Some comments on α (prob. of type I error). ①

The significance level of a test is the max^m. allowable type I error probability.

This will enable us to define a rejection region.

Disadvantages of significance level as a reliable counter-measure

i) One may not have a clear idea of what an appropriate max^m. allowable type I error should be for a given test statistic.

ii) Significance level may also be sensitive to minor changes in sample statistic.

We will generally not discuss β in this course.

iii) Recall β = probability of type II error. Increasing α reduces β & reducing α increases $\beta \Rightarrow$ there is always a trade-off.

Recall the table!

Truth (Actual situation of the population)

Decision	H_0 is TRUE	H_0 is false
Do NOT reject H_0	Correct decision w/ prob. $(1-\alpha)$	Incorrect decision w/ prob. β
Reject H_0	In-correct Decision w/ prob. α	Correct decision w/ prob. $(1-\beta)$

Type II error

Type I error

Some useful hypothesis tests.

Here is an application of t-Distribution in hypothesis tests!

* Two sample test for mean

$$X_1 \sim N(\mu_1, \sigma^2)$$
$$X_2 \sim N(\mu_2, \sigma^2)$$

Note σ^2 is same for both sets of random samples

Let there be n_1 samples of 1st population &
n_2 samples of 2nd population.

$$X_{1i} \sim N(\mu_1, \sigma^2) \quad ; \quad i = 1, 2, \ldots, n_1$$
$$X_{2j} \sim N(\mu_2, \sigma^2) \quad ; \quad j = 1, 2, \ldots, n_2$$

Step 1:

$H_0: \mu_1 = \mu_2$

double sided test

single sided test

$H_1: \mu_1 \neq \mu_2$ (or $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$)

Also set α (level of significance)

Step 2: Test statistic, t

+

step 3

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where

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

Under Ho,

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

Where S_j^2 is the sample variance of the j^{th} group.

Step 4 : Rejection (Critical region)

* Reject H_0 in favor of H_1 $(\mu_1 \neq \mu_2)$ if

$$|t| \geq t\left(\frac{\alpha}{2}, n_1 + n_2 - 2\right)$$

* Reject H_0 in favor of H_1 $(\mu_1 > \mu_2)$ if

$$t \geq t(\alpha, n_1 + n_2 - 2)$$

* Reject H_0 in favor of H_1 $(\mu_1 < \mu_2)$ if

$$t \leq -t(\alpha, n_1 + n_2 - 2)$$

example

In comparison of two gasoline brands, a consumer survey reveals the following :-

* A full tank of brand (A) requires 4 cans and covers 546 kms w/ a std. devn. of 31 kms.
* A full tank of brand (B) requires 4 cans and covers 492 kms w/ a std. devn. of 26 kms.

Assume that both brand populations (A) & (B) are sampled from Normal D's w/ equal variances.

Test: $H_0 : \mu_1 = \mu_2$ vs $H_1 : \mu_1 > \mu_2$ at $\alpha_{max} = 0.05$

This will tell customers which brand offers better mileage.

Soln:-

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 > \mu_2$

$\alpha = 0.05$

$\bar{X}_1 = 546$ $\quad\quad$ $\bar{X}_2 = 492$

$S_1 = 31$ $\quad\quad$ $S_2 = 26$

$n_1 = 4$ $\quad\quad$ $n_2 = 4$

$$S^2 = \frac{(4-1)31^2 + (4-1)26^2}{4+4-2} \Rightarrow S = 28.609$$

$$\therefore t \text{ (under } H_0) = \frac{546-492}{28.609\sqrt{\frac{1}{4}+\frac{1}{4}}} = 2.67$$

from table, $t(0.05, 6) = 1.943$ (one tail t - Dn)

Inference

$\because t_{cal} = 2.67 > t(0.05, 6) = 1.943$
(from table)

$\Rightarrow$ Reject H_0 in favor of H_1

— x —

Next Lecture → we will learn how to read values from sampling D^n tables!