t-tests

(Student t-tests)

Learning outcome

- To interpret t-test result

- Types of t-test :
  - One sample t-test
  - 2 independent samples t-test / 2 samples t-test
  - Paired-sample t-test

- Assumption (general)

1. Scale of measurement –
   - the data should be at interval or ratio level of measurement
2. Random sampling
   - The scores should be randomly sampled from the population of interest
3. Normality
   - The scores should be normally distributed in the population

- Steps in solving Biometry :

1. Hypothesis
2. Test statistics
3. Value of α
4. Critical value, t_t / F_t / χ_t, …..
5. Calculation
6. Conclusion

- $H_0 : \mu\ (\geq \text{ or } = \text{ or } \leq)\ A$
- $H_1 : \mu\ (< \text{ or } \neq \text{ or } >)\ A$; A= a number

- $H_0 : \mu_1\ (\geq \text{ or } = \text{ or } \leq)\ \mu_2$
- $H_1 : \mu_1\ (< \text{ or } \neq \text{ or } >)\ \mu_2$

$$t = \frac{\overline{X} - \mu}{s_{\overline{X}}} \qquad t = \frac{\overline{y_1} - \overline{y_2}}{s_{\overline{y_1} - \overline{y_2}}}$$

- t = t value (from calculation)
- $\overline{X}$ = Mean samples
- *μ = mean population*
- $s_{\overline{X}}$ = *standard error of samples.*

Standard deviation is known – z-test

Standard deviation is unknown – t-test

One sample test :

- Data from a single of participants
- Wish to know whether the mean of the population from which the sample is drawn is the same as the hypothesis mean (given μ)

$$t = \frac{\overline{X} - \mu}{\left(s / \sqrt{n} \right)} \qquad s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

$$\overline{X} = \frac{\Sigma X_i}{n}$$

$$s^2 = \frac{\Sigma X_i^2 - \frac{(\Sigma X_i)^2}{n}}{n-1} \quad \text{or} \quad = \frac{\Sigma (X_i - \overline{X})^2}{n-1}$$

$$s = \sqrt{s^2}$$

Example 1

- A student exposed 25 crabs to air at 24.3^0C. The student would like to determine whether the **mean** body temperature of the crabs is the same as the ambient air temperature of 24.3^0C.
- The data of body temperature of crabs are display as below :

25.8	26.2	24.3	25.5	23.9
24.6	23.9	24.6	25.4	27.0
26.1	24.5	23.3	26.3	24.8
22.9	24.0	25.5	23.5	22.9
25.1	27.3	28.1	24.8	25.4

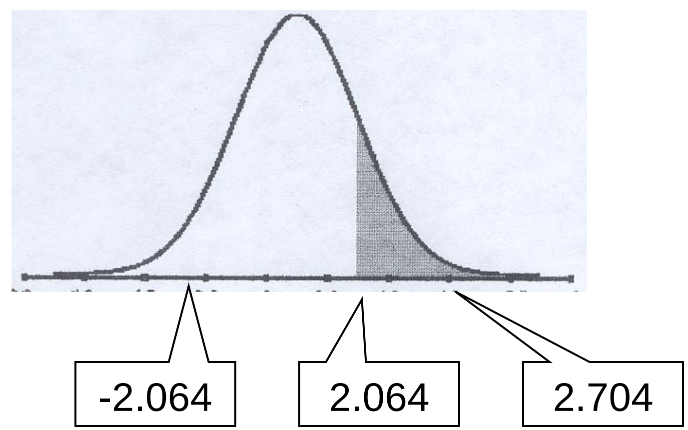
Solution 1

1. $H_0 : \mu = 24.3^0C$

   $H_a : \mu \neq 24.3^0C$
2. t-test (to compare the samples mean with 24.3^0C)
3. Value of α = 0.05
4. Critical value = t_t = $t_{\alpha(2),\nu} = t_{0.05(2),(25-1=24)} = \pm 2.064$
5. Calculation,

$$\overline{X} = 25.03 \qquad s_{\overline{X}} = 0.27$$

$$t = \frac{\overline{X} - \mu}{s_{\overline{X}}} = \frac{25.03 - 24.3}{0.27} = 2.704$$



6. $|t| \rangle t_t$, reject H_0

   The sample of crabs body temperature is not 24.3^0C

SPSS output

One-Sample Test

	Test Value = 24.3					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Crab temperature	2.713	24	.012	.72800	.1741	1.2819

Reject H_0 when $p < \alpha$

Reject H_0 when $p < 0.05$

Example 2

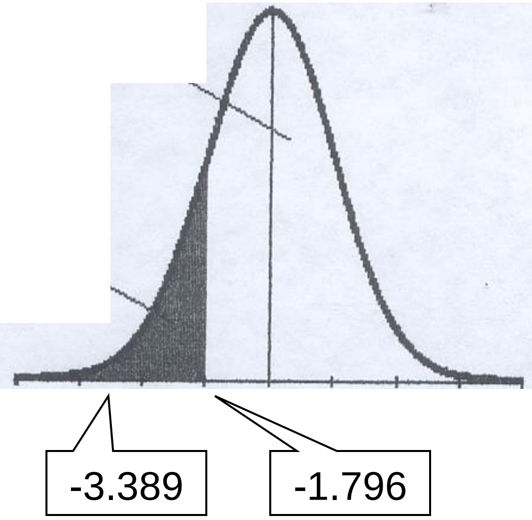
- The data are weight changes of human, tabulated after administration of a drug proposed to result in weight loss. Each weight change (in kg) is the weight after minus the weight before drug administration.

0.2	0.4	-0.1	-0.8
-0.5	-0.7	0.0	-1.2
-1.3	-1.6	-0.6	-1.1

Solution 2

1. $H_0 : \mu \geq 0$, $H_a : \mu < 0$
2. Test sta. : t test
3. α=0.05, n=12, ν=11, $\overline{X} = -0.61$
4. t_t=-1.796
5. t = -3.389
6. If $t \leq -t_t$, reject H_0. → reject H_0

   The drug proposed do cause weight loss in human.



SPSS output

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
weight loss	-3.329	11	.007	-.60833	-1.0106	-.2061

$p = 0.007/2 = 0.0035$ (1 tailed)

Tips: $p < \alpha$; Reject H_0

$p = 0.0035 < \alpha = 0.05$

$\therefore$ Reject H_0

Two samples test (independent t test):

- Different participants have performed in each of the different conditions (when the participants in one condition are different from the participants in the other condition)
- Referred as a between-subjects design

$$t = \frac{\overline{y_1} - \overline{y_2}}{s_{\overline{y_1} - \overline{y_2}}}$$

$$s_{\overline{y_1} - \overline{y_2}} = \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

$$s_p^2 = \frac{v_1 s_1^2 + v_2 s_2^2}{v_1 + v_2}$$

$$v = n - 1$$

$$s^2 = \frac{\Sigma Y^2 - \left(\frac{(\Sigma Y)^2}{n} \right)}{n - 1}$$

Example 3

- Thirteen persons were divided at random into two groups, one group of six and group of seven
- The members of the first group were given one kind of drug (called "B"), and the members of the second group were given another kind of drug (called "G"). Blood is to be taken from each person and the time it takes the blood to clot is to recorded.
- Is blood from persons treated with drug B has the same mean clotting time as does blood from persons treated with drug G?

Given drug B	Given drug G
8.8	9.9
8.4	9.0
7.9	11.1
8.7	9.6
9.1	8.7
9.6	10.4
	9.5

Solution 3

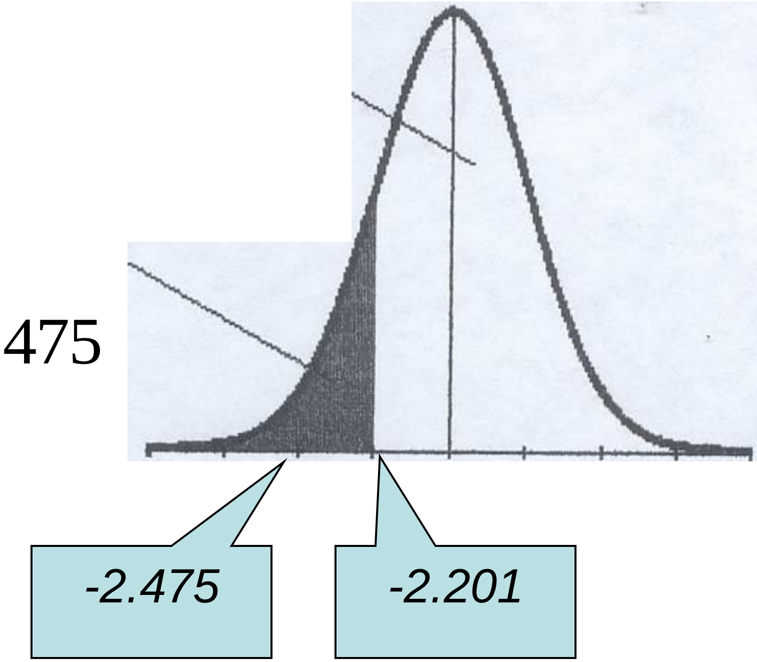
- $H_0 : \mu_B = \mu_G$; $H_0 : \mu_B \neq \mu_G$
- t-test
- $\alpha = 0.05$, $\overline{X_B} = 8.75\text{min}$ $\overline{X_G} = 9.74\text{min}$

$$s_p^2 = \frac{SS_B + SS_G}{\nu_B + \nu_G} = 0.5193\text{min}^2$$

$$s_{\overline{X_B} - \overline{X_G}} = \sqrt{\frac{s_p^2}{n_B} + \frac{s_p^2}{n_G}} = 0.40\text{min}$$

$$t_t = t_{0.05(2),11} = \pm 2.201$$

$$t = \frac{\overline{X}_B - \overline{X}_G}{s_{\overline{X}_B - \overline{X}_G}} = \frac{8.75 - 9.74}{0.4} = -2.475$$



If $t < -t_t$, *reject* H_0

The two drug gave different result at 0.05 sig.

Drug B clotting faster than drug G

SPSS output

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
Effectiveness	Equal variances assumed	.747	.406	-2.476	11	.031	-.99286	.40091	-1.87526	-.11045
	Equal variances not assumed			-2.545	10.701	.028	-.99286	.39006	-1.85430	-.13141

$p = 0.031$
(2-tailed as compared if drugs B &G have the same effectiveness)

Tips: $p < \alpha$; Reject H_0

$p = 0.031 < \alpha = 0.05$
$\therefore$ Reject H_0

Example 4

- The data are heights of plants, each grown with one of two different fertilizers

Present fertilizer	Never fertilizer
48.2	54.8
54.6	59.8
58.3	58.0
47.8	61.3
51.4	53.2
52.0	55.6
49.1	57.4
49.9	52.3
52.6	
55.2	

Solution 4

- $H_0 : \mu_1 \geq \mu_2$

  $H_0 : \mu_1 < \mu_2$
- t-test
- $\alpha = 0.05$

$$\overline{X_1} = 51.91\,\text{cm}; \quad \overline{X_2} = 56.55\,\text{cm}$$

$$s_p^2 = \frac{SS_1 + SS_2}{\nu_1 + \nu_2} = 10.71\,\text{cm}^2$$

$$s_{\overline{X_1} - \overline{X_2}} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = 1.55\,\text{cm}$$

$$t_t = t_{0.05(1),16} = -1.746$$

$$t = \frac{\overline{X_1} - \overline{X_2}}{s_{\overline{X_1} - \overline{X_2}}} = -2.99$$

- If $t < t_t$, reject H_0
- $\therefore$ newer fertilizer is much better than present fertilizer at $\alpha=0.05$.

Paired-sample t-test :

- Also known as repeated measures t-test / dependent-samples
- Only have data from one group of participants
- Referred as within-subject design
- To determine whether the difference between means for the two sets of scores is the same or different

- Additional assumption:
  - Normality of population difference scores
    - The difference between the scores for each participant should be normally distributed
    - Providing the samples size is not too small (30+), violations of this assumption are of little concern.

$$t = \frac{\bar{d}}{s_{\bar{d}}}$$

$$\bar{d} = \frac{\Sigma d}{n}$$

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

$$s_{\bar{d}} = \sqrt{\frac{s_d^2}{n}}$$

Example 5

- A student wish to test the null hypothesis that the left foreleg and left hind leg lengths of deer are equal :

Deer (j)	Hindleg length (cm) (X_{1j})	foreleg length (cm) (X_{2j})	Difference (cm) ($d_j=X_{1j}-X_{2j}$)
1	142	138	4
2	140	136	4
3	144	147	-3
4	144	139	5
5	142	143	-1
6	146	141	5
7	149	143	6
8	150	145	5
9	142	136	6
10	148	146	2

Solution 5

- $H_0 : \mu_d = 0$; $H_a : \mu_d \neq 0$
- Paired t-test
- $\alpha = 0.05$; $n = 10$; $\nu = 9$;

$$\bar{d} = 3.3\text{cm};$$

$$s_d^2 = 9.344\text{cm}^2;$$

$$s_{\bar{d}} = 0.97\text{cm};$$

$$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{3.3}{0.97} = 3.402$$

$$t_{0.05(2),9} = 2.262$$

If $t > t_t$, reject H_0

SPSS output

Paired Samples Test

Pair 1	Hindleg_length - foreleg_length	Paired Differences	Mean		3.30000
			Std. Deviation		3.05687
			Std. Error Mean		.96667
			95% Confidence Interval of the Difference	Lower	1.11325
				Upper	5.48675
		t			3.414
		df			9
		Sig. (2-tailed)			.008

$p = 0.08$

Tips: $p < \alpha$; Reject H_0

$p = 0.008 < \alpha = 0.05$

$\therefore$ Reject H_0

Thank you