T-test for difference of Mean

Suppose we want to test if two independent samples x_i ($i = 1, 2 \ldots n_1$) and y_j ($j = 1, 2, \ldots n_2$) of size n_1 and n_2 have been drawn from two normal populations with means μ_x and μ_y respectively.

Under the null hypothesis (H_0) that the samples have been drawn from the normal populations with means μ_x and μ_y under the assumption that the population variance are equal, ie

$$\sigma_x^2 = \sigma_y^2 = \sigma^2 \text{ (say)}$$

the statistic

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $S^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum_i (x_i - \bar{x})^2 + \sum_j (y_j - \bar{y})^2\right]$

is an unbiased estimate of the common population variance σ^2, follows Student's t-dist with $(n_1 + n_2 - 2)$ d.f.

Example Below are given the gain in weight (in kgs) of pigs fed on two diets A and B.

Gain in weight

Diet A	25 32 30 34 24 14 32 24 30 31 35 25
Diet B	44 34 22 10 47 31 40 30 32 35 18 21 35 29 22

Test, if two diets differ significantly as regards to their effect on increase in weight.

Null hypothesis $H_0: \mu_x = \mu_y$, ie there is no significant difference between the mean increase in weight due to ~~two~~ diets

Alternate hypo: $H_1: \mu_x \neq \mu_y$

Sol

Diet A | Diet B

x	$x-\bar{x}$	$(x-\bar{x})^2$	y	$y-\bar{y}$	$(y-\bar{y})^2$
25	-3	9	44	14	196
32	4	16	34	4	16
30	2	4	22	-8	64
34	6	36	10	-20	400
24	-4	16	47	17	289
14	-14	196	31	1	1
32	4	16	40	10	100
24	-4	16	30	0	0
30	2	4	32	2	4
31	3	9	35	5	25
35	7	49	18	-12	144
25	-3	9	21	-9	81
$\Sigma x = 336$	$\Sigma(x-\bar{x}) = 0$	$\Sigma(x-\bar{x})^2 = 380$	35	5	25
			29	-1	1
			22	-8	64
			$\Sigma y = 450$	$\Sigma(y-\bar{y}) = 0$	$\Sigma(y-\bar{y})^2 = 1410$

$$\bar{x} = \frac{336}{12} = 28, \quad \bar{y} = \frac{450}{15} = 30$$

$$S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2\right]$$

$$= \frac{1}{12+15-2}\left[380 + 1410\right] = 71.6$$

Under the null hypothesis

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{28-30}{\sqrt{71.6\left(\frac{1}{12}+\frac{1}{15}\right)}}$$

$$= \frac{-2}{\sqrt{10.74}} = -0.609$$

Tabulated $t_{0.05}$ for $(12+15-2) = 25$
deg of freedom $= 2.06$

Conclusion: $|t| = 0.609 <$ tabulated t
Ho may be accepted at 5% level of significance and we conclude that the two diets do not differ significantly as regards to their increase in weight.

Observation: What have you noted?
In single mean t-test, one sample was given and based on that sample, we were testing whether the sample can be considered to have come from a normal population with a hypothetical mean.
But in t-test for difference of mean, we have two samples of different sizes and we are testing whether they have been drawn from two different normal populations.

Ex ② Samples of two types of electric bulbs were tested for length of life and following data were obtained.

	Type I	Type II
Sample No	$n_1 = 8$	$n_2 = 7$
Sample means	$\bar{x}_1 = 1234$ hrs	$\bar{x}_2 = 1036$ hrs
Sample S.D.'s	$s_1 = 36$ hrs	$s_2 = 40$ hrs

Is the difference in means sufficient to warrant that type I is superior to type II regarding the length of life?

[Note - Here again we have two samples and we want to test whether they come from same normal population or type I is superior (This is right tailed)]

Scanned with CamScanner

Null hypothesis $H_0: \mu_x = \mu_y$, the two types I and II of electric bulbs are identical.

Alternative hypothesis: $H_1: \mu_x > \mu_y$, ie type I is superior to type II

Test statistics

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x_1-\bar{x}_1)^2 + \Sigma(y_1-\bar{y}_1)^2\right]$$

$$= \frac{1}{n_1+n_2-2}\left[n_1 s_1^2 + n_2 s_2^2\right]$$

$$= \frac{1}{13}\left[8 \times 36^2 + 7 \times 40^2\right]$$

$$= 1659.08$$

$$t = \frac{1234 - 1036}{\sqrt{1659.08\left(\frac{1}{8}+\frac{1}{7}\right)}} = \frac{198}{\sqrt{1659.08 \times 0.2679}}$$

$$= 9.39$$

Tabulated t for 13 deg of freedom at 5% level for right (single) tailed test is 1.77

$$9.39 > 1.77$$

Conclusion: Null hypothesis is rejected as calculated t is much greater than tabulated t value.

Hence type I is definitely superior.