SMALL SAMPLE TEST - 2

Prof. Nandini Rai
KJSCE

TESTING THE DIFFERENCE BETWEEN MEANS

- If the sample size $(n_1 + n_2 - 2)$ is small, the test statistic is computed as, $t = \frac{\bar{X}_1 - \bar{X}_2}{S.E.}$
- The statistic t so computed follows Student's t – distribution.
- The standard error of the difference between the two means is then given by
- $S.E. = s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

- If we are given unbiased standard deviations of the two samples, then,

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

- Where $S_1 = \sqrt{\frac{\sum(x_{i1}-\bar{x}_1)^2}{n_1-1}}$ and $S_2 = \sqrt{\frac{\sum(x_{i2}-\bar{x}_2)^2}{n_2-1}}$

- On the other hand, if we are given standard deviations of the two samples then,
- $s_p = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$
- Where $s_1 = \sqrt{\frac{\sum(x_{i1} - \bar{x}_1)^2}{n_1}}$ *and* $s_2 = \sqrt{\frac{\sum(x_{i2} - \bar{x}_2)^2}{n_2}}$

- In both the cases $S_p = \sqrt{\frac{\sum(x_{i1}-\overline{x_1})^2+\sum(x_{i2}-\overline{x_2})^2}{n_1+n_2-2}}$
- And $S.E. = s_p \times \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$

- When population standard deviation σ_1 and σ_2 are known, we can assume $\bar{X}_1 - \bar{X}_2$ to be normal with mean zero and S.E. $= \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$ and hence use Z-distribution.
- It is assumed that the two populations have the same standard deviation σ. If we cannot assume that $\sigma_1 = \sigma_2$, then the problem is not in syllabus.

EXAMPLE-1

If two independent random samples of sizes 15 and 8 have respectively the following means and population standard deviations, $\bar{x}_1 = 980, \bar{x}_2 = 1012, \sigma_1 = 75, \sigma_2 = 80$. Test the hypothesis that $\mu_1 = \mu_2$ at 5% level of significance.

When population standard deviations σ_1 and σ_2 are known, we can assume $\bar{x}_1 - \bar{x}_2$ to be normal with mean zero and $S.E. = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$ and hence, use z −distribution

- **Null hypothesis** $H_0: \mu_1 = \mu_2$
- **Alternative Hypothesis** $H_a: \ \mu_1 \neq \mu_2$
- **Calculation of test statistic:**
- $S.E. = 34.28$
- **Level of significance:** $\alpha = 0.05$
- **Critical value:** $z_\alpha = 1.96$
- **Decision:** Since the computed value of $|z| = 0.93$ is less than the table value 1.96, the hypothesis is accepted
- **The population means are equal** $\boldsymbol{\mu_1 = \mu_2}$

EXAMPLE-2

The means of two random samples of size 9 and 7 are 196.42 and 198.82 respectively. The sum of the squares of the deviation from the means are 26.94 and 18.73 respectively. Can the samples be considered to have been drawn from the same population?

SOLUTION

- Null hypothesis $H_0 : \mu_1 = \mu_2$
- Alternative Hypothesis $H_a : \ \mu_1 \neq \mu_2$
- Calculations of test statistic: Unbiased estimate of common population standard deviation is
- $S_p = \sqrt{\frac{\sum(x_i - \bar{x})^2 + \sum(y_i - \bar{y})^2}{n_1 + n_2 - 2}}$

- Standard error of the difference between the means $SE = s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
- $t = \frac{\bar{x}_1 - \bar{x}_2}{S.E.}$
- Level of significance: $\alpha = 0.05$
- Critical value: The table value of t at $\alpha = 0.05$ for $v = 9 + 7 - 2 = 14$ degree of freedom is 2.145
- Decision: Since the computed value of $|t| = 2.64$ is greater than the table value $t_\alpha = 2.145$, the null hypothesis is rejected
- The samples cannot be considered to have been drawn from the same population

EXAMPLE-3

Six guinea pigs injected with 0.5 mg of a medication took on an average 15.4 secs. to fall asleep with an unbiased standard deviation 2.2 secs., while six other guinea pigs injected with 1.5 mg. of the medication took over an average 11.2 secs. to fall asleep with an unbiased standard deviation 2.6 sec. Use 5% level of significance to test the null hypothesis that the difference in dosage has no effect.

SOLUTION

$\bar{x}_1 = 15.4, \bar{x}_2 = 11.2, S_1 = 2.2, S_2 = 2.6, n_1 = 6, n_2 = 6$

- **Null hypothesis** $H_0: \mu_1 = \mu_2$
- **Alternative Hypothesis** $H_a: \mu_1 \neq \mu_2$
- **Calculation of test statistic:** We are given unbiased standard deviations

∴ The unbiased estimate of the common population is given by

- $s_p = \sqrt{\frac{(n_1-1){S_1}^2+(n_2-1){S_2}^2}{n_1+n_2-2}}$
- $= 2.408$

- The standard error of the difference between the two means is given by
- $S.E. = s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
- $= 1.39$
- $\therefore t = \frac{\bar{x}_1 - \bar{x}_2}{S.E.}$
- $|t| = 3.02$
- Level of significance: $\alpha = 0.05$

- **Critical value**: The table value of t at $\alpha = 0.05$ for $v = 6 + 6 - 2 = 10$ degree of freedom is
- $t_\alpha = 2.228$
- **Decision**: Since the computed value of $|t| = 3.02$ is greater than the table value $t_\alpha = 2.28$
- the null hypothesis is rejected
- **The difference is significant**

EXAMPLE-4

Samples of two types of electric bulbs were tested for length of life and the following data were obtained. Test at 5% level of significance whether the difference in the sample means is significant.

	Type I	Type II
No of samples	8	7
Mean of the samples (in hours)	1134	1024
Standard deviation (in hours)	35	40

SOLUTION

We have $\bar{x}_1 = 1134, \bar{x}_2 = 1024, s_1 = 35, s_2 = 40, n_1 = 8, n_2 = 7$

- Null hypothesis $H_0: \mu_1 = \mu_2$
- Alternative Hypothesis H_a: $\mu_1 \neq \mu_2$
- Calculation of test statistic: $s_p = \sqrt{\frac{n_1 {s_1}^2 + n_2 {s_2}^2}{n_1 + n_2 - 2}}$
- $= 40.19$

- The standard error of the difference between the two means is given by

$$S.E. = s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

- $= 20.8$
- $\therefore t = \frac{\bar{x}_1 - \bar{x}_2}{S.E.}$
- $\therefore |t| = 5.288$

- Level of significance: $\alpha = 0.05$
- Critical value: The table value of t at $\alpha = 0.05$ for $v = 8 + 7 - 2 = 13$ degree of freedom is
- $t_\alpha = 2.16$
- Decision: Since the computed value of $|t| = 5.68$ is greater than the table value $t_\alpha = 2.16,$
- the hypothesis is rejected
- $\therefore$ The difference is significant

EXAMPLE-5

The heights of six randomly chosen sailors are in inches: $63, 65, 68, 69, 71$ and 72. The heights of ten randomly chosen soldiers are: $61, 62, 65, 66, 69, 69, 70, 71, 72$ and 73. Discuss in the light that these data throw on the suggestion that the soldiers on an average are taller than sailors.

SOLUTION

We first calculate the mean and standard deviation of the heights of both sailors and soldiers $\bar{x}_1$, $\bar{x}_2$, $\sum(x_1 - \bar{x}_1)^2$, $\sum(x_2 - \bar{x}_2)^2$

$$\bar{x}_1 = \frac{\sum x_1}{N} = \frac{408}{6} = 68, \bar{x}_2 = \frac{\sum x_2}{N} = \frac{678}{10} = 67.8$$

$$\sum(x_1 - \bar{x}_1)^2 = 60, \sum(x_2 - \bar{x}_2)^2 = 153.60$$

- $s_p = \sqrt{\frac{\sum(x_1-\bar{x}_1)^2+\sum(x_2-\bar{x}_2)^2}{n_1+n_2-2}}$
- $= 3.9$

- Null hypothesis $H_0: \mu_1 = \mu_2$
- Alternative Hypothesis H_a: $\mu_1 < \mu_2$
- Calculation of test statistic: $t = \frac{\bar{x}_1 - \bar{x}_2}{S.E.}$
- $\text{S}.E. = s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
- $= 2.014$
- $\therefore t = \frac{\bar{x}_1 - \bar{x}_2}{S.E.}$
- $= 0.099$
- Level of significance: $\alpha = 0.05$

- Critical value: The table value of t at $\alpha = 0.05$ for $v = 6 + 10 - 2 = 14$ degree of freedom is $t_\alpha = -1.761$ $(one\ tailed)$
- Decision: Since the computed value of $\mathrm{t} = 0.099$ is greater than the table value $t_\alpha = -1.761$,
- the hypothesis is accepted
- The means are equal i.e. the suggestion that the soldiers on the average are taller than sailors cannot be accepted