Commonly used Test

Test for One population.

Review: ① Hypothesis H_0, H_a

② Test Statistic GLRT $\left\{ \dfrac{\max_{\theta \in \Theta_0 \cup \Theta_a} L(\theta ; x)}{\max_{\theta \in \Theta_0} L(\theta ; x)} > k \right\}$

③ Null Distribution

Under H_0, the distribution of Test Statistic

$$2 \log \Lambda \sim \chi^2_r$$

④ P-value Rejection Region

⑤ Decision

Testing for the population mean / proportions

① population mean

H_0: $\mu = \mu_0$ $\quad H_0: \mu \geq \mu_0 \quad H_0: \mu \leq \mu_0$

H_a, $\mu \neq \mu_0$ $\quad H_a: \mu < \mu_0 \quad H_a: \mu > \mu_0$

→ Test Statistic: $T = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$

Under H_0, and normality. $T \sim T_{n-1}$

P-value: $P(|T_{n-1}| > |T|)$ $\quad P(T_{n-1} < T)$ $\quad P(T_{n-1} > T)$

Oil change: $H_0: \mu \geq 30$

$H_a: \mu < 30$

Type I $= \sup_{\Theta_0} P(R \mid \theta \in \Theta_0)$

$n = 36$.

$$T = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{\bar{x} - 30}{s/\sqrt{36}} = -3.59$$

$$\text{P-value} = P(T_{35} < -3.59) = 0.0005$$

We reject the H_0.

Proportion: Discrete

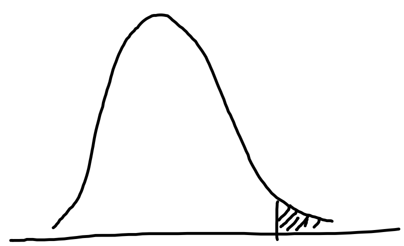
100 patients

U be the number of patients improved

$U \sim Bin(100, p)$ $\quad \hat{p} = \frac{Z}{n}$ $\quad \frac{np(1-p)}{n^2}$

$H_0: p \leq 0.5$

$H_a: p > 0.5$



Z_α

Test Statistic $Z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$

Under the null $Z \approx N(0,1)$ $\quad R = \{Z > Z_\alpha\}$

$$\text{P-value} := P(N(0,1) > z) = 1 - P(N(0,1) \leq Z)$$

$n = 100$

the number of success 55 $\quad \hat{p} = \frac{55}{100} = 0.55$

~~the~~ number of success 55 $\quad p = \frac{55}{100} = \underline{0.55}$

$$Z = \frac{0.55 - 0.50}{\sqrt{\frac{0.50 \times 0.50}{100}}} = \frac{0.05}{0.05} = \underline{1}$$

$$\text{p-value} = P(N(0,1) > 1) = 1 - \underline{\underline{P(N(0,1) \le 1)}} = 0.16$$

Fail-to-reject H_0. There is no sufficient evidence showing that H_0 is false.

Sample size calculation.

Type I $\le \alpha$ $\quad$ power $\ge 1-\beta$ for given effect size

Sample size

$\underline{H_0}: \quad \mu = \mu_0$

$H_a: \quad \mu \ne \mu_0 \qquad \mu = \mu_0 + \delta$

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \qquad R = \left\{ \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| > Z_{\frac{\alpha}{2}} \right\}$$

$$1-\beta = P\left(R \,\middle|\, \underline{\mu = \mu_0 + \delta}\right) = P\left(\left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| > Z_{\frac{\alpha}{2}} \right)$$

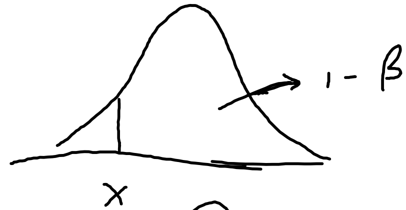
$$= P\left(\left| \frac{\bar{X} - (\mu_0 + \delta) + \delta}{\sigma/\sqrt{n}} \right| > Z_{\frac{\alpha}{2}} \right) = P\left(\left| \frac{\bar{X} - (\mu_0+\delta)}{\sigma/\sqrt{n}} + \frac{\delta}{\sigma/\sqrt{n}} \right| > Z_{\frac{\alpha}{2}} \right)$$

$$= P\left(\left| Z + \frac{\delta}{\sigma/\sqrt{n}} \right| > Z_{\frac{\alpha}{2}} \right)$$

$$= P\left(Z > Z_{\frac{\alpha}{2}} - \frac{\delta}{\sigma/\sqrt{n}} \right) + P\left(Z < -Z_{\frac{\alpha}{2}} - \frac{\delta}{\sigma/\sqrt{n}} \right)$$

$$= P\left(Z > Z_{\frac{\alpha}{2}} - \frac{\delta}{\sigma/\sqrt{n}}\right) + P\left(Z < -Z_{\frac{\alpha}{2}} - \frac{\delta}{\sigma/\sqrt{n}}\right)$$

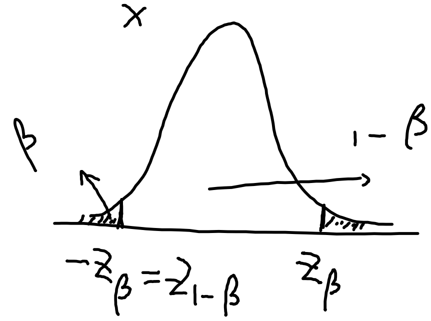
$$1-\beta \approx P\left(Z > Z_{\frac{\alpha}{2}} - \frac{\delta}{\sigma/\sqrt{n}}\right)$$



$$-Z_\beta = Z_{1-\beta} = Z_{\frac{\alpha}{2}} - \frac{\delta}{\sigma/\sqrt{n}}$$

$$\frac{\delta}{\sigma/\sqrt{n}} = Z_{\frac{\alpha}{2}} + Z_\beta$$

$$n = \frac{\sigma^2}{\delta^2}\left(Z_{\frac{\alpha}{2}} + Z_\beta\right)^2$$



Example: 6.4.3

H_0: $\mu = 1500$ ml

H_a: $\mu \neq 1500$ ml $\qquad \sigma^2 = 10{,}000$ ml^2

Set $\alpha = 0.05$

(a) Detect a difference of 50 ml. sample size = 10

What is the power.

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \qquad R = \left\{ \left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| > Z_{\frac{\alpha}{2}} \right\} \quad 1.96$$

$$\text{Power} = P\left(\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| > 1.96 \;\middle|\; \mu = 1550 \right)$$

$$= P\left(\left|\frac{\bar{X} - 1550 + 50}{\sigma/\sqrt{n}}\right| > 1.96\right) = P\left(\left| Z + \frac{50}{\sigma/\sqrt{n}} \right| > 1.96\right)$$

$$= P\left(\left| Z + \frac{50}{100/\sqrt{10}} \right| > 1.96 \right) = 0.35$$

(b) want $1-\beta = 80\%$. $\beta = 0.20$

$\alpha = 0.05 \quad \beta = 0.20 \quad \delta = 50 \quad \sigma = 100$ ml

$$n = \frac{\sigma^2}{\delta^2}\left(Z_{\frac{\alpha}{2}} + Z_\beta\right)^2 = 31.4$$

$n \simeq 32$

Midterm exam: 23rd 21-23

2:45 - 4:45 pm

20

Two Populations:

populations μ_1 μ_2

H_0. $\mu_1 - \mu_2 = \delta_0$

H_a: $\mu_1 - \mu_2 \neq \delta_0$

Independent Samples If $\sigma_1^2 = \sigma_2^2$

$$T_{pooled} = \frac{\bar{y}_1 - \bar{y}_2 - \delta_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}$$

Under the null, $T_{pooled} \sim T_{n_1+n_2-2}$

If $\sigma_1^2 \neq \sigma_2^2$, $V(\bar{y}_1 - \bar{y}_2) = V(y_1) + V(y_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

$$T = \frac{\bar{y}_1 - \bar{y}_2 - \delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Under the null, $T \sim T_{sat}$

Conservative d.f. $\min(n_1 - 1, n_2 - 1)$

Satterthwaite Approximation (Welch)

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

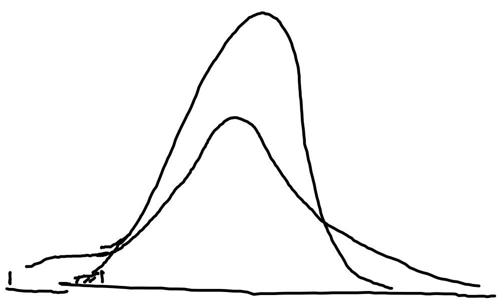
$$\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}$$

H_0: $\frac{\sigma_1^2}{\sigma_2^2} = 1$

H_a: $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$

Test Statistic. $F = \frac{S_1^2}{S_2^2}$

Under the null, $F \sim F_{n_1-1,\, n_2-1}$



Example 6.4.4. $n = 102$ men

$n_1 = 50$ control

$n_2 = 52$ patients

$P = 6033$ genes

$i = 1, 2, \cdots n_j$

$j = 1, 2$

$g = 1, 2, \cdots 6033$

$$y_{ijg} \sim N(\mu_{jg}, \sigma_{jg}^2)$$

$y_{ijg} = \mu_{jg} + \epsilon_{jg}$ where $\epsilon_{jg} \sim N(0, \sigma_{jg})$

H_0: $\mu_{1g} = \mu_{2g}$

H_a: $\mu_{1g} \neq \mu_{2g}$

} Multiple Testing

Focus on one gene. 1720

H_0: $\mu_{1,1720} = \mu_{2,1720}$

H_a: $\mu_{1,1720} \neq \mu_{2,1720}$

Two population. Matched Pairs

y_{ij} be the number of errors i-th sentence

$j - \begin{cases} 1 \text{ Native English} \\ 2. \text{ Native Greek} \end{cases}$

y_{i1} and y_{i2} are independent?

H_0: $\mu_1 - \mu_2 = \delta_0$

H_a. $\mu_1 - \mu_2 \neq \delta_0$

$d_i = y_{i1} - y_{i2}$ $\qquad \bar{d} \qquad s_d$

$$T = \frac{\bar{d}_i - \delta_0}{s_d / \sqrt{n}} \qquad n = 32$$

H_0: $T \sim T_{n-1}$