Paired t-test For difference of Mean

Consider the case when (i) the sample sizes are equal i.e $n_1 = n_2 = n$ (say), and (ii) the two samples are not independent but sample observations are paired together, i.e, the pair of observations (x_i, y_i) $(i=1,2,\cdots n)$ corresponds to the same (ith) sample unit. The problem is to test if the sample means differ significantly or not.

For example, suppose we want to test the efficacy of a particular drug, say, for inducing sleep. Let x_i and y_i $(i=1,2,\cdots n)$ be the readings, in hours of sleep, on the ith individual before and after the drug is given respectively. Here instead of applying the difference of mean test discussed in previous class, we apply paired t-test given below

Here we consider the increments $d_i = x_i - y_i$ $(i=1,2 \cdots n)$

Under the null hypothesis, H_0 that increments are due to fluctuations of sampling, i.e., the drug is not responsible for these increments

the statistic : $$d = \frac{\bar{d}}{\frac{S}{\sqrt{n}}}$$

where $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$

follows student's t-distribution with $(n-1)$ degree of freedom.

Remark: Some students have confusion with difference of mean t-test and paired t-test for difference of mean. You need to be careful to understand it:

Ex ① A certain stimulus administered to each of the 12 patients resulted in the following increase of Blood pressure:
5, 2, 8, −1, 3, 0, −2, 1, 5, 0, 4 and 6
Can we conclude that the stimulus, in general, be accompanied by an increase in blood pressure?

(**Remark:** Before solving, it let us understand that here the increase in blood pressure corresponds to the same individual, i.e there are 12 patients and the difference in blood pressure is given for each of them)

Sol: Here we are given the increments in blood pressure i.e $d_i (= x_i - y_i)$

Null hypothesis $H_0: \mu_x = \mu_y$ i.e., there is no significant difference in the blood pressure readings of the patients before and after the drug. In other words the given increments are just by chance and not due to stimulus.

Alternative hypothesis $H_1: \mu_x < \mu_y$, i.e the stimulus results in an increase in blood pressure.

Test statistic $t = \dfrac{\bar{d}}{S/\sqrt{n}}$

Page

$\bar{d} = \frac{1}{n}\Sigma d$, and $S^2 = \frac{1}{n-1}\left[\Sigma d^2 - \frac{(\Sigma d)^2}{n}\right]$

d	5	2	8	-1	3	0	-2	1	5	0	4	6
d^2	25	4	64	1	9	0	4	1	25	0	16	36

$\Sigma d = 31$ and $\Sigma d^2 = 185$

$\bar{d} = \frac{31}{12} = 2.58$

$S^2 = \frac{1}{12-1}\left[\Sigma d^2 - \frac{(\Sigma d)^2}{n}\right]$

$= \frac{1}{11}\left[185 - \frac{(31)^2}{12}\right] = 9.5382$

$t = \frac{\bar{d}}{S/\sqrt{n}} = \frac{2.58}{\frac{\sqrt{9.5382}}{\sqrt{12}}} = \frac{2.58 \times \sqrt{12}}{\sqrt{9.5382}}$

$= 2.89$

Tabulated $t_{0.05}$ for 11 d.f $= 1.80$

Calulated t > Tabulated t

Null hypothesis is rejected

Hence we conclude that, the stimulus, will in general be accompanied by an increase in blood pressure.

Scanned with CamScanner

Page

Ex (2) In a certain experiment to compare two types of animal foods A and B the following result of increase in weights were observed in animals:

Animal number		1	2	3	4	5	6	7	8	Total
Increase weight in pound	Food A	49	53	51	52	47	50	52	53	407
	Food B	52	55	52	53	50	54	54	53	423

(i) Assuming that the two samples of animals are ~~independent~~ independent, can we conclude that food B is better than food A?

(ii) Also examine the case when the same set of eight animals were used in both the fields.

Sol: Null hypothesis, H_0: If the increase in weights due to food A and B are denoted by X and Y respectively, then $H_0: \mu_X = \mu_Y$, i.e. there is no significant difference in increase in weight due to diets A and B.

Alternative hypothesis $H_1: \mu_X < \mu_Y$

(i) If two samples are assumed to be independent we will apply t-test for difference of means to test H_0

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

Food A			Food B		
X	d=X−50	d^2	Y	D=Y−52	D^2
49	−1	1	52	0	0
53	3	9	55	3	9
51	1	1	52	0	0
52	2	4	53	1	1
47	−3	9	50	−2	4
50	0	0	54	2	4
52	2	4	54	2	4
53	3	9	53	1	1
Total	7	37	Total	7	23

$\bar{x} = 50 + \frac{7}{8}$
$= 50.875$

$\bar{y} = 52 + \frac{7}{8}$
$= 52.875$

and $\Sigma(x-\bar{x})^2 = \Sigma d^2 - \frac{(\Sigma d^2)}{n_1} = 37 - \frac{49}{8}$
$= 30.875$

and $\Sigma(y-\bar{y})^2 = \Sigma D^2 - \frac{(\Sigma D)^2}{n_2}$
$= 23 - \frac{49}{8} = 16.875$

$S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2\right]$
$= \frac{1}{14}[30.875 + 16.875] = 3.41$

Tabulated $t_{0.05}$ for $(8+8-2) = 14$ d.f $= 1.76$

$t = \frac{\bar{x}-\bar{y}}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{50.875 - 52.875}{\sqrt{3.41\left(\frac{1}{8}+\frac{1}{8}\right)}} = -2.17$

The critical region for left tail test is $t < -1.76$
Since calculated t is less than -1.76, H_0 is rejected at 5% level of significance.
So we conclude that food B is superior.

Date
Page

(11) If the same set of animals is used in both the cases, then reading X and Y are not independent but they are paired together and we apply the paired t-test for H_0

$H_0: \mu_X = \mu_Y$

$$t = \frac{\bar{d}}{s/\sqrt{n}}$$

									Total
X	49	53	51	52	47	50	52	53	
Y	52	55	52	53	50	54	54	53	
$d = X - Y$	−3	−2	−1	−1	−3	−4	−2	0	−16
d^2	9	4	1	1	9	16	4	0	44

$$\bar{d} = \frac{\Sigma d}{n} = \frac{-16}{8} = -2$$

$$S^2 = \frac{1}{n-1}\left[\Sigma d^2 - \frac{(\Sigma d)^2}{n}\right]$$

$$= \frac{1}{7}\left(44 - \frac{256}{8}\right) = 1.714$$

$$t = \frac{-2}{\frac{\sqrt{1.714}}{\sqrt{8}}} = \frac{-2 \times \sqrt{8}}{\sqrt{1.714}} = -4.32$$

Tabulated $t_{0.05}$ for $(8-1) = 7$ d.f for one tail test is 1.90

Calculated $|t| > 1.90$

Null hypothesis is rejected

So we conclude that food B is superior.

Scanned with CamScanner