

T.H. para a diferença de médias, $\mu_1 - \mu_2$

Hipóteses:

$$\begin{array}{c|c|c} H_0: \mu_1 - \mu_2 = k & H_0: \mu_1 - \mu_2 = k & H_0: \mu_1 - \mu_2 = k \\ H_1: \mu_1 - \mu_2 < k & H_1: \mu_1 - \mu_2 > k & H_1: \mu_1 - \mu_2 \neq k \end{array}$$

Estatística de teste: $\bar{X}_1 - \bar{X}_2$

σ_1 e σ_2 conhecidos:

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

i) Populações normais:

$$\left(X_1 \sim N(\mu_1, \sigma_1^2) \quad \wedge \quad X_2 \sim N(\mu_2, \sigma_2^2) \right) \xRightarrow{TAN} Z \sim N(0,1)$$

ii) Grandes amostras e populações quaisquer:

$$(n_1 \geq 30 \quad \wedge \quad n_2 \geq 30) \xRightarrow{TLC, TAN} Z \sim N(0,1)$$

T.H. para a diferença de médias, $\mu_1 - \mu_2$

σ_1 e σ_2 desconhecidos ($\sigma_1 = \sigma_2$):

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad , \quad S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

i) Populações normais:

$$\left(X_1 \sim N(\mu_1, \sigma_1^2) \quad \wedge \quad X_2 \sim N(\mu_2, \sigma_2^2) \right) \Rightarrow T \sim T(n_1 + n_2 - 2)$$

ii) Grandes amostras e populações quaisquer:

$$(n_1 \geq 30 \quad \wedge \quad n_2 \geq 30) \Rightarrow T \sim N(0,1)$$

T.H. para a diferença de médias, $\mu_1 - \mu_2$

σ_1 e σ_2 desconhecidos ($\sigma_1 \neq \sigma_2$):

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

i) Populações normais:

$$(X_1 \sim N(\mu_1, \sigma_1^2) \wedge X_2 \sim N(\mu_2, \sigma_2^2)) \Rightarrow T \sim T(m), \quad m \cong \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}}$$

ii) Grandes amostras e populações quaisquer:

$$(n_1 \geq 30 \quad \wedge \quad n_2 \geq 30) \Rightarrow T \sim N(0,1)$$

T.H. para a diferença de médias, $\mu_1 - \mu_2$

Região crítica (RC):

H_0 verdadeira $\Rightarrow \mu_1 - \mu_2 = k$

$$\begin{array}{l|l|l} P(\bar{X}_1 - \bar{X}_2 < c | \mu_1 - \mu_2 = k) = \alpha & P(\bar{X}_1 - \bar{X}_2 > c | \mu_1 - \mu_2 = k) = \alpha & P(\bar{X}_1 - \bar{X}_2 < c_1 | \mu_1 - \mu_2 = k) = \frac{\alpha}{2} \\ \text{RC} = \{\bar{X}_1 - \bar{X}_2 : \bar{X}_1 - \bar{X}_2 < c\} & \text{RC} = \{\bar{X}_1 - \bar{X}_2 : \bar{X}_1 - \bar{X}_2 > c\} & P(\bar{X}_1 - \bar{X}_2 > c_2 | \mu_1 - \mu_2 = k) = \frac{\alpha}{2} \\ & & \text{RC} = \{\bar{X}_1 - \bar{X}_2 : \bar{X}_1 - \bar{X}_2 < c_1 \\ & & \quad \vee \bar{X}_1 - \bar{X}_2 > c_2\} \end{array}$$

Decisão no método da região crítica:

$$\bar{x}_1 - \bar{x}_2 \in \text{RC} \Rightarrow \text{Rejeitar } H_0$$

T.H. para a diferença de médias, $\mu_1 - \mu_2$

Valor p :

$$P(\bar{X}_1 - \bar{X}_2 \leq \bar{x}_1 - \bar{x}_2 | \mu_1 - \mu_2 = k) \quad \bigg| \quad P(\bar{X}_1 - \bar{X}_2 \geq \bar{x}_1 - \bar{x}_2 | \mu_1 - \mu_2 = k) \quad \bigg| \quad 2 \min\{P(\bar{X}_1 - \bar{X}_2 \leq \bar{x}_1 - \bar{x}_2 | \mu_1 - \mu_2 = k), \\ P(\bar{X}_1 - \bar{X}_2 \geq \bar{x}_1 - \bar{x}_2 | \mu_1 - \mu_2 = k)\}$$

Decisão no método do valor de prova:

$$\text{Valor } p < \alpha \quad \Rightarrow \quad \text{Rejeitar } H_0$$

Exercício

Uma turma de 40 alunos da escola 1 e outra de 50 alunos da escola 2 foram submetidas a uma mesma prova. Na 1ª turma, a nota média foi 74 e o desvio padrão 8. Na 2ª turma, a nota média foi 78 e o desvio padrão 7. Serão melhores os resultados da 2ª escola?

Hipóteses:

$$H_0 : \mu_1 = \mu_2 \Leftrightarrow \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 < \mu_2 \Leftrightarrow \mu_1 - \mu_2 < 0$$

Estatística de teste: $\bar{X}_1 - \bar{X}_2$

$\bar{X}_1$: "Nota média por aluno numa amostra de 40 alunos da escola 1"

$\bar{X}_2$: "Nota média por aluno numa amostra de 50 alunos da escola 2"

$$\begin{aligned} \begin{cases} n_1 = 40 \\ n_2 = 50 \end{cases} &\Rightarrow \bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) \wedge T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1) \\ \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} &= \sqrt{\frac{8^2}{40} + \frac{7^2}{50}} \end{aligned}$$

Exercício

Região crítica (RC):

$$H_0 \text{ verdadeira} \Rightarrow \mu_1 - \mu_2 = 0$$

$$\alpha = P(\text{Rejeitar } H_0 \mid H_0 \text{ verdadeira}) \quad ; \quad \alpha = 0.05 \quad (\text{Por defeito})$$

$$P(\bar{X}_1 - \bar{X}_2 < c \mid \mu_1 - \mu_2 = 0) = \alpha = 0.05 \Leftrightarrow \Phi\left(\frac{c - 0}{\sqrt{\frac{8^2}{40} + \frac{7^2}{50}}}\right) = 0.05$$
$$\Leftrightarrow \frac{c}{\sqrt{\frac{8^2}{40} + \frac{7^2}{50}}} = -1.65 \Leftrightarrow c = -2.65$$

$$RC = \{\bar{X}_1 - \bar{X}_2 : \bar{X}_1 - \bar{X}_2 < -2.65\}$$

Decisão:

$$\bar{x}_1 - \bar{x}_2 = 74 - 78 = -4 \in RC$$

Conclusão: H_0 deve ser rejeitada ao nível de 5%, ou seja, pode concluir-se que a 2ª escola tem melhores resultados.

Exercício

Nota:

$$\text{Valor } p = P(\bar{X}_1 - \bar{X}_2 \leq -4 \mid \mu_1 - \mu_2 = 0) = \Phi\left(\frac{-4 - 0}{\sqrt{\frac{8^2}{40} + \frac{7^2}{50}}}\right) = \Phi(-2.49) = 0.0064 < \alpha$$

T.H. para a diferença de proporções, $p_1 - p_2$

Hipóteses:

$$\begin{array}{c|c|c} \begin{array}{l} H_0: p_1 - p_2 = k \\ H_1: p_1 - p_2 < k \end{array} & \begin{array}{l} H_0: p_1 - p_2 = k \\ H_1: p_1 - p_2 > k \end{array} & \begin{array}{l} H_0: p_1 - p_2 = k \\ H_1: p_1 - p_2 \neq k \end{array} \end{array}$$

Estatística de teste: $\hat{P}_1 - \hat{P}_2$

$$\begin{cases} n_1 \geq 30 \\ n_2 \geq 30 \end{cases} \xrightarrow{TLC, TAN} \hat{P}_1 - \hat{P}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$$

$$\begin{array}{l} p_1 \neq p_2 \\ (k \neq 0) \end{array} \Rightarrow Z = \frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{P}_1(1 - \hat{P}_1)}{n_1} + \frac{\hat{P}_2(1 - \hat{P}_2)}{n_2}}} \sim N(0,1)$$

$$\begin{array}{l} p_1 = p_2 \\ (k = 0) \end{array} \Rightarrow Z = \frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\hat{P}(1 - \hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1) , \quad \hat{P} = \frac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2}$$

T.H. para a diferença de proporções, $p_1 - p_2$

Região crítica (RC):

$$H_0 \text{ verdadeira} \Rightarrow p_1 - p_2 = k$$

$$\begin{array}{l|l|l} P(\hat{P}_1 - \hat{P}_2 < c | p_1 - p_2 = k) = \alpha & P(\hat{P}_1 - \hat{P}_2 > c | p_1 - p_2 = k) = \alpha & P(\hat{P}_1 - \hat{P}_2 < c_1 | p_1 - p_2 = k) = \frac{\alpha}{2} \\ \text{RC} = \{\hat{P}_1 - \hat{P}_2 : \hat{P}_1 - \hat{P}_2 < c\} & \text{RC} = \{\hat{P}_1 - \hat{P}_2 : \hat{P}_1 - \hat{P}_2 > c\} & P(\hat{P}_1 - \hat{P}_2 > c_2 | p_1 - p_2 = k) = \frac{\alpha}{2} \\ & & \text{RC} = \{\hat{P}_1 - \hat{P}_2 : \hat{P}_1 - \hat{P}_2 < c_1 \\ & & \quad \vee \hat{P}_1 - \hat{P}_2 > c_2\} \end{array}$$

Decisão no método da região crítica:

$$\hat{p}_1 - \hat{p}_2 \in \text{RC} \Rightarrow \text{Rejeitar } H_0$$

T.H. para a diferença de proporções, $p_1 - p_2$

Valor p :

$$P(\hat{P}_1 - \hat{P}_2 \leq \hat{p}_1 - \hat{p}_2 | p_1 - p_2 = k) \quad \left| \quad P(\hat{P}_1 - \hat{P}_2 \geq \hat{p}_1 - \hat{p}_2 | p_1 - p_2 = k) \quad \left| \quad 2 \min \left\{ P(\hat{P}_1 - \hat{P}_2 \leq \hat{p}_1 - \hat{p}_2 | p_1 - p_2 = k), \right. \right. \\ \left. \left. P(\hat{P}_1 - \hat{P}_2 \geq \hat{p}_1 - \hat{p}_2 | p_1 - p_2 = k) \right\} \right.$$

Decisão no método do valor de prova:

$$\text{Valor } p < \alpha \quad \Rightarrow \quad \text{Rejeitar } H_0$$