Hypothesis Testing II

Dr. Md. Israt Rayhan
Professor
Institute of Statistical Research and Training (ISRT)
University of Dhaka
Email: israt@isrt.ac.bd

Example: Two-Tail Test

The average cost of a blood test in a hospital is said to be Tk.168. A random sample of 25 patients resulted in

$\bar{x}$ = Tk.172.50 and

s = Tk.15.40. Test at the

α = 0.05 level.

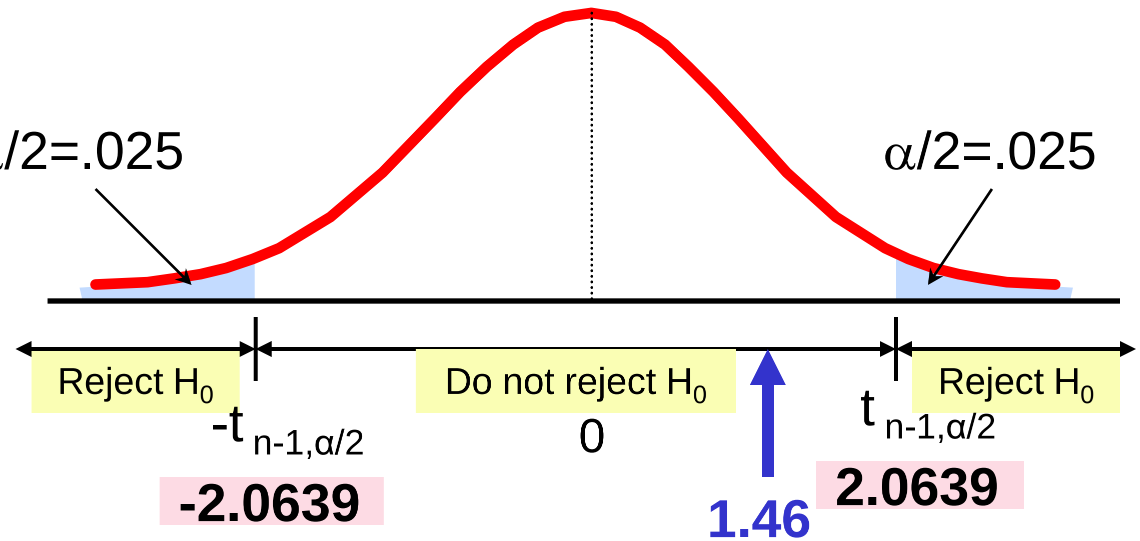
(Assume the population distribution is normal)

$H_0: \mu = 168$

$H_1: \mu \neq 168$

Example Solution: Two-Tail Test

H_0: $\mu = 168$
H_1: $\mu \neq 168$



- $\alpha = 0.05$
- **n = 25**
- **σ is unknown, so use a t statistic**
- **Critical Value:**

$t_{24\,,\,.025} = \pm\, 2.0639$

$$t_{n-1} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{172.50 - 168}{\frac{15.40}{\sqrt{25}}} = 1.46$$

Do not reject H_0: not sufficient evidence that true mean cost is different than Tk.168

Matched Pairs

Matched Pairs

Tests Means of 2 Related Populations

- Paired or matched samples
- Repeated measures (before/after)
- Use difference between paired values:

$$d_i = x_i - y_i$$

- Assumptions:
  - Both Populations Are Normally Distributed

Test Statistic: Matched Pairs

Matched Pairs

The test statistic for the mean difference is a t value, with n – 1 degrees of freedom:

$$t = \frac{\bar{d} - D_0}{\frac{s_d}{\sqrt{n}}}$$

Where

D_0 = hypothesized mean difference
s_d = sample standard dev. of differences
n = the sample size (number of pairs)

Matched Pairs Example

- Assume you send your Doctors to a training workshop. Has the training made a difference in the number of complaints? You collect the following data:

Doctors	Number of Complaints: Before (1)	After (2)	(2) - (1) Difference, d_i
DMC	6	4	- 2
Mitford	20	6	-14
SMC	3	2	- 1
RMC	0	0	0
CMC	4	0	- 4
			-21

$$\bar{d} = \frac{\sum d_i}{n}$$

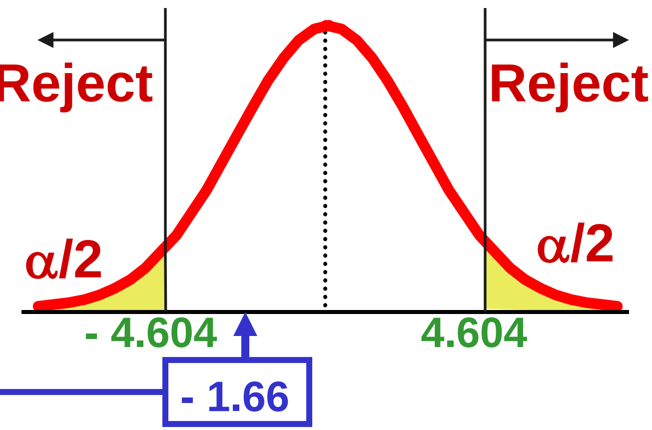
$$= -4.2$$

$$S_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}}$$

$$= 5.67$$

Matched Pairs: Solution

- Has the training made a difference in the number of complaints (at the α = 0.01 level)?



H_0: $\mu_x - \mu_y = 0$
H_1: $\mu_x - \mu_y \neq 0$

α = .01 $\quad\quad \bar{d}$ = - 4.2

Critical Value = ± 4.604

d.f. = n - 1 = 4

Test Statistic:

$$t = \frac{\bar{d} - D_0}{s_d/\sqrt{n}} = \frac{-4.2 - 0}{5.67/\sqrt{5}} = \boxed{-1.66}$$

Decision: Do not reject H_0

(t stat is not in the reject region)

Conclusion: There is not a significant change in the number of complaints.

Difference Between Two Means

Population means, independent samples

Goal: Form a confidence interval for the difference between two population means, $\boxed{\mu_x - \mu_y}$

- Different data sources
  - Unrelated
  - Independent
    - Sample selected from one population has no effect on the sample selected from the other population

Pooled Variance t Test: Example

You are an analyst. Is there a difference in the ICU patients dying in DMC and SMC per day? You collect the following data:

	DMC	SMC
Number	**21**	**25**
Average Dying	**3.27**	**2.53**
Sample std dev	**1.30**	**1.16**

Assuming both populations are approximately normal with equal variances, is there a difference in average number of dying ($\alpha = 0.05$)?

Calculating the Test Statistic

The test statistic is:

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\left(3.27 - 2.53\right) - 0}{\sqrt{1.5021\left(\frac{1}{21} + \frac{1}{25}\right)}} = \boxed{2.040}$$

$$S_p^2 = \frac{\left(n_1 - 1\right)S_1^{\,2} + \left(n_2 - 1\right)S_2^{\,2}}{\left(n_1 - 1\right) + \left(n_2 - 1\right)} = \frac{\left(21 - 1\right)1.30^2 + \left(25 - 1\right)1.16^2}{(21\text{-}1) + (25 - 1)} = 1.5021$$

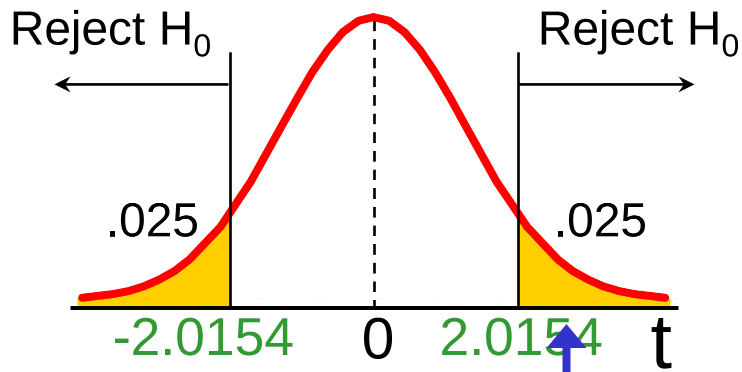
Solution

H_0: $\mu_1 - \mu_2 = 0$ i.e. ($\mu_1 = \mu_2$)
H_1: $\mu_1 - \mu_2 \neq 0$ i.e. ($\mu_1 \neq \mu_2$)

$\alpha = 0.05$

df = 21 + 25 - 2 = 44

Critical Values: t = ± 2.0154



2.040

Test Statistic:

$$t = \frac{3.27 - 2.53}{\sqrt{1.5021\left(\frac{1}{21} + \frac{1}{25}\right)}} = \boxed{2.040}$$

Decision:

Reject H_0 at $\alpha = 0.05$

Conclusion:

There is evidence of a difference in means.