# Inferência estatística para uma e duas amostras

# Teste de hióteses

$$Z = \frac{\bar{X} - \mu}{\sigma}$$

# Teste de hióteses

| | | Verdade sobre a população | |
|---|---|---|---|
| | | Ho (verdadeira) | Ha (verdadeira) |
| Deciçõa basejada na amostra | Aceita Ho (Rejeita Ha) | ✓ Decisão correta ($1-\alpha$) | Error tipo II ($\beta$) |
| | Rejeita Ho (Aceita Ha) | Error de tipo I ($\alpha$) | ✓ Decisão correta ($1-\beta$) |

ste um equilíbrio entre esses dois tipos de erros, no sentido de que ao tentar-se minimizar α, aumenta-s

go, não e´ possíıvel minimizar estas duas probabilidades simultaneamente.

eitar Ho é considerado o erro mais sério, do que  erroneamente aceitar.

# Procedimento para se efetuar um teste de hipóteses

1. Enunciar as hipóteses $H_o$ e $H_1$;
2. Fixar o limite de erro $\alpha$ e identificar a variável do teste;
3. Determinar as áreas de aceitação (RA) e rejeição (RR) em função do nível ⍰ pelas tabelas estatísticas;
4. Por meio dos elementos amostrais avaliar o valor da variável do teste;
5. Concluir pela aceitação ou rejeição.
6. Elaborar uma conclusão em relação ao problema que está sendo testado.

# Teste de hióteses

- Graus de liberdade: Referem-se a liberdade de variação de um conjunto de escores, por exemplo: uma amostra com 6 elementos, 5 podem varia e 1 fica fixo, Logo os graus de liberdade podem ser representados por gl (N –1).

# Inferência estatística para uma amostra

| | Para n > 30 | Para n ≤ 30 |
|---|---|---|
| cia<br>acional<br>cida | • $Z_c = \frac{\bar{X}-\mu_o}{\frac{\sigma}{\sqrt{n}}}$<br>• $IC_{(\mu,1-\alpha)} = \left(\bar{X} - Z^*.\frac{\sigma}{\sqrt{n}} \text{ and } \bar{X} + Z^*.\frac{\sigma}{\sqrt{n}}\right)$ | • $T_c = \frac{\bar{X}-\mu_o}{\frac{\sigma}{\sqrt{n}}}; \quad T^*_{(n-1),1-\frac{\alpha}{2}}$<br>• $IC_{(\mu,1-\alpha)} = \left(\bar{X} - T^*.\frac{\sigma}{\sqrt{n}} \text{ and } \bar{X} + T^*.\right.$ |
| cia<br>acional<br>nhecida | • $Z_c = \frac{\bar{X}-\mu_o}{\frac{s}{\sqrt{n}}}$<br>• $IC_{(\mu,1-\alpha)} = \left(\bar{X} - Z^*.\frac{s}{\sqrt{n}} \text{ and } \bar{X} + Z^*.\frac{s}{\sqrt{n}}\right)$ | • $T_c = \frac{\bar{X}-\mu_o}{\frac{s}{\sqrt{n}}}$<br>• $IC_{(\mu,1-\alpha)} = \left(\bar{X} - T^*.\frac{s}{\sqrt{n}} \text{ and } \bar{X} + T^*\right.$ |
| rção | $\mu_o = np_o; \quad \sigma = \sqrt{np_o(1-p_o)}$<br>• $Z_c = \frac{X-np_o}{\sqrt{np_o(1-p_o)}} = \frac{\frac{X}{n}-p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} = \frac{\hat{P}-p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}}$<br>• $IC_{(\mu,1-\alpha)} = \left(\hat{P} - Z^*.\sqrt{\frac{p_o(1-p_o)}{n}} \text{ and } \hat{P} + Z^*.\sqrt{\frac{p_o(1-p_o)}{n}}\right)$ | |

# Teste de hióteses Z e t

- Uma amostra aleatória de 36 copos de um certo vinho mostrou que tinha um conteúdo médio líquido de 223.5 ml, com desvio padrão de 3,6 ml. Testar a hipótese de que $\mu = 225$ ml contra a alternativa $\mu < 225$ ml, com o nível de significância de $\alpha = 0{,}05$.
- Uma amostra aleatória de 8 copos de um certo vinho mostrou que tinha um conteúdo médio líquido de 223.5 ml, com desvio padrão de 3,6 ml. Testar a hipótese de que $\mu = 225$ ml contra a alternativa $\mu < 225$ ml, com o nível de significância de $\alpha = 0{,}05$.

# Exemplo (proporção)

1. Um fabricante afirma que no máximo 10% dos seus produtos são defeituosos. Um órgão de defesa do consumidor testa uma amostra de 81 desses itens, detectando 12 de defeituosos. Testar se o fabricante esta certo.
2. O nível de aprovação da qualidade das refeições servidas em um restaurante universitário era de 20%, quando houve uma movimentação geral dos estudantes que forçou a direção do restaurante a fazer mudanças. Feitas as mudanças, sorteia-se uma amostra de 64 estudantes usuários do restaurante e 25 aprovam a qualidade da comida. Você diria, ao nível de significância de 5%, que as mudanças surtiram efeito?
3. O consumidor exige que a fração de defeituosos de certo produto não exceda 0,05. O nível de significância $\alpha$ exigido é de $\alpha = 0, 05$. Uma amostra de tamanho 200 é observada. Dentre os 200, 4 são defeituosos (2%). Podemos concluir que a exigência do consumidor é satisfeita?

# Teste de hióteses

- Graus de liberdade: Referem-se a liberdade de variação de um conjunto de escores, por exemplo: uma amostra com 6 elementos, 5 podem variar e 1 fica fixo, Logo os graus de liberdade podem ser representados por gl = (N –1).

# 1. Teste para a diferença entre duas proporções populacionais p1 e p2

- $\mu_o = np_o;$ $\sigma_d = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

- $Z_c = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$

- $IC_{(\mu,1-\alpha)} = \left( \begin{array}{c} (p_1 - p_2) - Z^* . \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad and \\ (p_1 - p_2) + Z^* . \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \end{array} \right)$

# 1. Teste para a diferença entre duas proporções populacionais p1 e p2

p://www.lampada.uerj.br/arquivosdb/_book/compara%C3%A7%C3%A3o-de-propor%C3%A7%C3%B5es.html

- $\mu_o = np_o;$ $\sigma_d = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

- $Z_c = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_o)}{n}}}$

- $IC_{(\mu, 1-\alpha)} = \left( \begin{array}{c} (p_1 - p_2) - Z^* . \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad and \\ (p_1 - p_2) + Z^* . \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \end{array} \right)$

- Uma empresa produz ceral matinal e deseja saber se as duas fabricas apresentam diferença significativa entre a porcentagem de productos aceitaveis pelo controle de qualidade.  Para isso foi selecionado aleatoriamente 1000 prudutos da fabrica A encontrandose que em 12 deles houve algum tipo de defeito. Já a fabrica na B foram selecionados 1200 serviços e 20 receberam algum tipo de rejeição.

- Numa pesquisa de opinião, 64 dentre 298 homens declararam apreciar certo produto, acontecendo o mesmo com 28 dentre 253 mulheres. Ao nível de 5% de significância, os homens e as mulheres apreciam igualmente a revista?

# 2. Teste para a diferença de duas médias (n≥30)

a) $\sigma^2$ conhecida; grupos pareados; n1 ≠ n2

$$H_o: \mu_1 - \mu_2 = d \quad ou \quad \mu_1 - \mu_2 = 0 \quad ou \quad \mu_1 = \mu_2$$
$$H_1: \mu_1 - \mu_2 \neq d \quad ou \quad \mu_1 - \mu_2 \neq 0 \quad ou \quad \mu_1 = \mu_2$$

$$(\overline{X_1} - \overline{X_2}) \cong N\left(\mu_1 - \mu_2; \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right); \qquad \sigma_d = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$Z_c = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sigma_d}$$

$$IC_{(\mu,1-\alpha)} = \left( (\overline{X_1} - \overline{X_2}) \pm Z^*{}_{\left(1-\frac{\alpha}{2}\right)} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$$

# 3. Teste para a diferença de duas médias (n < 30)

a) $\sigma_1^2$ = $\sigma_2^2$ desconhecida e semelhantes; grupos não pareados; n1 = n2

$\mu_1 - \mu_2 = d$ ou $\mu_1 - \mu_2 = 0$ ou $\mu_1 = \mu_2$

$\mu_1 - \mu_2 \neq d$ ou $\mu_1 - \mu_2 \neq 0$ ou $\mu_1 = \mu_2$

$$(\overline{X_1} - \overline{X_2}) \cong N\left(\mu_1 - \mu_2; \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$T_c = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$IC_{(\mu,1-\alpha)} = \left((\overline{X_1} - \overline{X_2}) \pm T^*_{\left(n_1+n_2-2,1-\frac{\alpha}{2}\right)} \cdot \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$$

# 4. Teste para a diferença de duas médias (n < 30)

a) $\sigma_1^2 \neq \sigma_2^2$ desconhecida; grupos não pareados; n1 ≠ n2

$\mu_1 - \mu_2 = d$ ou $\mu_1 - \mu_2 = 0$ ou $\mu_1 = \mu_2$

$\mu_1 - \mu_2 \neq d$ ou $\mu_1 - \mu_2 \neq 0$ ou $\mu_1 = \mu_2$

$- \overline{X_2}) \cong N\left(\mu_1 - \mu_2; \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right);$

$$graus\ de\ liberdade\ \ \nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

$$T_c = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$IC_{(\mu, 1-\alpha)} = \left( (\overline{X_1} - \overline{X_2}) \pm T^*_{\left(\nu, 1 - \frac{\alpha}{2}\right)} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right)$$

# 5. Teste para a diferença de duas médias (n < 30)

a) $\sigma_1^2 = \sigma_2^2$ desconhecida; grupos Pareadas; n1 = n2

$\mu_1 - \mu_2 = d$ ou $\mu_1 - \mu_2 = 0$ ou $\mu_1 = \mu_2$

$\mu_1 - \mu_2 \neq d$ ou $\mu_1 - \mu_2 \neq 0$ ou $\mu_1 = \mu_2$

$- \overline{X_2}) \cong N\left(\mu_1 - \mu_2; \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right);$ $d = (\overline{X_1} - \overline{X_2});$ $s_d = \sqrt{\frac{1}{n_1 - 1} \cdot \sum_{i=1}^{n} (d_i - \overline{D})^2}$

$c$

$$IC_{(\mu, 1-\alpha)} = \left( (\overline{X_1} - \overline{X_2}) \pm T^*_{\left((n-1), 1-\frac{\alpha}{2}\right)} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right)$$

# 6. Teste de Normalidade

1. Teste de Shapiro-Wilk
2. Teste Anderson Darling,
3. Kolgomorof smitrrow, etc
4. QQ-Plot

H0: A amostra provém de uma população Normal $p>0.0$

H1: A amostra não provém de uma população Normal $p<0.0$

# 6. Igualdade (homocedasticidade) de variâncias $\sigma_1^2 = \sigma_2^2$

$$F_c = \frac{\sigma^2_{maior}}{\sigma^2_{menor}}$$

$$F^*_{(GLmaior,GLmenor)}$$

1. Teste F de igualdade de variâncias
2. Teste de levene
3. Teste de Barlet, etc

H0: Variâncias iguais $p>0.025$

H1: Variâncias diferentes $p<0.025$

H0: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$

H1: pelo menos uma $\sigma_i^2$ é diferente

$H0: \mu 1 = \mu 2$ $p>0.025$

$Ha: \mu 1 \neq \mu 2$ $p<0.025$

Alfa = 0.05

Alfa/2 = 0.025

Se acepta Ho si p_value > 0.025

Se rechaza Ho si p_value < 0.025

$H0: d = 0$

$Ha: d \neq 0$