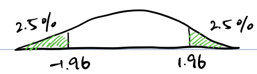
Lecture 03

Large Sample Testing

$H_0: \theta = \theta_0 \qquad H_1: \theta \neq \theta_0$

$$Z = \frac{\hat{\theta} - \theta_0}{\sqrt{Var(\hat{\theta})}} \sim N(0, 1)$$

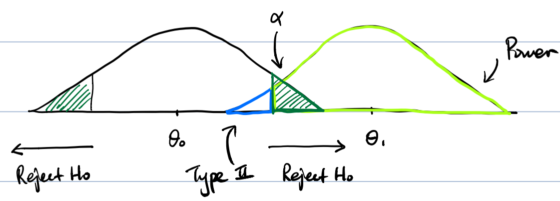


	H_0 true	H_1 true
Fail to reject		Type II Error $= \beta$
Reject H_0	Type I Error $= \alpha$	Power

✓ $\alpha = Pr(\text{Type I}) = Pr(\text{Reject } H_0 \mid H_0 \text{ True})$

$\beta = Pr(\text{Type II}) = Pr(\text{Fail to reject } H_0 \mid H_1 \text{ True})$

$\text{Power} = 1 - \beta$



✓ $\alpha \uparrow \Rightarrow \text{Power} \uparrow$

$\Rightarrow \beta \downarrow$

Confidence Interval $\quad \hat{\theta} \sim N(\theta, \sigma^2) \qquad \sigma^2 = \frac{\Phi^2}{N}$ ← a variance component

the large sample variance can be expressed as a function of the sample size N

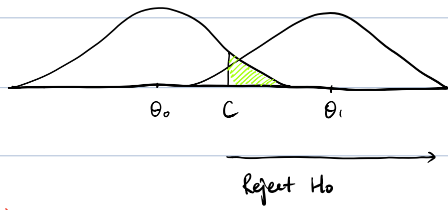
$\hat{\theta} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\hat{Var}(\hat{\theta})} = 100 \times (1-\alpha)\%$ CI

$\hat{\theta} \pm MOE$ Margin of Error $\quad MOE = Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\Phi^2}{N}} \qquad N = \left(\frac{Z_{1-\alpha/2}\,\Phi}{MOE}\right)$

$H_0: \theta = \theta_0$

$H_1: \theta > \theta_0$

Right-tailed



Under H_0 $\hat{\theta} \sim N(\theta_0, \sigma_0^2)$

Under H_1 $\hat{\theta} \sim N(\theta_1, \sigma_1^2)$

$$\alpha = Pr(\text{Type I error}) = Pr(\underbrace{\hat{\theta} > c}_{\text{find } c} \mid H_0 \text{ true})$$

$$= Pr\left(\frac{\hat{\theta} - \theta_0}{\sigma_0} > \frac{c - \theta_0}{\sigma_0} \,\middle|\, H_0 \text{ true}\right)$$

$$= 1 - Pr\left(Z \leq \frac{c - \theta_0}{\sigma_0} \,\middle|\, H_0 \text{ true}\right)$$

$$\alpha = 1 - \Phi\left(\frac{c - \theta_0}{\sigma_0}\right)$$

$$\Rightarrow 1 - \alpha = \Phi\left(\frac{c - \theta_0}{\sigma_0}\right) \Rightarrow Z_{1-\alpha} = \frac{c - \theta_0}{\sigma_0} \Rightarrow C = \theta_0 + Z_{1-\alpha}\sigma_0$$

Reject H_0 if $\hat{\theta} > \theta_0 + Z_{1-\alpha}\sigma_0$

$$\beta = Pr(\text{Type II error}) = \underbrace{Pr(\text{Fail to reject } H_0 \mid H_1 \text{ true})}_{\text{use } c \text{ find } Z_{1-\beta}}$$

$$= Pr(\hat{\theta} \leq \theta_0 + Z_{1-\alpha}\sigma_0 \mid H_1 \text{ true})$$

$$= Pr\left(\frac{\hat{\theta} - \theta_1}{\sigma_1} \leq \frac{(\theta_0 + Z_{1-\alpha}\sigma_0) - \theta_1}{\sigma_1} \,\middle|\, H_1 \text{ true}\right)$$

$$= \Phi\left(\frac{\theta_0 + Z_{1-\alpha}\sigma_0 - \theta_1}{\sigma_1}\right)$$

$$\Rightarrow Z_\beta = \frac{\theta_0 + Z_{1-\alpha}\sigma_0 - \theta_1}{\sigma_1} \qquad Z_{1-\beta} = \frac{\theta_1 - \theta_0 - Z_{1-\alpha}\sigma_0}{\sigma_1}$$

$$\Rightarrow |\Delta| = |\theta_1 - \theta_0| = Z_{1-\alpha}\sigma_0 + Z_{1-\beta}\sigma_1$$ Δ = detectable difference

$$|\Delta| = |\theta_1 - \theta_0| = Z_{1-\alpha/2}\sigma_0 + Z_{1-\beta}\sigma_1$$ for two sided test

$H_0: \theta = \theta_0$

$H_1: \theta < \theta_0$

Left-tailed

$\alpha = Pr(\text{Reject } H_0 \mid H_0 \text{ true}) = Pr(\theta < C \mid H_0 \text{ true}) = Pr\left(\frac{\theta - \theta_0}{\sigma_0} < \frac{c - \theta_0}{\sigma_0} \mid H_0 \text{ true}\right)$

$= \Phi\left(\frac{c-\theta_0}{\sigma_0}\right)$

$\Rightarrow Z_\alpha = \frac{c - \theta_0}{\sigma_0} = -Z_{1-\alpha}$

$\Rightarrow C = \theta_0 - Z_{1-\alpha} \cdot \sigma_0$

$\beta = Pr(\text{Fail to reject} \mid H_1 \text{ true}) = Pr(\theta > C \mid H_0 \text{ true}) = Pr\left(\frac{\theta - \theta_1}{\sigma_1} > \frac{c - \theta_1}{\sigma_1} \mid H_1 \text{ true}\right)$

$1 - \beta = \Phi\left(\frac{c - \theta_1}{\sigma_1}\right)$

$\Rightarrow Z_{1-\beta} = \frac{c - \theta_1}{\sigma_1}$

$\Rightarrow Z_{1-\beta} = \frac{\theta_0 - Z_{1-\alpha} \cdot \sigma_0 - \theta_1}{\sigma_1}$

$|\Delta| = Z_{1-\alpha} \sigma_0 + Z_{1-\beta} \cdot \sigma_1$

$$\sigma_0^2 = Var(\hat{\theta} \mid H_0 \text{ true}) = \sqrt{\frac{\Phi_0^2}{N}}$$

$$\sigma_1^2 = Var(\hat{\theta} \mid H_1 \text{ true}) = \sqrt{\frac{\Phi_1^2}{N}}$$

$$|\Delta| = Z_{1-\alpha}\frac{\Phi_0}{\sqrt{N}} + Z_{1-\beta}\frac{\Phi_1}{\sqrt{N}} = \frac{Z_{1-\alpha}\Phi_0 + Z_{1-\beta}\Phi_1}{\sqrt{N}}$$

$$N = \left(\frac{Z_{1-\alpha}\cdot\Phi_0 + Z_{1-\beta}\cdot\Phi_1}{|\Delta|}\right)^2$$

$|\Delta| \uparrow \Rightarrow N \downarrow$

power $\uparrow \Rightarrow N \uparrow$ sample size

$$Z_{1-\beta} = \frac{\sqrt{N}|\Delta| - Z_{1-\alpha}\Phi_0}{\Phi_1}$$

$$Pr(Z \le Z_{1-\beta}) = 1-\beta$$

$\frac{\Phi}{|\Delta|}$ = noncentrality parameter

Binomial Distribution

$H_0: \pi_1 = \pi_2$

Sample fraction $\xi_1 = \frac{n_1}{N}$ $\xi_2 = \frac{n_2}{N}$ $\xi_1 + \xi_2 = 1$

$H_1: \pi_1 \neq \pi_2$

$\xi_1 = \xi_2 = \frac{1}{2}$

$$\sigma_0^2 = \frac{\pi(1-\pi)}{n_1} + \frac{\pi(1-\pi)}{n_2} = \frac{\pi(1-\pi)}{N\cdot\xi_1} + \frac{\pi(1-\pi)}{N\cdot\xi_2} = \frac{\Phi_0^2}{N}$$

Balanced Design

$$\sigma_1^2 = \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2} = \frac{1}{N}\left[\frac{\pi_1(1-\pi_1)}{\xi_1} + \frac{\pi_2(1-\pi_2)}{\xi_2}\right] = \frac{\Phi_1^2}{N}$$

$\Rightarrow$ More Power

$$\Rightarrow \Phi_0 = \sqrt{\frac{\pi(1-\pi)}{\xi_1} + \frac{\pi(1-\pi)}{\xi_2}}$$

under $\xi_1 = \xi_2 = \frac{1}{2}$

$$\Phi_0 = \sqrt{4\pi(1-\pi)}$$

$$\Phi_1 = \sqrt{\frac{\pi_1(1-\pi_1)}{\xi_1} + \frac{\pi_2(1-\pi_2)}{\xi_2}}$$

$$\Phi_1 = \sqrt{2\pi_1(1-\pi_1) + 2\pi_2(1-\pi_2)}$$

$$N = \left(\frac{z_{1-\alpha}\Phi_0 + z_{1-\beta}\Phi_1}{\underset{|\pi_1 - \pi_2|}{\underset{\parallel}{|\Delta|}}}\right)^2$$