Test of Hypothesis (2)

Md. Ismail Hossain Riday

Steps (t Test)

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

1. Identify the hypotheses
2. Choose the value of α; level of significance
3. Appropriate test statistic $(Z \; or \; t)$ and determine the value.
4. Compare this value with critical value obtained by using α. the critical value depends on the number of degrees of freedom.
5. Decision

Decision: We may reject H_0 if,

a) For left tailed: $t_{cal} \leq -(t_{\alpha,DF})$
b) For right tailed: $t_{cal} \geq +(t_{\alpha,DF})$

c) For two tailed: $|t_{cal}| \geq t_{\frac{\alpha}{2},DF}$

Degrees of Freedom

- Degrees of freedom are the maximum number of logically independent values, which may vary in a data sample.

Degrees of Freedom

- Consider a data sample consisting of five integers.
- The values of the five integers must have an average of six.
- We can consider four independent values; this number, four, represents the degrees of freedom.

Degrees of Freedom

- Difference between total sample and total number of parameters used in the test.

$$H_0: \mu = \mu_0$$

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

There is one parameter

Sample size = n

$$\underbrace{n \quad - \quad 1}_{\text{Degrees of Freedom}}$$

Degrees of Freedom

- Difference between total sample and total number of parameters used in the test.

H_0: $\mu_1 = \mu_2$

n_1 n_2

There are two parameter

Sample size $= n_1 + n_2$

$$\underbrace{n_1 + n_2 - 2}_{\text{Degrees of Freedom}}$$

Degrees of Freedom

- Difference between total sample and total number of parameters used in the test.

$H_0: \mu = \mu_0$

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

There is one parameter

Sample size = n

$$\underbrace{\boxed{n}\ \boxed{-}\ \boxed{1}}_{\text{Degrees of Freedom}}$$

Recap...

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

1. Identify the hypotheses
2. Choose the value of α; level of significance
3. Appropriate test statistic $(Z \text{ or } t)$ and determine the value.
4. Compare this value with critical value obtained by using α. the critical value depends on the number of degrees of freedom.
5. Decision

$$df = (n-1)$$

Decision: We may reject H_0 if,

a) For left tailed: $t_{cal} \leq -(t_{\alpha,DF})$
b) For right tailed: $t_{cal} \geq +(t_{\alpha,DF})$

c) For two tailed: $|t_{cal}| \geq t_{\frac{\alpha}{2},DF}$

Example

$$H_0: \mu = 100$$

$$H_1: \mu \neq 100$$

$$\alpha = 0.05$$

- A random sample of 10 boys had the following I. Q's: 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Do these data support the assumption of a population mean I. Q. of 100?

$$t_{CAL} = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$t_{CAL} = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{(97.2 - 100)}{\frac{14.27}{\sqrt{10}}} = -0.62$$

$$\bar{X} = \frac{\sum X_i}{n} = \frac{70 + 120 + \cdots + 100}{10} = 97.2$$

$$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}} = 14.27$$

Example

$H_0: \mu = 100$

$H_1: \mu \neq 100$

$\alpha = 0.05$

$$t_{CAL} = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = -0.62$$

- A random sample of 10 boys had the following I. Q.'s: 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Do these data support the assumption of a population mean I. Q. of 100?

$$t_{TAB} = t\left(\frac{\alpha}{2}, DF\right)$$

$$t_{TAB} = t(0.025, DF)$$

$$t_{TAB} = t\big(0.025, (n-1)\big)$$

$$t_{TAB} = t(0.025{,}9)$$

$$t_{TAB} = 2.262$$

Decision: We may reject H_0 if,

a) For left tailed: $t_{cal} \leq -(t_{\alpha,DF})$

b) For right tailed: $t_{cal} \geq +(t_{\alpha,DF})$

c) For two tailed: $|t_{cal}| \geq t_{\frac{\alpha}{2},DF}$

$$|t_{CAL}| < t_{TAB}$$

$$May\ not\ reject\ H_0$$

Example

$H_0: \mu = 110$

$H_1: \mu < 110$

$\alpha = 0.01$

$$t_{CAL} = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

- Is the temperature required to damage a computer on the average less than 110 degrees? Because of the price of testing, twenty computers were tested to see what minimum temperature will damage the computer. The damaging temperature averaged 109 degrees with a standard deviation of 3 degrees. (Use 99% confidence level)

DF = 20 − 1
= 19

Errors in Test

- Any decision we make based on a hypothesis test may be incorrect

- Because we have used partial information obtained from a sample to draw conclusion about the entire population.

Errors in Test

- Two types of error can be made

a. Type I error

b. Type II error

Type I Error α

- The probability of rejecting null hypothesis when it is true.

- For example, Let's say a medical researcher is conducting a clinical trial to test a new drug's effectiveness in treating a certain condition. Null hypothesis is “The drug has no effect”, and alternative is “The drug has effect”.

- If, in reality, the drug has no effect (null hypothesis is true), but the statistical analysis leads the researcher to reject the null hypothesis, it would be a Type I error.

Type II Error $\boxed{\beta}$

- The probability of accepting null hypothesis when it is false.

- For example, Let's say a medical researcher is conducting a clinical trial to test a new drug's effectiveness in treating a certain condition. Null hypothesis is “The drug has no effect”, and alternative is “The drug has effect”.

- If, in reality, the drug has effect (null hypothesis is true), but the statistical analysis leads the researcher to accept the null hypothesis, it would be a Type II error.

Type I and II

	Actual situation		
Decision from sample			

Type I and II

	Actual situation		
Decision from sample		H_0 is true	H_0 is false

Type I and II

	Actual situation		
Decision from sample		H_0 is true	H_0 is false
	H_0 accepted		
	H_0 rejected		

Type I and II

	Actual situation		
Decision from sample		H_0 is true	H_0 is false
	H_0 accepted		
	H_0 rejected		

Type I and II

		Actual situation	
Decision from sample		H_0 is true	H_0 is false
	H_0 accepted	No error	
	H_0 rejected		

Type I and II

	Actual situation		
Decision from sample		H_0 is true	H_0 is false
	H_0 accepted	No error	
	H_0 rejected		

Type I and II

	Actual situation		
Decision from sample		H_0 is true	H_0 is false
	H_0 accepted	No error	Type II error
	H_0 rejected		

Type I and II

	Actual situation		
Decision from sample		H_0 is true	H_0 is false
	H_0 accepted	No error	Type II error
	H_0 rejected		

Type I and II

	Actual situation		
Decision from sample		H_0 is true	H_0 is false
	H_0 accepted	No error	Type II error
	H_0 rejected	Type I error	

Type I and II

		Actual situation	
Decision from sample		H_0 is true	H_0 is false
	H_0 accepted	No error	Type II error
	H_0 rejected	Type I error	

Type I and II

	Actual situation		
		H_0 is true	H_0 is false
Decision from sample	H_0 accepted	No error	Type II error
	H_0 rejected	Type I error	No error

Reduce errors

- We can reduce type II error by increasing sample size.

- We can reduce type I error by decreasing level of significance.

P-value

- For left tailed: $p\ value = P(Z_{cal})$
- For right tailed: $p\ value = 1 - P(Z_{cal})$
- For two tailed: $p\ value = 2 \times (1 - P(|Z_{cal}|)$

- Probability value (P-value)
- A measure used in statistical hypothesis testing
- To quantify the evidence against a null hypothesis.

If $P\text{-}value < \alpha$,
we may reject H_0

P-value

If $P\text{-}value < \alpha$, we may reject H_0

- Calculate p-value for left tailed $Z_{cal} = -1.19$

Left tailed test

- $P\text{-}value = P(-1.19) = 0.1170$

- Calculate p-value for two tailed $Z_{cal} = -1.71$

Two tailed test

- $P\text{-}value = 2 \times (1 - P(|-1.71|) = ???$

Mathematical exercise

To access additional mathematical problems,

please refer to the PDF lecture notes.

Thank You