

Student	Before Program	After Program	D	D ²
1	520	555	35	1225
2	490	510	20	400
3	600	585	-15	225
4	620	645	25	625
5	580	630	50	2500
6	560	550	-10	100
7	610	645	35	1225
8	480	520	40	1600
			$\sum D = 180$	$\sum D^2 = 7900$

Recall that

For single samples:

$$t_{obt} = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{\text{score} - \text{mean}}{\text{standard error}}$$

For related samples:

$$t_{obt} = \frac{\bar{D} - \mu_D}{s_{\bar{D}}}$$

where:

$$s_{\bar{D}} = \frac{s_D}{\sqrt{N}} \quad \text{and} \quad s_D = \sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{N-1}}$$

Mean of D:
$$\bar{D} = \frac{\sum D}{N} = \frac{180}{8} = 22.5$$

Standard deviation of D:
$$s_D = \sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{N-1}} = \sqrt{\frac{7900 - \frac{(180)^2}{8}}{8-1}} = 23.45$$

Standard error:
$$s_{\bar{D}} = \frac{s_D}{\sqrt{N}} = \frac{23.45}{\sqrt{8}} = 8.2908$$

$$t_{obt} = \frac{\bar{D} - \mu_D}{s_{\bar{D}}}$$

Under H_0 , $\mu_D = 0$, so:
$$t_{obt} = \frac{\bar{D}}{s_{\bar{D}}} = \frac{22.5}{8.2908} = 2.714$$

From Table (next page): for $\alpha = 0.05$, one-tailed, with $df = 7$

$$t_{critical} = 1.895$$

$2.714 > 1.895 \rightarrow$ reject H_0

The program is effective

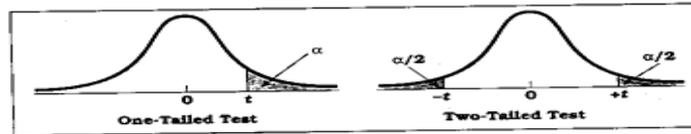


Table D.6 Percentage Points of the t Distribution (Source: The entries in this table were computed by the author.)

df	Level of Significance for One-Tailed Test								
	.25	.20	.15	.10	.05	.025	.01	.005	.0005
	Level of Significance for Two-Tailed Test								
	.50	.40	.30	.20	.10	.05	.02	.01	.001
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	63.662
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	3.496
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.390
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

t-Value:

t is a measure of: How difficult is it to believe the null hypothesis?

High t

Difficult to believe the null hypothesis - accept that there is a real difference.

Low t

Easy to believe the null hypothesis - have not proved any difference.

In Conclusion

Student 's t-test will be used:

- **When Sample size is small**
- **and for the following situations:**
 - to compare the single sample mean with the population mean
 - to compare the sample means of two independent samples
 - to compare the sample means of paired samples