Mathematical Statistics II

Ch.8 Tests of Statistical Hypotheses

Jungsoon Choi

jungsoonchoi@hanyang.ac.kr

Table of Contents

Introduction

Introduction

- Hypothesis: a statement about the population
- Testing hypothesis: the process to choice if the hypothesis is true or not.
- Null hypothesis(H_0)
- Alternative hypothesis(H_1)
- simple vs composite hypothesis

Q) new argue.

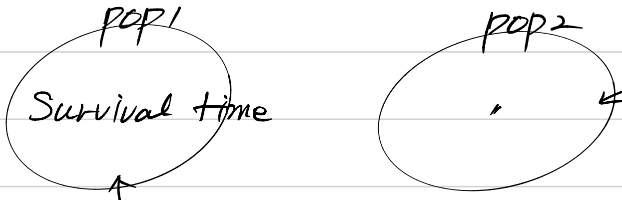
↓

collect data

↓

data analysis. (Statistical Testing).

H_0: No diff vs H_1: new argue



For the current drug μ_Y.
old

For new drug μ_X

H_0 $\mu_X = \mu_Y$

H_1 : $\mu_X > \mu_Y$

Simple H_0 : $\mu = 35$ H_1 : $\mu = 45$

composite H_0 : $\mu < 45$ H_1 : $\mu > 45$

	Testing result	
True	H_0 true	H_1 true
H_0 true	$1-\alpha$	Type I error $= \alpha$
H_1 true	error $= \beta$	$1-\beta$

can not reduce type I & II error at the same time.

Type I & II errors

- Type I error, significance level (α) $=$ P(reject H_0 | H_0 is true)
- Type II error (β) $=$ P(accept H_0 | H_1 is true)
- Power $= 1 - \beta =$ P(reject H_0 | H_1 is true)

	Test results	
True	accept H_0	accept H_1
H_0 true	$1-\alpha$	α
H_1 true	β	$1-\beta$

accept H_1

Type I error (α) = P(reject H_0 | H_0 true)

Type II error (β) = P(accept H_0 | H_1 is true).

" reject H_1

power = $1-\beta$.

Test statistic $u(X_1, \cdots, X_n)$ under H_0 is true

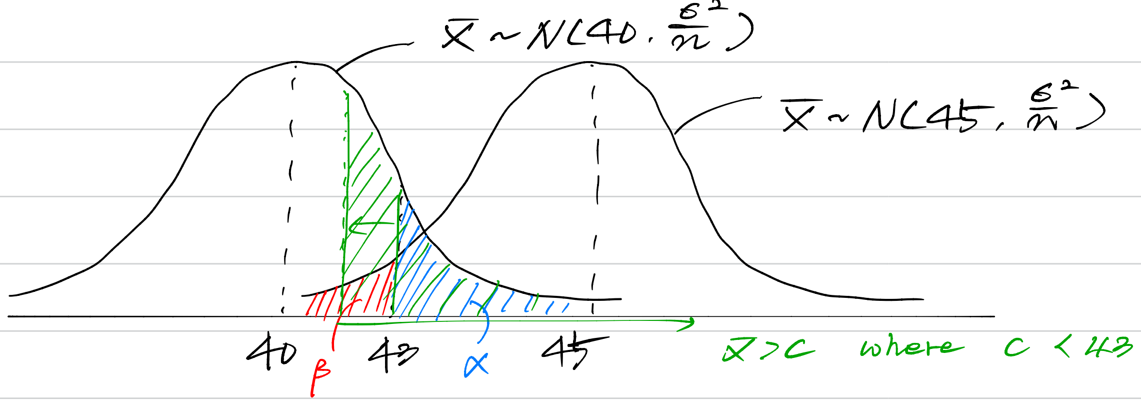
rejection (critical) region : reject H_0 if $u(x_1, \cdots, x_n) \in C$

Example) $X_i \sim N(\mu, \sigma^2)$ iid, $i = 1, \cdots, n$ σ^2 : known.

$H_0 : \mu = 40$ vs $H_1 : \mu = 45$ $\bar{x}_1 = 42$, $\bar{x}_2 = 44$

i) Under H_0 is true, $\bar{X} \sim N(40, \frac{\sigma^2}{n})$

ii) under H_1 is true, $\bar{X} \sim N(45, \frac{\sigma^2}{n})$



If the rejection region is $\bar{X} > 43$, then

(α) Type I error = $P(\bar{X} > 43 \mid \bar{X} \sim N(40, \frac{\sigma^2}{n}))$

(β) Type II error = $P(\bar{X} \leq 43 \mid \bar{X} \sim N(45, \frac{\sigma^2}{n}))$

rejection region reject H_0 if $u(x_1, \cdots, x_n) \in C$

As Type I error (α) increase

i) rejection region $\uparrow$ (eg $\bar{X} > 42$)

ii) Type II error (β) decreases

$\Rightarrow$ $\alpha \uparrow \Leftrightarrow \beta \downarrow$

cf) $MSE(\hat{\theta}) = Var(\hat{\theta}) + \{Bias(\hat{\theta})\}^2$

1) $Bias(\hat{\theta}) = 0$

2) $\hat{\theta}$ among U.E with the smallest variance

↳ Best estimator

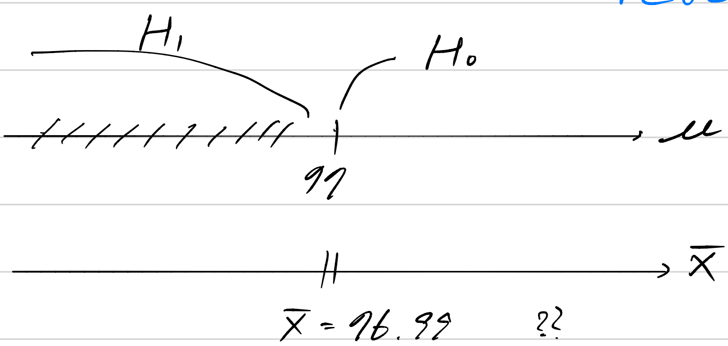
Type I error make more risky situation.
Fix Type I error (α) value, reduce Type II error (β)
α = 0.05, 0.01

Ex) $X_i \sim N(\mu, 4)$ $i = 1, \cdots, n = 25$

$H_0 : \mu = 77$ vs $H_1 : \mu < 77$ $\quad \alpha = 0.05$

reject H_0 if $\bar{X} < C$ s.t $c < 77$

rejection region.



Under H_0 is true, $\bar{X} \sim N(77, \frac{4}{25})$

$$0.05 = P(\bar{X} < c_{0.05} \mid \bar{X} \sim N(77, \tfrac{4}{25}))$$

$$= P\left(\frac{\bar{X} - 77}{2/5} < \frac{c_{0.05} - 77}{2/5}\right) = P\left(Z < \frac{c_{0.05} - 77}{2/5}\right)$$

$$= \Phi\left(\frac{c_{0.05} - 77}{2/5}\right)$$

$$\frac{c_{0.05} - 77}{2/5} = -1.645$$

$$c_{0.05} = 77 - 1.645 \times \frac{2}{5} = 76.342$$

reject H_0 if $\bar{X} < 76.342$ when $\alpha = 0.05$

depend on Type I error

$\bar{x}_1 = 75 \Rightarrow$ reject H_0

$\bar{x}_2 = 76.5 \Rightarrow$ accept H_0

Some Concepts

- Test statistic: $u(X_1, \ldots, X_n)$ which is computed under the assumption that null hypothesis is true.
- Rejection (critical) region: $u(X_1, \ldots, X_n) \in C$ then reject H_0
  i.e reject H_0 if $u(X_1, \ldots, X_n) \in C$

How to conduct the tests of statistical hypotheses

- Establish the null hypothesis and alternative hypothesis.
- Find the rejection (critical) region $(u(X_1, \ldots, X_n) \in C)$ given the Type I error, α.
- Given the sample data, compute a test statistic.

Example

Let $X_1, \ldots, X_n \sim N(\mu, 4)$ with $n = 25$.

$$H_0 : \mu = 77 \quad vs \quad H_1 : \mu < 77$$

Find the rejection region when $\alpha = 0.05$.

$\bar{X}$ under $H_0 : \mu = 77$

$\Leftrightarrow$ ~ under $\bar{X} \sim N(77, \frac{4}{25})$

$\Leftrightarrow$ $\frac{\bar{X} - 77}{2/5} = Z \sim N(0,1)$ $\quad Z = \frac{\bar{X} - 77}{2/5} < -1.645$

test statistic

P-value

Definition

Let $C = \{(x_1, \ldots, x_n) | u(x_1, \ldots, x_n) \leq c\}$ be the rejection region.

Let u^* be the observed value of $u(X_1, \ldots, X_n)$.

p-value=observed significance level

- α=P$[u(X_1, \ldots, X_n) \leq \overset{16.342}{C_\alpha} \mid \mathrm{H}_0 \text{ is true}]$
- p-value=P$[u(X_1, \ldots, X_n) \leq \underset{16}{u^*} \mid \mathrm{H}_0 \text{ is true}]$
- P-value does not depend on α.

$\alpha = P(\bar{X} \le c_\alpha \mid H_0 \text{ true})$

$p\text{-value} = P(\bar{X} \le \underset{16}{\bar{x}} \mid H_0 \text{ is true}).$

$\alpha = P(u(X_1, \cdots, X_n) \le c_\alpha \mid H_0 \text{ is true})$

$p\text{-value} = P(u(X_1, \cdots X_n) \le u^* \mid H_0 \text{ is true}).$

u^* = observed value of $u(X_1, \cdots, X_n)$ from the dataset.

$H_0 : \bar{X} \sim N(17, \frac{4}{25})$

$\bar{X} < 16.342.$ when $\alpha = 0.05$

not test statistic

$Z = \dfrac{\bar{X} - 17}{2/5} < -1.645.$

test statistic

(H_0 is true 가 내포되어야 함).

Review).

(α). Type I error significance level
$= P(\text{reject } H_0 \mid H_0 \text{ true})$.

(β) Type II error $= P(\text{accept } H_0 \mid H_1 \text{ true})$.

Test statistic : $u(X_1, \cdots, X_n)$ under H_0 true

Rejection region : reject H_0 if $u(X_1 \cdots X_n) \in C$

* p-value

Assume that $\alpha = P[u(X_1 \cdots X_n) \le C_\alpha \mid H_0 \text{ is true}]$

p-value $= P[u(X_1 \cdots X_n) \le u^* \mid H_0 \text{ is true}]$.

where u^* is the observed value of $u(X_1 \cdots X_n)$ using the data.

[Note].

1). Using the definition of Type I error, we can obtain the rejection region before collecting the dataset.

2). After collecting the dataset, p-value can be calculated. To calculate a p-value we don't need the Type I error information But to make the statistical decision, we use

i) data info + rejection region

ii) p-value + Type I error.

Example

Let $X_1, \ldots, X_n \sim N(\mu, 4)$.

$$H_0 : \mu = 77 \quad vs \quad H_1 : \mu < 77$$

The sample mean is 76.1 from 25 samples. Find the p-value and make the conclusion under $\alpha = 0.05$.

cf). Reject H_0 if $\bar{X} < 76.342$ when $\alpha = 0.05$

Example). $\bar{x} = 76.1$ $\quad n = 25$

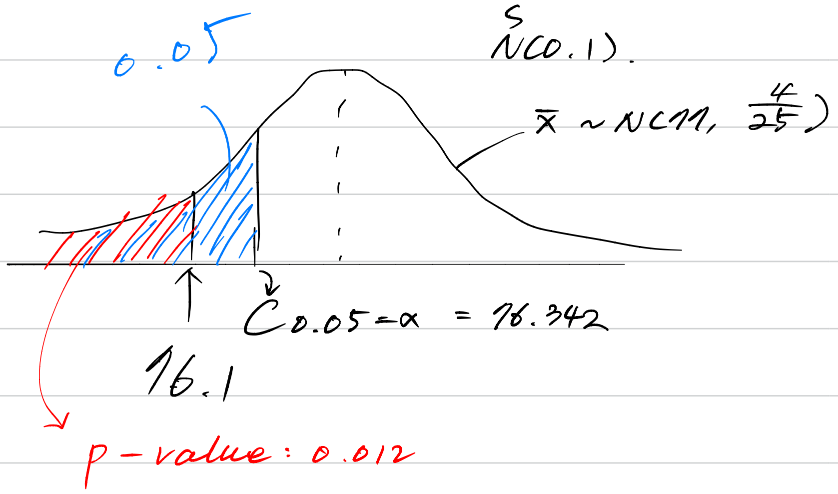
reject H_0 if $\bar{X} < C_\alpha$

p-value $= P(\bar{X} < \overset{\mu^*}{76.1} \mid H_0 : \text{true})$.

$= P(\bar{X} < 76.1 \mid \bar{X} \sim N(\mu, \frac{4}{25}))$

$= P(\frac{\bar{X} - \mu}{2/5} = Z < \frac{76.1 - \mu}{2/5})$

$= P(\underset{\sim N(0,1)}{Z} < -2.25) = 0.012$.



Compare (under Type I error(α) $= 0.05$).

i) $\bar{x} = 76.1$ vs $C_{0.05} = 76.342$

ii) p-value $= 0.012$ vs $\alpha = 0.05$

Ch8.1 Tests about One Mean

Tests about One Mean

Example 8.1-1

Let X be the breaking strength of a steel bar. We assume that $X \sim N(50, 36)$ from the Method 1 and $X \sim N(55, 36)$ from the Method 2. Then, $X \sim N(\mu, 36)$.

$$H_0 : \mu = 50 \quad vs \quad H_1 : \mu = 55$$

Suppose that the rejection region is $C = \{\overline{x} | \overline{x} > 53\}$ and the sample size is 16. Find the type I and type II errors.

Type I error : $P(\bar{X} > 53 \mid H_0 : \text{true})$.

$$P(\bar{X} > 53 \mid \bar{X} \sim N(50, \tfrac{36}{16}))$$

$$P\left(\frac{\bar{X}-50}{6/4} > \frac{53-50}{6/4}\right)$$

$$P\left(Z > \frac{3}{3/2} = 2\right) = P(Z > 2)$$

Type II error

$$P(\bar{X} < 53 \mid H_1 : \text{true}).$$

$$P(\bar{X} < 53 \mid \bar{X} \sim N(55, \tfrac{36}{16}))$$

$$P\left(Z < \frac{53-55}{6/4}\right) = P\left(Z < -\frac{4}{3}\right).$$

$X_i \sim N(\mu, 36)$ iid $\quad i = 1, \cdots, 16$

$H_0 : \mu = 50, \quad H_1 : \mu = 55$

$\bar{X} \sim N(\mu, \frac{36}{16} = \frac{9}{4} = (\frac{3}{2})^2)$

i) $\alpha = P(\bar{X} > 53 \mid \mu = 50) = P(\bar{X} > 53 \mid \bar{X} \sim N(50, (\frac{3}{2})^2))$

$\quad = P\left(Z = \frac{\bar{X}-50}{3/2} > \frac{53-50}{3/2}\right) = P(Z > 2) = 0.0228$

ii) $\beta = P(\text{accept } H_0 \mid H_1 : \text{true}) = P(\bar{X} \leq 53 \mid \bar{X} \sim N(55, (\frac{3}{2})^2))$

$\quad = P\left(Z = \frac{\bar{X}-55}{3/2} \leq \frac{53-55}{3/2}\right) = P(Z < -\frac{4}{3}) \doteqdot 0.0913$

test statistic. ($\because$ under H_0 true).

Not test statistic. ($\because$ under H_1 true).

Example 8.1-2

Suppose that the population follows a $N(\mu, 100)$ distribution. We have the sample mean of $\overline{x} = 62.75$ from 52 samples.

$$H_0 : \mu = 60 \quad vs \quad H_1 : \mu > 60$$

Find the p-value.

$X_i \sim N(\mu, 100) \quad i = 1, \cdots, 52, \quad \text{iid}.$

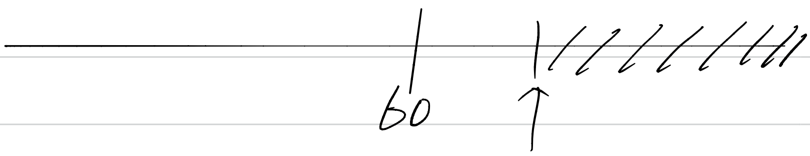
$\bar{X} \sim N(\mu, \frac{100}{52}) \quad \bar{x} = 62.75.$

$\text{p-value} = P(\bar{X} > 62.75 \ \mu^* \mid H_0 : \text{true}).$

$= P(\bar{X} > 62.75 \mid \bar{X} \sim N(60, \frac{100}{52})).$

$= P(Z = \frac{\bar{X} - 60}{10/\sqrt{52}} > \frac{62.75 - 60}{10/\sqrt{52}}) = P(Z > 1.983).$

$\fallingdotseq 0.0237$



if $\bar{X} > C_\alpha$, reject H_0.

Example 8.1-3

Let $X \sim N(\mu, \sigma^2)$ be the size of tumor in 15 days induced in a mouse.

$$H_0 : \mu = 4 \quad vs \quad H_1 : \mu \neq 4$$

We have 9 samples and $\alpha = 0.1$.

- Find the rejection region.
- Given the data information, $\overline{x} = 4.3$ and $s = 1.2$, compute the p-value and make the conclusion.

$\bar{X} = 4.3 \qquad s = 1.2. \qquad \frac{4.3-4}{1.2/3} = \frac{0.3}{1.2/3} = \frac{1}{4/3}$

$= \frac{3}{4}$

$= 0.75.$

$$\text{p-value} = P(\bar{X} > 4.3 \mid H_0 : \text{true}).$$

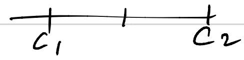
$$= P\left(\frac{\bar{X}-4}{1.2/3} > \frac{4.3-4}{1.2/3} = 0.75 \right).$$

$$= P(T > 0.75).$$

$X \sim N(\mu, \sigma^2)$. $\quad \bar{X} \sim N(\mu, \frac{\sigma^2}{9})$

$H_0 : \mu = 4 \qquad H_1 : \mu \neq 4$.

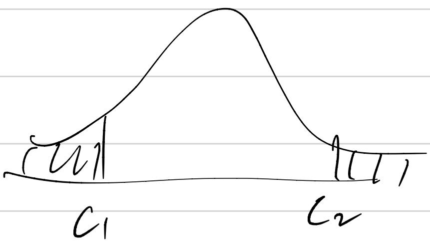
$\alpha = 0.1$.



$$\alpha = P(\bar{X} \geq C_2 \text{ or } \bar{X} \leq C_1 \mid \mu = 4)$$

$$0.1 = P\left(\frac{\bar{X}-4}{\sigma/3} \geq \frac{C_2-4}{\sigma/3} \text{ or } \frac{\bar{X}-4}{\sigma/3} \leq \frac{C_1-4}{\sigma/3} \right)$$

$$= P\left(Z \geq \frac{C_2-4}{\sigma/3} \text{ or } Z \leq \frac{C_1-4}{\sigma/3} \right).$$



$C_1 = -C_2$.

$$\therefore P\left(Z \geq \frac{C_2-4}{\sigma/3} \right) = 0.05.$$

$$Z_{0.05} = 1.645$$

$$\frac{C_2-4}{\sigma/3} = Z_{0.05}.$$

$$C_2 = 4 + Z_{0.05} \cdot \frac{\sigma}{3}$$

i). σ known.

rejection region $=$ $\bar{X} \geq 4 + Z_{0.05} \cdot \frac{\sigma}{3}$

or

$\bar{X} \leq 4 - Z_{0.05} \cdot \frac{\sigma}{3}$

ii) σ unknown.

$$0.1 = P\left(\frac{\bar{X}-4}{S/3} \geq \frac{C_2-4}{S/3} \text{ or } \frac{\bar{X}-4}{S/3} \leq \frac{C_1-4}{S/3} \right)$$

$$0.05 = P\left(T \geq \frac{C_2-4}{S/3} = t_{0.05}(8) \right)$$

$$C_2 = 4 + t_{0.05}(8) \frac{S}{3}$$

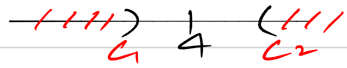
$\therefore$ rejection region : $\bar{X} \geq 4 + t_{0.05}(8) \frac{S}{3}$

$\bar{X} \leq 4 - t_{0.05}(8) \frac{S}{3}$

$X_i \sim N(\mu, \sigma^2)$ iid $\quad i = 1, \cdots, 9 \quad \alpha = 0.1$

$\Rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

$H_0 : \mu = 4$ vs $H_1 : \mu \neq 4$



i) reject H_0 if $\bar{X} < c_1$ or $\bar{X} > c_2$

$\alpha = 0.1 = P(\bar{X} > c_2$ or $\bar{X} < c_1 \mid \mu = 4)$.

$= P(\bar{X} > c_2$ or $\bar{X} < c_1 \mid \bar{X} \sim N(4, \frac{\sigma^2}{9}))$

$= P\left(T = \frac{\bar{X}-4}{s/3} > \frac{c_2 - 4}{s/3} = d_2 \text{ or } T < \frac{c_1 - 4}{s/3} = d_1\right)$

$(T \sim t(8))$.

$= P(T < d_1) + P(T > d_2) = P(|T| > d_2) = 2P(T > d_2)$.

we consider the equal probability for $P(T < d_1)$ & $P(T > d_2)$.

$\therefore d_1 = -d_2$ (b/c dist^n of T is sym).

$d_2 = t_{0.05}(8) = 1.86 \qquad d_1 = -1.86$.

$\Rightarrow c_2 = 4 + 1.86 \times \frac{s}{3} \qquad c_1 = 4 - 1.86 \times \frac{s}{3}$

$\therefore$ Rejection region is

i) $\bar{X} < 4 - 1.86 \times \frac{s}{3}$ or $\bar{X} > 4 + 1.86 \times \frac{s}{3}$

ii) $|T| = \left|\frac{\bar{X}-4}{s/3}\right| > t_{0.05}(8) = 1.86$

2) p-value $= P(|T| > |t|)$ where $T = \frac{\bar{X}-4}{s/3}$

$= P(|T| > 0.75) \qquad t = \frac{\bar{x}-4}{s/3} = \frac{0.3}{0.4}$

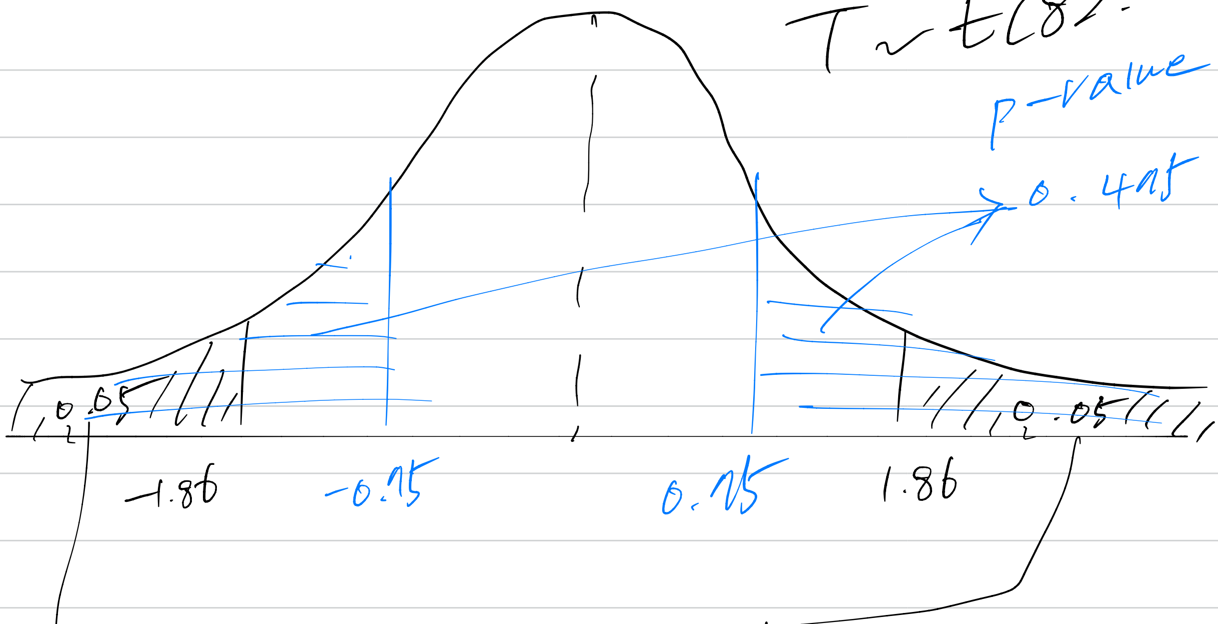
$= 2P(T > 0.75) \qquad = 0.75$

$\approx 0.475 > \alpha = 0.1$

In (1) question, $|t| = 0.75 \leq t_{0.05}(8) = 1.86$.

$\Rightarrow$ Not reject H_0 when $\alpha = 0.1$

$T \sim t(82)$.



sum of two areas $= 0.1 = \alpha$.

$H_0 : \mu = 4$ vs $H_1 : \mu \neq 4$

Rejection H_0 if $\bar{X} > c_2$ or $\bar{X} < c_1$

$$\alpha = 0.1 = P(\bar{X} > c_2 \text{ or } \bar{X} < c_1 \mid \mu = 4).$$

$$= \underbrace{P(T < d_1)}_{0.05} + \underbrace{P(T > d_2)}_{0.05}$$

$$= P(|T| > d_2) = 2P(T > d_2)$$

where $d_2 = \dfrac{c_2 - 4}{S/3}$

equal prob. for $P(T < d_1)$ & $P(T > d_2)$

$\therefore d_1 = -d_2$

$d_2 = t_{0.05}(8) = 1.86.$ $\quad d_1 = -1.86$

$c_2 = 4 + 1.86 \times \frac{S}{3}$ $\quad c_1 = 4 - 1.86 \cdot \frac{S}{3}$

$\therefore$ Rejection region is

① $\bar{X} < 4 - 1.86 \times \frac{S}{3}$ or $\bar{X} > 4 + 1.86 \times \frac{S}{3}$

② $|T| = \left| \dfrac{\bar{X} - 4}{S/3} \right| > t_{0.05}(8) = 1.86.$

p-value $= P(|T| > |t|)$ where $t = \dfrac{\bar{x} - 4}{S/3} = 0.75$

$= P(|T| > 0.75) = 2P(T > 0.75) \approx 0.475 > \alpha$

$T \sim t(8).$ $\quad \alpha = 0.1$

$\Rightarrow$ Not reject H_0 when $\alpha = 0.1$

① rejection region $\quad 3.25 < 4.3 < 4.75$

② p-value vs α $\quad 0.475 > 0.1$

Example 8.1-5 (paired t-test)

Let W be the time difference in running between the before and after the program. Suppose that $W \sim N(\mu_W, \sigma_W^2)$

$$H_0 : \mu_W = 0 \quad vs \quad H_1 : \mu_W > 0$$

We have $\overline{w} = 0.0788$ and $s_w = 0.2549$ from 24 samples and $\alpha = 0.05$.

- Find the rejection region.
- Compute the test statistic and make the conclusion.

$W \sim N(\mu_w, \sigma_w^2)$

$H_0 : \mu_w = 0$, $H_1 : \mu_w > 0$ Reject H_0 if $\overline{w} > c_\alpha$

$\alpha = 0.05 = P(\overline{w} > c_\alpha \mid \mu_w = 0)$.

$= P\left(T = \frac{\overline{w} - 0}{s_w/\sqrt{24}} \geq \frac{c_\alpha}{s_w/\sqrt{24}} \right)$, $T \sim t(23)$.

test statistic $t_{0.05}(23) = 1.714$

Reject H_0 if $\overline{w} > c = 1.714 \times \frac{s}{\sqrt{24}}$

$T = \frac{\overline{w}}{s_w/\sqrt{24}} \Rightarrow \frac{0.0788}{0.2549/\sqrt{24}} \doteqdot 1.518$

H_0 is not rejected $(\because 1.518 < 1.714)$.

write down

H_0, H_1, rejection region