- Recall: last time we defined what is hypothesis testing.
  - parameter $\theta \in \Theta$: controls how to generate data.
  - Random Variable. $X \in \mathcal{X}$: data.
  - Hyothesis: H_0: a subset $\Theta_0 \subset \Theta$. (null hypothesis).
  - Test: R: rejection ~~re~~ region. $R \subset \mathcal{X}$.

    (we reject H_0 if $X \in R$)

  - test say accept / test say "reject"

	$X \notin R$	$X \in R$
$\theta \in \Theta$	✓	type I error
$\theta \notin \Theta_0$	type II error	✓

too cautious. or reject too often.

test too loose.

- (power function of a test):

$$\beta(\theta) = \mathbb{P}_\theta(X \in R)$$

the probability to trigger rejection under θ.

(size of a test):

$$\underset{\sim}{\alpha} = \sup_{\theta \in \Theta_0} \beta(\theta).$$

the worst case scenario: what is the largest probability to reject H_0 when H_0 is actually true.

Ideally: we want $\beta(\theta)$ to be <u>low</u> when $\theta \in \Theta_0$ (H_0 is true.)

and <u>high</u> when $\theta \notin \Theta_0$

- Example of test :

- Say $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma)$, we want to ask : is $\mu \leq 0$?

Hypothesis :

$$\Theta_0 = \{\mu \leq 0\}. \qquad \text{null hypothesis}$$

$$\Theta_1 = \{\mu > 0\} \qquad \text{alternative hypothesis}$$

Test :

$$R = \{ \underbrace{(X_1, \cdots, X_n)}_{\in \mathbb{R}^n} : T(x_1, \cdots, x_n) > \underbrace{c}_{\text{to be decided later.}} \}$$

$T(x_1, \cdots, x_n) = \frac{1}{n}(X_1 + \cdots + X_n).$

c : critical threshold.

power function:

$$\beta(\mu) = \mathbb{P}_\mu(\text{do rejection})$$

$$= \mathbb{P}_\mu\Big(\underbrace{\tfrac{1}{n}(X_1 + \cdots + X_n)}_{\bar{X}} > c\Big) \qquad X_i \sim N(\mu, \sigma)$$

CDF

$$= \underline{\mathbb{P}_\mu(\bar{X} > c)}. \qquad \bar{X} \sim N(\mu, \tfrac{\sigma}{\sqrt{n}}).$$

make a Gaussian R.V. standard.

$$= \mathbb{P}_\mu\Big(\underbrace{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}_{N(0,1)} > \underbrace{\frac{c - \mu}{\sigma/\sqrt{n}}}_{\text{number.}}\Big) \qquad \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$\boxed{\beta(\mu) = 1 - \Phi\left(\frac{c - \mu}{\sigma/\sqrt{n}}\right)}$$

CDF

$$\Phi(x) = \mathbb{P}(Z < x)$$

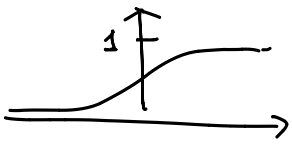
Z : std normal $N(0,1)$

$$\bar{X} > c \iff \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{c - \mu}{\sigma/\sqrt{n}}.$$

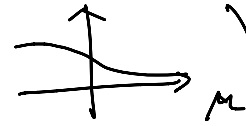
$$\mathbb{P}(Z > c_1) = 1 - \mathbb{P}(Z < c_1) = 1 - \Phi(c_1)$$

$Z \sim N(0,1)$

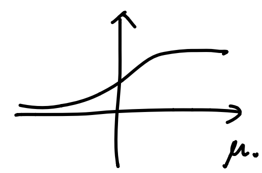
$$\text{size} = \sup_{\mu \in \Theta_0} \beta(\mu) = \sup_{\mu \le 0} \beta(\mu) = \beta(0) = 1 - \Phi\left(\frac{c}{\sigma/\sqrt{n}}\right)$$

($\Phi(x)$: 

$\checkmark$

, $\Phi\left(\frac{c-\mu}{\sigma/\sqrt{n}}\right)$: 

$\checkmark$

$1 - \Phi\left(\frac{c-\mu}{\sigma/\sqrt{n}}\right)$: 

$\checkmark$

$\beta(\mu)$ is $\nearrow$ function.)

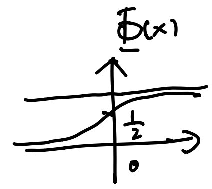
For a size α test, we want.

$$1 - \Phi\left(\frac{c}{\sigma/\sqrt{n}}\right) = \alpha$$

$$1 - \alpha = \Phi\left(\frac{c}{\sigma/\sqrt{n}}\right)$$

$$\Phi^{-1}(1-\alpha) = \frac{c}{\sigma/\sqrt{n}}$$

$$c = \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}(1-\alpha).$$

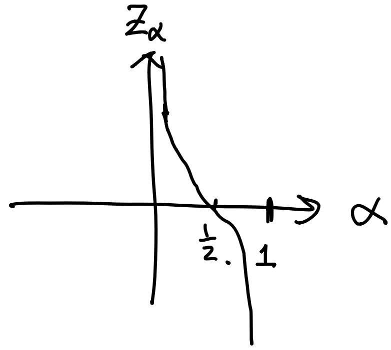


say. $\alpha = 0.5$, then $\Phi^{-1}(1-\frac{1}{2}) = \Phi^{-1}(\frac{1}{2}) = 0$

then $c = 0$.

$$\Phi(x) = \int_{-\infty}^{x} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dx$$

$$Z_\alpha = \Phi^{-1}(1-\alpha)$$



$$c = \frac{\sigma}{\sqrt{n}} \cdot Z_\alpha$$

to have smaller α, we need to have bigger c.

Remark : this is an example where we can adjust the test R_α depending on the desired size of the test (α-value).

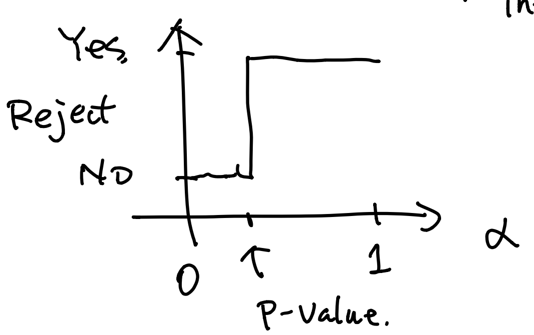
Definition (p-value) : (assume we have a rejection region R_α dep. on α).

Given the data X, what is the smallest α that we reject H_0.

$$\text{p-value} = \underbrace{\inf}_{\uparrow \text{ infimum.}} \{\alpha : X \in R_\alpha\}.$$

inf = min
sup = max

e.g. $\inf\left(\left(\frac{1}{2}, \frac{2}{3}\right)\right) = \frac{1}{2}$



- p-value < 0.01 : strong evidence against H_0.

  p-value > 0.1 : little or no evidence against H_0

warning : large p-value means either H_0 is true ; or the test is too week.

Reading Further :

- Wassermann : All of statistics. (prob + stat).
- Dobrow : intro to stochastic process in R. (lots of examples)
- (math) Rick Durrett : Introduction to probability.

Final :

- vector space, linear alg. 30%.
- differential eq. special fcn. 50%.
- probability (only things in Boas). 20%

- zoom channel: math 121B.