Guia Rápido - Inferência Estatística I

arthursilva@alu.ufc.br
4r7hu3.github.io

O que é Inferência?

A Inferência Estatística busca fazer estimativas, que podem ser intervalares e pontuais, de **parâmetros populacionais** e/ou da distribuição ou forma/lei de um conjunto de dados a partir de uma **amostra**. Um exemplo de uso prático muito importante são as eleições, que buscam inferir — através da coleta de informações de uma amostra de tamanho n — a quantidade X de votos que irão receber candidatos a cargos políticos. Além dessa aplicação, a Inferência trabalha com testes de hipótese (intrinsecamente ligados à estimação), os quais são empregados em praticamente toda a área do conhecimento humano, não limitando-se às áreas de ciências exatas.

Preliminares

Alguns conceitos importantes para a Inferência:

1. *População* - O conjunto universo. O todo;
2. *Amostra* - Qualquer subconjunto da população. Parte da população;
3. *Parâmetro* - Dados que referem-se à população;
4. *Estatística* - Dados que referem-se à amostra;
5. *Amostra Aleatória Simples* - Conjunto de n variáveis aleatórias independentes e identicamente distribuídas (I.I.D);
6. *Espaço Paramétrico* - É o conjunto Θ em cujos os parâmetros variam;
7. *Suporte* - É o conjunto A em cuja a v.a varia;
8. *Estimador* - Função dos valores amostrais para estimar um parâmetro desconhecido[1];
9. *Estimativa* - Valor estimado propriamente dito;
10. *Viés* - Mede a distância absoluta (erro) da média do estimador em relação ao parâmetro estimado[2];
    10.1 Estimador não-viesado: $E[\hat{\theta}] = \theta, \ \forall \theta \in \Theta$
    10.2 Estimador viesado: $E[\hat{\theta}] \neq \theta, \ \forall \theta \in \Theta$
    10.3 Estimador assintoticamente não viesado: $\lim_{n \to \infty} E[\hat{\theta}] = \theta$
    10.4 Estimador consistente: $\lim_{n \to \infty} E[\hat{\theta}] = \theta$ e $\lim_{n \to \infty} Var[\hat{\theta}] = 0$
11. *Erro Quadrático Médio* - O EQM calcula a expectativa do quadrado do erro do estimador em relação ao parâmetro a ser estimado;
12. *Erro Padrão* - É a raíz quadrada da varância de um estimador;

Definição formal de *Estatística*: função da amostra que **não** depende de parâmetros desconhecidos.
Exemplos:
Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n de uma v.a X.

- $max(x_1, x_2, \ldots, x_n)$ é uma estatística;
- $min(x_1, x_2, \ldots, x_n)$ é uma estatística;
- $\bar{X}$ é uma estatística;
- S^2 é uma estatística.

Note que nenhuma das funções acima depende de parâmetros desconhecidos!

Para fixar

Agora, vamos anotar algumas formas e provas muito importantes para a Inferência:

Fórmulas

- $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$

- $Z = \frac{X - \mu}{\sigma} \sim N(0;1)$ e $Z^2 \sim \chi^2_{(1)}$

- $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$

- $\hat{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$

- $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

- $\frac{(n-1)S^2}{n} \sim \chi^2_{(n)}$

- $2\theta S \sim \chi^2_{(2n)}$, com $S \sim Gama(n; \theta)$

- $t = \frac{Z}{\sqrt{\frac{V}{v}}} \sim t_v, \text{ com } V \sim \chi^2_v$

- $Y = -ln(U) \sim Exp(1), \text{onde } U \sim U(0,1)$

- $X = \frac{\frac{U}{n-1}}{\frac{V}{m-1}} \sim F(n-1; m-1)$, com $\frac{U}{n-1} = \frac{S_1^2}{\sigma_1^2}$ e $\frac{V}{m-1} = \frac{S_2^2}{\sigma_2^2}$

- $G = \frac{1}{X} \sim F(m-1; n-1)$

- $\sum_{i=1}^{n} X_i \sim Gama(n{,}\theta)$, com $X \sim Exp(\theta)$

- $\sum_{i=1}^{n} X_i \sim Bin(n;p)$, com $X \sim Ber(p)$

- $B(\hat{\theta}) = E[\hat{\theta}] - \theta$, é o **viés** do estimador $\hat{\theta}$

- $EQM(\theta) = E[(\hat{\theta} - \theta)^2] = Var[\hat{\theta}] + B^2(\hat{\theta})$ é o **erro quadrático médio** do estimador $\hat{\theta}$[3]

- $f(x_1, x_2, \ldots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$ é a **distribuição conjunta da amostra**

- $L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) I_{\Theta}(\theta)$ é a **função de verossimilhança**[4]

- $\int_0^{\infty} u^{a-1}\, e^{-bu^c}\, dv = \dfrac{\Gamma(\frac{a}{c})}{cb^{\frac{a}{c}}}$ é a **função gama generalizada**
  para lembrar:
  1. $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$
  2. $\Gamma(n) = (n-1)!,\ \forall n \in \mathbb{Z}^+$
  3. $\Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi}$

Algumas demonstrações

- $\bar{X}$ é um estimador viesado ou não viesado para μ? Considere uma v.a $X \sim N(\mu; \sigma^2)$.

  Solução: $E[\bar{X}] - \mu = \dfrac{1}{n} E[\sum_{i=1}^{n} X_i] - \mu = \dfrac{1}{n} \times n \times E[X] - \mu = 0$. Portanto, $\bar{X}$ é **não viesado**!
  OBS: Note que X_i são n v.a's I.I.D

- S^2 é um estimador vieasado ou não viesado para σ^2? Considere uma v.a $X \sim N(\mu; \sigma^2)$.

  Solução: $\dfrac{1}{n-1} E[\sum_{i=1}^{n} (X_i - \bar{X})^2]$.

  Temos que $\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} (X_i^2 - 2X_i\bar{X} + \bar{X}^2) = \sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$.

Agora temos que $E[\sum_{i=1}^{n}(X_i - \bar{X})^2] = E[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2]$

Note que $E[X^2] = Var[X] + E^2[X] \ = \sigma^2 + \mu^2$, e que $E[\bar{X}^2] = Var[\bar{X}] + E^2[\bar{X}] = \dfrac{\sigma^2}{n} + \mu^2$

Finalmente: $\dfrac{1}{n-1}E[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2] = \dfrac{1}{n-1}[(n\sigma^2 + n\mu^2) - (\sigma^2 + n\mu^2)] = \dfrac{1}{n-1}\sigma^2(n-1) = \sigma^2$

Portanto, S^2 é **não viesado**!

- $\hat{\sigma^2}$ é um estimador não viesado para σ^2? Considere uma v.a $X \sim N(\mu; \sigma^2)$.

  Solução: Já sabemos que $E[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2] = \sigma^2(n-1)$.

  Portanto, $\dfrac{1}{n}E[\sum_{i=1}^{n}(X_i^2 - n\bar{X})^2] = \dfrac{\sigma^2(n-1)}{n} \neq \sigma^2$.
  OBS: Note que $B(\hat{\sigma^2}) = -\dfrac{\sigma^2}{n}$, ou seja, **subestima** o parâmetro!

- Prove que $E[(\hat{\theta} - \theta)^2] = Var[\hat{\theta}] + B^2(\hat{\theta})$.

  Solução: Primeiro, vamos considerar $\hat{\theta} - \theta = \hat{\theta} - E[\hat{\theta}] + E[\hat{\theta}] - \theta$.

  Então, $(\hat{\theta} - \theta)^2 = (\hat{\theta} - E[\hat{\theta}])^2 + 2(\hat{\theta} - E[\hat{\theta}])(E[\hat{\theta}] - \theta) + (E[\hat{\theta}] - \theta)^2$.

  Logo, $E[(\hat{\theta} - \theta)^2] = E[(\hat{\theta} - E[\hat{\theta}])^2] + E[(E[\hat{\theta}] - \theta)^2] = Var[\hat{\theta}] + B^2(\hat{\theta})$.

  OBS: Note que $E[2(\hat{\theta} - E[\hat{\theta}])(E[\hat{\theta}] - \theta)] = 2(E[\hat{\theta}] - E[\hat{\theta}])(E[\hat{\theta}] - E[\theta]) = 0$.

- Encontre a distribuição de S^2. Considere $W \sim Gama(\alpha; \lambda)$ e utilize sua $f.g.m.$ para achar o resultado.

  Solução: Sabemos que $S^2 = \dfrac{\sigma^2 V}{n-1}$, onde $V \sim \chi^2_{(n-1)}$. Usando a função geradora de momentos ficamos com o seguinte:

  $$M_{S^2}(t) = E\left[e^{tS^2}\right] = E\left[e^{t\frac{\sigma^2 V}{n-1}}\right] = M_V(\tfrac{t\sigma^2}{n-1}) = \left[1 - \tfrac{2t\sigma^2}{n-1}\right]^{-(n-1)/2}.$$

  Sabemos que a f.g.m de W pode ser posta na seguinte forma: $\left[\frac{1}{1-\frac{t}{\lambda}}\right]^{\alpha} = [1 - \frac{t}{\lambda}]^{-\alpha}$.

Analisando atentamente, temos que $\alpha = (n-1)/2$ e $\lambda = \frac{n-1}{2\sigma^2}$.

Logo, $S^2 \sim Gama\left(\frac{n-1}{2}; \frac{n-1}{2\sigma^2}\right)$. Vamos testar?

$$E[S^2] = \frac{\alpha}{\lambda} = \frac{\frac{n-1}{2}}{\frac{n-1}{2\sigma^2}} = \sigma^2.$$
$$Var[S^2] = \frac{\alpha}{\lambda^2} = \frac{\frac{n-1}{2}}{\frac{(n-1)^2}{4\sigma^4}} = \frac{2\sigma^4}{n-1}.$$

Eficiência entre estimadores

Sejam T_1 e T_2 dois estimadores de θ . Se $\forall \theta \in \Theta$ a seguinte desigualdade ocorrer, para pelo menos um valor de θ, então o **melhor estimador** será aquele de **menor EQM**.

$$EQM(T_1) \leq EQM(T_2)$$

Se T_1 e T_2 forem não viesados, o melhor será aquele de **menor variância**.

Família exponencial

A *Família exponencial* constitui um importante grupo de funções (f.d.p e f.m.p) com características bem definidas pelos seguintes critérios:

1. O suporte A da distribuição **não** dependende do parâmetro desconhecido, digamos, θ;
2. $\log(f(x|\theta)) = c(\theta)T(x) + d(\theta) + h(x)$.

É importante destacar que nem sempre as componentes vão estar explícitas na forma acima, por exemplo, há funções onde $h(x) = 0$, sendo assim omitida da equação.

Fatos importantes decorrentes da família exponencial:

1. $E[T(x)] = -\dfrac{d'(\theta)}{c'(\theta)}$;
2. Seja $S = \sum_{i=1}^{n} T(X_i)$. S é uma **estatística suficiente e completa** para θ;
3. $E[S] = nE[T(x)] \ \ e \ \ Var[S] = nVar[T(x)]$;
4. $M_S(t) = [M_X(t)]^n$, que usamos para encontrar a **distribuição amostral** de S.

Para encontrar o UMVUE de $g(\theta)$, sabendo que a distribuição pertence à família exponencial, encontra-se uma função de S, $h(S)$, de tal forma que $E[h(S)] = g(\theta)$.

Exercícios Resolvidos 1

1. Seja $X \sim Ber(p)$. Mostre que X pertence à família exponencial.

   1.1. $f(x|p) = p^x(1-p)^{1-x}I_{\{0,1\}}(x),\ \theta = [0,1]$

   1.2. $\log f(x|p) = x\log p + \log(1-p) - x\log(1-p) = x\big[\log p - \log(1-p)\big] + \log(1-p)$

   1.3. $T(x) = x;\ c(p) = \log p - \log(1-p);\ d(p) = \log(1-p);\ h(x) = 0$

   Vemos que o suporte da distribuição **independe** de p. Logo, está provado que X pertence à família exponencial.

   Além disso, $S = \sum_{i=1}^{n} x_i$ é uma estatística suficiente e completa para p. Sua distribuição é: $M_S(t) = [M_X(t)]^n = \left[pe^t + (1-p)\right]^n \sim Bin(n,p)$

   Vamos encontrar o **UMVUE**: $E[S] = np \quad \therefore \quad E\left[\dfrac{S}{n}\right] = p \quad \therefore \quad \bar{X}$ é o nosso estimador procurado.

2. Seja $X \sim Gama(3;\theta)$. Essa v.a pertence à família exponencial? Verifique.

   2.1. $f(x|\theta) = \dfrac{\theta^3}{2}x^2e^{-\theta x}I_{(0,\infty)}(X), \Theta = (0,\infty)$

   2.2. $\log f(x|\theta) = 3\log\theta - \log 2 + 2\log x - \theta x$

   2.2. $T(x) = x;\ c(\theta) = -\theta;\ d(\theta) = 3\log\theta;\ h(x) = 2\log x - \log 2$

   Podemos ver que o suporte da nossa v.a **independe** do parâmetro. Portanto, provamos que X pertence à família exponencial.

   Além disso, $S = \sum_{i=1}^{n} x_i$ é uma estatística suficiente e completa, com distribuição: $M_S(t) = \left[\dfrac{\theta}{\theta - t}\right]^{3n}$, $Gama(3n;\theta)$

Mais definições

Considere X uma v.a com f.m.p ou f.d.p $f(x|\theta)$, com suporte A independente de θ.

- $V = \dfrac{\partial \log f(X|\theta)}{\partial\theta}$ é a **função escore**.
  OBS: $E[V] = 0$.

- $I_F(\theta) = Var[V] = E[V^2] = E^2\left[\frac{\partial \log f(X|\theta)}{\partial \theta}\right] = -E\left[\frac{\partial^2 \log f(X|\theta)}{\partial \theta^2}\right]$ é a **informação de Fisher**.[5]
  Note que como $E[V] = 0$, logo, $Var[V] = E[V^2]$.

- $LICR(\theta) = \frac{1}{nI_F(\theta)}$ é o **Limite Inferior de Crámer-Rao**, denotado também por LI.

- Se T é um estimador de $g(\theta)$, então $Var[T] \geq \frac{[g'(\theta)]^2}{nI_F(\theta)}$, que é chamada de **desigualdade da informação** (caso geral).

- $e(T) = \frac{LICR(\theta)}{Var[T]}$ é a eficiência do estimador.
  Se $e(T) = 1$, então o estimador é dito ser **eficiente**!

- Dizemos que T é uma **estatística suficiente** para θ se, e somente se, a distribuição conjunta da amostra puder ser fatorada dessa forma:
  $\prod_{i=1}^{n} f(x_i|\theta) = h(x_1,x_2,\dots,x_n)\ g_\theta(T(x_1,x_2,\dots,x_n))$ [6]

- **Teorema de Rao-Blackwell**: Seja S uma estatística suficiente de θ e T um estimador não-viesado de $g(\theta)$. Então $T^* = E[T|S]$ geralmente é o melhor estimador para θ.

- **Teorema de Lehmann-Scheffé**: T^* é o **UMVUE** (uniformly minimum variance unbiased estimador) [7].

Exercícios Resolvidos 2

1. Seja $X \sim Exp(\theta)$. Encontre a função escore e a informação de Fisher.

   1.1. $f(x|\theta) = \theta e^{-\theta X} I_A(X), A = (0,\infty), \Theta = (0,\infty)$

   1.2 $\log f(x|\theta) = \log\theta - \theta X$

   1.3 $V = \frac{1}{\theta} - X$. Observe que $E[V] = 0$

1.4 $I_F(\theta) = Var[V] = Var\left[\frac{1}{\theta} - X\right] = Var[X] = \frac{1}{\theta^2}$

2. Seja X uma v.a de seguinte distribuição: $f(x|\theta) = e^{-(x-\theta)} I_{(\theta,\infty)}(x), \theta > 0$. Encontre uma estatística suficiente para o parâmetro.

   Vemos que o suporte é dependente do parâmetro desconhecido, logo, não podemos usar a família exponencial. Usaremos o critério da fatoração, encontrando $h(\mathbf{x}) = e^{-s}$ e $g(\theta, y_1) = e^{n\theta} I_{(\theta,\infty)}(y_1)$. y_1 é nossa estatística suficiente:

$$L(\theta) = \prod_{i=1}^{n} e^{-(x_i-\theta)} I_{(\theta,\infty)}(x_i) = e^{-s} e^{n\theta} \prod_{i=1}^{n} I_{(\theta,\infty)}(x_i)$$
$$I_{(\theta,\infty)}(x_i) = 1 \Leftrightarrow x_1 > \theta; x_2 > \theta; \ldots ; x_n > \theta$$
$$\therefore$$
$$y_1 = min(x_1, x_2, \ldots, x_n) > \theta$$
$$\therefore$$
$$I_{(\theta,\infty)}(y_1)$$
$$\therefore$$
$$e^{-s} e^{n\theta} I_{(\theta,\infty)}(y_1)$$

3. Seja $T = 2\bar{X}$ um estimador não viesado para θ, parâmetro de $X \sim U[0;\theta]$. Seja o novo estimador $T_1 = E[2X|Y_n]$, onde Y_n é uma estatística suficiente e completa para o parâmetro. Qual nome recebe esse novo estimador?

   $E[2X|Y_n]$ é um estimador não viesado que é função de uma estatística suficiente. Logo, pelo teorema de *Lehmann-Sheffé* ou pelo teorema de *Rao-Blackwell*, T_1 é o nosso **UMVUE**.

4. Seja $X_1, \ldots, X_n$ uma A.A.S. de $X \sim Bin(2, \theta)$.
   a) Encontre o LICR.
   b) Encontre uma estatística suficiente para θ.
   c) Obtenha o **UMVUE** para θ.
   d) Descubra se o estimador é eficiente ou não.

   4.1. $f(X|\theta) = \binom{2}{X} \theta^X (1-\theta)^{2-X} I_{\{0,1,2\}}(X)$

   4.2. $\log f(X|\theta) = \log \binom{2}{X} + X \log \theta + (2 - X) \log (1 - \theta)$

   4.3. $V = \frac{X}{\theta} - \frac{(2-X)}{(1-\theta)} = \frac{X(1-\theta) - \theta(2-X)}{\theta(1-\theta)} = \frac{X - 2\theta}{\theta(1-\theta)}$

   4.4. $I_F(\theta) = Var[V] = \frac{2\theta(1-\theta)}{\theta^2(1-\theta)^2} = \frac{2}{\theta(1-\theta)}$

4.5. $LICR = \dfrac{1}{n\frac{2}{\theta(1-\theta)}} = \dfrac{\theta(1-\theta)}{2n}$

Vamos encontrar a estatística suficiente pelo critério da fatoração (poderia usar família exponencial):

$L(\theta) = \prod_{i=1}^{n} \binom{2}{X_i}\theta^{X_i}(1-\theta)^{2-X_i} = \theta^s(1-\theta)^{2n-s}\prod_{i=1}^{n}\binom{2}{X_i},\ \text{com } s = \sum_{i=1}^{n} X_i.\ L(\theta) = \left[\dfrac{\theta}{(1-\theta)}\right]^s (1-\theta)^{2n}\prod_{i=1}^{n}\binom{2}{X_i} = g(\theta,s)h(\mathbf{x}).$

Assim, nossa estatística suficiente é $s = \sum_{i=1}^{n} X_i$.

4.6. Sabemos que $s \sim Bin(2n;\theta), \text{com } E[s] = 2n\theta$. Então, $E\left[\dfrac{s}{2n}\right] = \theta = E\left[\dfrac{\bar{X}}{2}\right]$ é nosso **UMVUE** procurado.

4.7. $Var\left[\dfrac{\bar{X}}{2}\right] = \dfrac{1}{4}\dfrac{2\theta(1-\theta)}{n} = \dfrac{\theta(1-\theta)}{2n} = LICR.$ Nosso estimador é eficiente!

Métodos de estimação

Método dos momentos

É o método mais simples e mais antigo já utilizado, e que consiste em igualar o primeiro momento populacional ao primeiro momento amostral. Em outras palavras, basta igualar $\bar{X}$ a μ: $E[X] = \bar{X}$

Método da máxima verossimilhança

Para encontrar o estimador de máxima verossimilhança, em geral, segue-se os seguintes passos:

1. Encontrar $L(\theta;\boldsymbol{x})$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$
2. Encontrar a função log-verossimilhança: $l(\theta;\boldsymbol{x}) = \log L(\theta;\boldsymbol{x})$
3. Derivar $l(\theta;\boldsymbol{x})$ em relação a θ e igualar a 0: $\dfrac{\partial l(\theta;\boldsymbol{x})}{\partial\theta} = 0$
4. Derivar novamente para provar que o ponto encontrado no passo 3 é o máximo: $\dfrac{\partial^2 l(\theta;\boldsymbol{x})}{\partial\theta^2}$

Note que, em suma, estamos **maximizando** a função de versossimilhança, de forma que seja o mais provável possível encontrar o ponto em nossa amostra.

Propriedades dos EMV:

- O EMV é sempre função de uma estatística suficiente para o parâmetro
- Se $\hat{\theta}$ é um EMV, então $g(\hat{\theta})$ é o EMV para $g(\theta)$ [8]

- Para grandes amostras: $\sqrt{n}\left(g(\hat{\theta})-g(\theta)\right) \overset{a}{\sim} N\left(0, \dfrac{[g'(\theta)]^2}{I_F(\theta)}\right)$ [9]

Método dos mínimos quadrados

Ao contrário dos estimadores de máxima verossimilhança, onde queremos maximizar uma função, nos mínimos quadrados, como o próprio nome sugere, vamos **minimizar** uma soma de quadrados, que na verdade é a soma dos quadrados dos desvios (X - E[X])². Veja os passos a seguir:

1. Encontre $S(x)=\sum_{i=1}^{n}(X_i-E[X_i])^2$
2. Faça $S'(x)=0$
3. Mostre que $S''(x)>0$ para provar que o ponto encontrado no passo 2 é de mínimo

Exercícios Resolvidos 3

1. Seja $X \sim Exp(\theta)$. Encontre o estimador pelo método dos momentos.

$$\mu=\bar{X} \quad \therefore \quad \frac{1}{\theta}=\bar{X} \quad \therefore \quad \theta_{MM}=\frac{1}{\bar{X}}$$

2. Seja $X \sim U(0;\theta)$. Encontre o estimador pelo método dos momentos.

$$\mu=\bar{X} \quad \therefore \quad \frac{\theta}{2}=\bar{X} \quad \therefore \quad \theta_{MM}=2\bar{X}$$

3. Ache o estimador pelo método da máxima verossimilhança para a v.a do item 1.

   3.1. $L(\theta)=\prod_{i=1}^{n} f(x_i|\theta)=\prod_{i=1}^{n}\theta e^{-\theta x_i}=\theta^n e^{-\theta s}$
   com $s=\sum_{i=1}^{n} X_i$

   3.2. $l(\theta)=\log L(\theta)=n\log\theta-\theta s$

   3.3. $l'(\theta)=\dfrac{n}{\theta}-s$

   3.4. $\dfrac{n}{\theta}=s \quad \therefore \quad \dfrac{1}{\theta}=\bar{X} \quad \therefore \quad \theta_{MV}=\dfrac{1}{\bar{X}}$

   3.5. $l''(\theta)=-\dfrac{n}{\theta^2}<0$, portanto, ponto de máximo

4. Ache o estimador de máxima verossimilhança para a seguinte v.a:
   $f(x|\theta)=\dfrac{x}{\theta^2}\exp(-\dfrac{x}{\theta})I_A(x), A=(0,\infty), \theta>0$. Verifique se ele é eficiente.
   4.1. $L(\theta)=\prod_{i=1}^{n} f(x_i|\theta)=\prod_{i=1}^{n}\dfrac{x_i}{\theta^2}\exp(-\dfrac{x_i}{\theta})=\theta^{-2n}e^{-\frac{s}{\theta}}\prod_{i=1}^{n}x_i$, com $s=\sum_{i=1}^{n}x_i$

   4.2. $l(\theta)=-2n\log\theta-\dfrac{s}{\theta}+\sum_{i=1}^{n}\log x_i$

4.3. $l'(\theta) = \frac{-2n}{\theta} + \frac{s}{\theta^2}$

4.4. $\theta_{MV} = \frac{s}{2n} = \frac{\bar{X}}{2}$

4.5. $l''(\theta) = \frac{2n}{\theta^2} - \frac{2s}{\theta^3} = \frac{2n}{\theta^2} - \frac{2n\bar{X}}{\theta^3} < 0$, logo, ponto de máximo.

Veremos agora se ele é eficiente:

4.6. $\log f(x|\theta) = \log X - 2\log\theta - \frac{x}{\theta}$

4.7. $V = -\frac{2}{\theta} + \frac{x}{\theta^2}$

4.8. $Var[V] = \frac{1}{\theta^4} Var[X] = \frac{2}{\theta^2} = I_F(\theta)$, pois $Var[X] = 2\theta^2; X \sim Gama(\frac{1}{\theta}, 2)$

4.9. $LICR = \frac{1}{n\frac{2}{\theta^2}} = \frac{\theta^2}{2n}$

4.9.1 $e\left(\frac{\bar{X}}{2}\right) = \frac{\frac{\theta^2}{2n}}{Var\left[\frac{\bar{X}}{2}\right]} = 1$. Nosso estimador **é eficiente**

5. Ache o estimador pelo método dos mínimos quadrados para a v.a do item 1.

5.1. $S(\theta) = \sum_{i=1}^{n}(X_i - E[X_i])^2 = \sum_{i=1}^{n}(X_i - \frac{1}{\theta})^2 = \sum_{i=1}^{n} X_i^2 - 2n\bar{X}\frac{1}{\theta} + \frac{n}{\theta^2}$

5.2. $S'(\theta) = \frac{2n\bar{X}}{\theta^2} - \frac{2n}{\theta^3}$

5.3. $\theta_{MM} = \frac{1}{\bar{X}}$

5.4. $S''(\theta) = \frac{-4n\bar{X}}{\theta^3} + \frac{6n}{\theta^4} > 0$, portanto, ponto de mínimo

6. Considere o modelo de regressão: $\beta\sqrt{X_i} + u_i, i = 1, 2, n$, com $u_i's$ independentes. Considere também que $E[u_i] = 0$ e $Var[u_i] = \sigma^2$. Encontre o EMQ de β.

$S(\beta) = \sum_{i=1}^{n}(Y_i - E[Y_i])^2$. Vamos encontrar a esperança e variância de Y_i primeiro:

6.1. $E[Y_i] = E[\beta\sqrt{X_i} + u_i] = E[\beta\sqrt{X_i}] + E[u_i] = \beta\sqrt{X_i}$

6.2. $Var[\beta\sqrt{X_i} + u_i] = Var[u_i] = \sigma^2$

Agora, vamos prosseguir.

6.3. $S(\beta) = \sum_{i=1}^{n}(Y_i^2 - 2Y_i\beta\sqrt{X_i} + \beta^2 X_i) = \sum_{i=1}^{n} Y_i^2 - 2\beta\sum_{i=1}^{n} Y_i\sqrt{X_i} + \beta^2\sum_{i=1}^{n} X_i$

6.4. $S'(\beta) = -2\sum_{i=1}^{n} Y_i\sqrt{X_i} + 2\beta\sum_{i=1}^{n} X_i \quad \therefore \quad \beta = \dfrac{\sum_{i=1}^{n} Y_i\sqrt{X_i}}{\sum_{i=1}^{n} X_i}$

6.5. $S''(\beta) = 2\sum_{i=1}^{n} X_i > 0$, logo, ponto de mínimo

7. Para o item 1., encontre o EMV de $P(X > 2) = g(\theta)$ e sua distribuição em grandes amostras.

7.1. $P(X > 2) = 1 - P(X \leq 2) = 1 - [1 - e^{-2\theta}] = e^{-2\theta} = g(\theta)$

7.2. Pelo princípio da invariância, $g(\hat{\theta}) = g(\theta) \quad \therefore \quad g(\hat{\theta}) = e^{-\frac{2}{\bar{X}}}$

7.3. Já vimos anteriormente que $I_F(\theta) = \dfrac{1}{\theta^2}$. É fácil ver também que $\left[g'(\theta)\right]^2 = 4e^{-4\theta}$. Logo, em grandes amostras:

$$\sqrt{n}\big(g(\hat{\theta}) - g(\theta)\big) \overset{a}{\sim} N\big(0; 4\theta^2 e^{-4\theta}\big)$$

Estimação intervalar

Estimações intervalares, ou seja, entre intervalos, são construídas através do método da **quantidade pivotal**, que é uma variável aleatória especial, cuja distribuição **não** depende de parâmetros desconhecidos.

Tais estimativas partem da vertente frequentista da estatística, sob a seguinte interpretação: para um número grande de vezes, e sob as mesmas circunstâncias de realização, o parâmetro a ser estimado encontra-se dentro de X% dos intervalos encontrados, onde X é o nível de confiança pré-estabelecido.

Geralmente, são estabelecidos níveis $\gamma = 1 - \alpha$ de confiança de 90%, 95% e 99%, onde α é o nível de significância. Porém, outros níveis de confiança podem também ser utilizados. Entretando, deve-se observar a influência do mesmo no comprimento do intervalo, como será visto a seguir.

Intervalos Bilaterais

1. Seja $X_1, X_2, \ldots, X_n$ uma A.A com distribuição $N(\mu, \sigma^2)$.

- Intervalo de confiança para μ, com σ^2 **conhecida** [10]:

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}};\ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

Quantidade pivotal: $Q = \dfrac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0;1)$

- Intervalo de confiança para μ, com σ^2 **desconhecida**:

$$\left[\bar{X} - t_{\alpha/2}\frac{S}{\sqrt{n}};\ \bar{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}\right]$$

Quantidade pivotal: $Q = \dfrac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{(n-1)}$

- Intervalo de confiança para σ^2:

$$\left[\frac{(n-1)S^2}{q_2};\ \frac{(n-1)S^2}{q_1}\right]$$

Quantidade pivotal: $Q = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

2. Seja $X_1, X_2, \ldots, X_n$ uma A.A com distribuição $Exp(\theta)$.

Intervalo de confiança para θ:

$$\left[\frac{q_1}{2S};\ \frac{q_2}{2S}\right]$$

Quantidade pivotal: $Q = 2\theta S \sim \chi_{(2n)}, S = \sum_{i=1}^{n} X_i \sim Gama(n;\theta)$

3. Seja $X_1, X_2, \ldots, X_n$ uma A.A com distribuição $Ber(p)$.

   Intervalo de confiança de MV para p:

$$\left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}};\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

   Onde $\hat{p}$ é o estimador de máxima verossimilhança para p.
   Quantidade pivotal: $Q = \dfrac{\hat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0;1)$
   É importante salientar que existem outros intervalos para a proporção populacional, como o **conservador** e o de **MV com fator de correção**. Porém, por terem notações muito extensas, não serão colocados aqui.

4. Seja $X_1, X_2, \ldots, X_n$ uma A.A de tamanho n com distribuição $N(\mu_1, \sigma_1^2)$. Temos que $\bar{X}$ e S_1^2 são **independentes**. Agora, seja $Y_1, Y_2, \ldots, Y_m$ uma A.A de tamanho m com distribuição $N(\mu_2, \sigma_2^2)$. Temos que $\bar{Y}$ e S_2^2 são **independentes**. Temos também que $U = \frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2_{(n-1)}$ e $V = \frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2_{(n-1)}$. As variáveis $\bar{X}, \bar{Y}, U$ e V são **independentes** entre si. [11]

   - Intervalo de confiança para $\mu_1 - \mu_2$, com σ_1^2 e σ_2^2 **conhecidas**:

$$\left[(\bar{X} - \bar{Y}) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}};\ (\bar{X} - \bar{Y}) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right]$$

   Quantidade pivotal: $Q = \dfrac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n} + \dfrac{\sigma_2^2}{m}}} \sim N(0,1)$

   - Intervalo de confiança para $\mu_1 - \mu_2$, com $\sigma_1^2 = \sigma_2^2$ **deconhecidas**:

$$\left[(\bar{X}-\bar{Y})-t_{\alpha/2}S_p\sqrt{\frac{1}{n}+\frac{1}{m}};\ (\bar{X}-\bar{Y})+t_{\alpha/2}S_p\sqrt{\frac{1}{n}+\frac{1}{m}}\right]$$

Onde $S_p=\sqrt{\dfrac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}$.

Quantidade pivotal: $Q=\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n}+\frac{1}{m}}}\sim t_{(n+m-2)}$

- Intervalo de confiança para $\mu_1-\mu_2$, com $\sigma_1^2\neq\sigma_2^2$ **deconhecidas**:

$$\left[(\bar{X}-\bar{Y})-t_{\alpha/2}\sqrt{\frac{S_1^2}{n}+\frac{S_2^2}{m}};\ (\bar{X}-\bar{Y})+t_{\alpha/2}\sqrt{\frac{S_1^2}{n}+\frac{S_2^2}{m}}\right]$$

Quantidade pivotal: $Q=\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n}+\frac{S_2^2}{m}}}\sim t_{(r)}$, onde $r=\dfrac{(A+B)^2}{\frac{A^2}{n-1}+\frac{B^2}{m-1}}$, com $A=\dfrac{S_1^2}{n}$ e $B=\dfrac{S_2^2}{m}$. r é o número inteiro mais próximo.

- Intervalo de confiança para $\dfrac{\sigma_1^2}{\sigma_2^2}$:

$$\left[f_1\frac{S_2^2}{S_1^2};\ f_2\frac{S_2^2}{S_1^2}\right]$$

Quantidade pivotal: $Q=\dfrac{S_1^2}{S_2^2}\dfrac{\sigma_2^2}{\sigma_1^2}\sim F(n-1;m-1)$

5. Seja $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ uma A.A de $(X,Y)\sim N(\mu_1,\mu_2;\rho,\sigma_1^2,\sigma_2^2)$.

Intervalo de confiança para $D=X-Y$:

$$\left[\bar{D} - t_{\alpha/2}\frac{S_D}{\sqrt{n}};\ \bar{D} + t_{\alpha/2}\frac{S_D}{\sqrt{n}}\right]$$

Quantidade pivotal: $Q = \dfrac{\bar{D} - \mu_D}{S_D/\sqrt{n}} \sim t_{(n-1)}$

Quando temos amostras advindas de processos SEM REPOSIÇÃO, multiplicamos os limites pelo fator de correção $\dfrac{N-n}{N-1}$. [12]

Note que, nos intervalos para a média populacional, o valor de $\bar{X}$ pode ser encontrado da seguinte forma: $\bar{X} = \dfrac{li + ls}{2}$

Intervalos Unilaterais

A definição de intervalos de confiança, sejam eles bilaterais ou unilaterais, está intrinsecamente ligada à testes de hipótese, pois é a partir de hipóteses pré formuladas que se constroem tais intervalos.

Há dois tipos de hipóteses: H_0 e H_1, que correspondem respectivamente à hipótese nula ("principal") e à hipótese alternativa ("secundária").

As mesmas ainda podem ser classificadas em duas categorias:

- Simples: quando os valores supostos **são exatos** (ou seja, são iguais a um valor X)
- Compostas: quando os valores **não são exatos** (ou seja, são maiores, menores ou diferentes de um valor X)

O tipo de intervalo que é construído para testar as hipóteses, é baseado na direção da hipótese alternativa.

Quando H_1 é da forma $H_1 : \mu \neq X$, construímos intervalos bilaterais, pois o valor real pode estar tanto acima quanto abaixo do valor suposto em H_0.

Já quando H_1 é da forma $H_1 : \mu > X$ ou $H_1 : \mu < X$, usamos intervalos unilaterais, pois o valor real só se encontra ou acima ou abaixo do valor suposto em H_0.

1. Para $H_1 : \mu > X$, temos um intervalo unilateral **à direita** ou **superior**.
2. Para $H_1 : \mu < X$, temos um intervalo unilateral **à esquerda** ou **inferior**.

Para construir esses intervalos, procedemos da mesma forma que os intervalos bilaterais já vistos, com a diferença de que, se a hipótese alternativa for da primeira forma, então temos $IC[-\infty, ls]$. Caso tenhamos a segunda forma, ficamos com $IC[li, \infty]$.

Para proporções, a diferença é que substituímos $-\infty \to 0, \text{ e } \infty \to 1$.

Testes de Hipótese

Como visto anteriormente, testes de hipótese estão diretamente relacionados à construção de intervalos de confiança. Porém, para a construção dos mesmos, precisamos estabelecer previamente as hipóteses nula e alternativa, a fim de poder analisar e elucidar um resultado posteriormente.

Entretanto, ao optar por uma hipótese, nos sujeitamos a dois tipos de erro, que são eles:

- *Erro de tipo I*: rejeitar H_0 sendo H_0 verdade
- *Erro de tipo II*: não rejeitar H_0 sendo H_0 falsa

Na Inferência, o cálculo da probabilidade de ambos os erros recebem nomenclaturas especiais, sendo elas:

- α o **tamanho do erro de tipo I**, que é $P(Rejeitar\ H_0|H_0\ verdade)$
- β o **tamanho do erro de tipo II**, que é $P(Não\, rejeitar\ H_0|H_0\ falsa)$

Note que α também é o *nível de significância*, geralmente fixado em 5%, 2% ou 1%.

Outro fator muito relevante neste conteúdo é o cálculo do *valor-p*, sendo o mesmo a probabilidade de encontrar um valor pelo menos tão extremo quanto o observado na amostra, **respeitando a direção** de H_1.

Assim, o *valor-p* é, na verdade, o menor nível de significância encontrado, ou seja, pode ser comparado ao α estabelecido para saber se deve-se rejeitar ou não H_0.

Outra medida importante no estudo dos testes de hipótese é a **função poder**, possuindo a forma $\pi(\theta) = P(Rejeitar\ H_0|\theta)$, sendo θ um valor qualquer para o parâmetro estimado, digamos, θ. Assim:

- $\pi(\theta_0) = \alpha$
- $\pi(\theta_1) = 1 - \beta \Rightarrow P(\text{Rejeitar } H_0|\theta_1) = 1 - P(\text{Não Rejeitar } H_0|\theta_1) = 1 - \beta$

Assim, a função poder e a construção de gráficos de poder são basedas no complementar do tamanho do erro de tipo II.

Métodos

A fim de resumir os procedimentos possíveis para se analisar hipóteses, listo a seguir as formas vistas na disciplina.

1. *Intervalos de Confiança*: o valor de H_0 encontra-se dentro ou fora do intervalo encontrado?
2. *Região Crítica (R.C) e Região de Aceitação (R.A)*: a estatística encontra-se na R.C ou na R.A?
3. *Valor-p*: $\alpha > \hat{\alpha}$? Ou $\alpha < \hat{\alpha}$?
4. *Estatística de teste*: $zcalc > ztab$? Ou $zcalc < ztab$?[13]

Procedimentos em uma amostra

1. $H_0 : \mu = \mu_0, \; \sigma^2$ **conhecida**

   Estatística de teste: $z_0 = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

H_1	**Rejeita H_0 se**
$\mu \neq \mu_0$	$\lvert z_0 \rvert > z_{\alpha/2}$
$\mu > \mu_0$	$z_0 > z_\alpha$
$\mu < \mu_0$	$z_0 < -z_\alpha$

2. $H_0 : \mu = \mu_0, \; \sigma^2$ **desconhecida**

   Estatística de teste: $t_0 = \dfrac{\bar{x} - \mu_0}{S / \sqrt{n}}$

H_1	**Rejeita H_0 se**
$\mu \neq \mu_0$	$\lvert t_0 \rvert > t_{\alpha/2}, (n-1)$
$\mu > \mu_0$	$t_0 > t_\alpha, (n-1)$
$\mu < \mu_0$	$t_0 < -t_\alpha, (n-1)$

3. $H_0 : \sigma^2 = \sigma_0^2$

   Estatística de teste: $\chi_0^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

H_1	Rejeita H_0 se
$\sigma^2 \neq \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha/2}, (n-1)$ ou $\chi_0^2 < \chi^2_{1-\alpha/2}, (n-1)$
$\sigma^2 > \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha}, (n-1)$
$\sigma^2 < \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha}, (n-1)$

4. $H_0 : p = p_0$

   Estatística de teste: $z_0 = \dfrac{x - np_0}{\sqrt{np_0 q_0}}$, onde x é o número de sucessos na amostra

H_1	Rejeita H_0 se
$p \neq p_0$	$\lvert z_0 \rvert > z_{\alpha/2}$
$p > p_0$	$z_0 > z_\alpha$
$p < p_0$	$z_0 < -z_\alpha$

Procedimentos em duas amostras

1. $H_0 : \mu_1 - \mu_2 = \Delta_0$, $\boldsymbol{\sigma_1^2}$ e $\boldsymbol{\sigma_2^2}$ **conhecidas**

   Estatística de teste: $z_0 = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n} + \dfrac{\sigma_2^2}{m}}}$

H_1	Rejeita H_0 se
$\mu_1 - \mu_2 \neq \Delta_0$	$\lvert z_0 \rvert > z_{\alpha/2}$
$\mu_1 - \mu_2 > \Delta_0$	$z_0 > z_\alpha$
$\mu_1 - \mu_2 < \Delta_0$	$z_0 < -z_\alpha$

2. $H_0 : \mu_1 - \mu_2 = \Delta_0$, $\boldsymbol{\sigma_1^2 = \sigma_2^2}$ **desconhecidas**

Estatística de teste: $t_0 = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{S_p\sqrt{\dfrac{1}{n} + \dfrac{1}{m}}}$

H_1	**Rejeita H_0 se**
$\mu_1 - \mu_2 \neq \Delta_0$	$\|t_0\| > t_{\alpha/2}, (n+m-2)$
$\mu_1 - \mu_2 > \Delta_0$	$t_0 > t_\alpha, (n+m-2)$
$\mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_\alpha, (n+m-2)$

3. $H_0 : \mu_1 - \mu_2 = \Delta_0$, $\boldsymbol{\sigma_1^2 \neq \sigma_2^2}$ **desconhecidas**

Estatística de teste: $t_0 = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\dfrac{S_1^2}{n} + \dfrac{S_2^2}{m}}}$

H_1	**Rejeita H_0 se**
$\mu_1 - \mu_2 \neq \Delta_0$	$\|t_0\| > t_{\alpha/2}, (v)$
$\mu_1 - \mu_2 > \Delta_0$	$t_0 > t_\alpha, (v)$
$\mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_\alpha, (v)$

Sendo $v = \dfrac{\left(\dfrac{S_1^2}{n} + \dfrac{S_2^2}{m}\right)^2}{\dfrac{(S_1^2/n)^2}{n-1} + \dfrac{(S_2^2/m)^2}{m-1}}$

4. $H_0 : \mu_D = 0$ (**populações dependentes**)

Estatística de teste: $t_0 = \dfrac{\bar{d}}{s_D/\sqrt{n}}$

H_1	**Rejeita H_0 se**
$\mu_D \neq 0$	$\|t_0\| > t_{\alpha/2}, (n-1)$
$\mu_D > 0$	$t_0 > t_\alpha, (n-1)$

H_1	**Rejeita H_0 se**
$\mu_D < 0$	$t_0 < -t_{\alpha}, (n-1)$

5. $H_0 : \sigma_1^2 = \sigma_2^2$

   Estatística de teste: $f_0 = S_1^2 / S_2^2$

H_1	**Rejeita H_0 se**
$\sigma_1^2 \neq \sigma_2^2$	$f_0 > f_{\alpha/2}, (n-1, m-1)$ ou $f_0 < f_{1-\alpha/2}, (n-1, m-1)$
$\sigma_1^2 > \sigma_2^2$	$f_0 > f_{\alpha}, (n-1, m-1)$

6. $H_0 : p_1 = p_2$

   Estatística de teste: $z_0 = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left[\dfrac{1}{n} + \dfrac{1}{m}\right]}}$

H_1	**Rejeita H_0 se**
$p_1 \neq p_2$	$\lvert z_0 \rvert > z_{\alpha/2}$
$p_1 > p_2$	$z_0 > z_{\alpha}$
$p_1 < p_2$	$z_0 < -z_{\alpha}$

Teste mais poderoso

O *lema de Neyman-Pearson* nos diz que o teste mais poderoso será aquele de região crítica dada por

$$A^* = \left\{ \mathbf{x}; \frac{L_1(\mathbf{x})}{L_0(\mathbf{x})} \geq c \right\}$$

Caso A^* **não** dependa do parâmetro a ser testado, então dizemos que o teste realizado é, na verdade, o teste uniformemente mais poderoso (**UMP**).

---

1. logo, **todo estimador é uma estatística**, mas nem toda estatística é um estimador. ↩
2. um estimador pode subestimar ou superestimar um parâmetro. ↩
3. se o estimador for **não viesado**, então o EQM será a variância do estimador. ↩
4. é a distribuição conjunta da amostra analisada como função de θ ↩
5. em algumas situações, é mais fácil usar a última forma, pois a variância torna-se mais difícil de encontrar! ↩
6. esse resultado é conhecido como **critério da fatoração de Neyman**, e é mais utilizado quando o suporte da distribuição depender do parâmetro; do contrário, encontra-se a estatística suficiente através da família exponencial ↩
7. A ideia é que dado um estimador não-viesado e que seja função de uma estatística suficiente, então, tal estimador é o UMVUE ↩
8. este é o princípio da **invariância** ↩
9. distribuição assintótica dos EMV ↩
10. $e = z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$ é o erro amostral. Quando o mesmo é especificado e queremos achar o tamanho da amostra, temos o seguinte: $n = \left\lceil \left(\dfrac{z_{\alpha/2}\ \sigma}{e} \right)^2 \right\rceil$. A situação é similar quando a variância é desconhecida, ou para quando temos duas populações ↩
11. há uma regra prática para decidir se as variâncias são iguais ou distintas: se $\dfrac{\max\left(S_1^2, S_2^2\right)}{\min\left(S_1^2, S_2^2\right)} < 4$, então, considerar **variâncias iguais** ↩
12. nesses casos, para encontrar o tamanho amostral, usamos agora $n = \dfrac{z_{\alpha/2}^2\ \sigma^2\ N}{(N-1)e^2 + z_{\alpha/2}^2 \sigma^2}$ para estimar a média, e $n = \dfrac{z_{\alpha/2}^2\ \bar{p}\ \bar{q}\ N}{(N-1)e^2 + z_{\alpha/2}^2\ \bar{p}\ \bar{q}}$ para estimar a proporção ↩
13. este é um tipo de região crítica, utilizando a estatística de teste ao invés da estatística observada ↩