ESTATÍSTICA II

(ver página da disciplina na Internet)

Linhas programáticas

1. Inferência Estatística: estimação por pontos e por intervalos, teste de hipóteses, testes não paramétricos;
2. Introdução à Econometria: O modelo de regressão linear, Inferência sobre o modelo, Previsão, Tópicos sobre a forma funcional, heterocedasticidade

Bibliografia

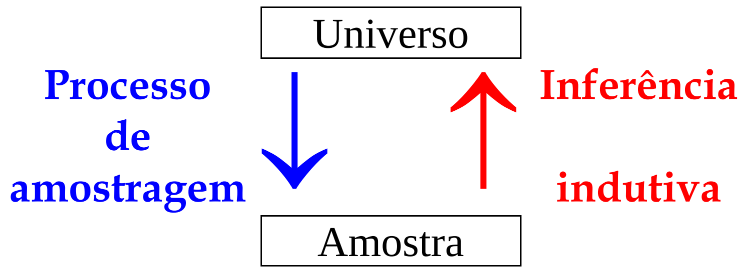
Murteira, B; Silva Ribeiro, C. ; Andrade e Silva, J. ; Pimenta, C. e F. Pimenta; *Introdução à Estatística*, 3ª ed., Escolar Editora, 2015.

Wooldridge, J.M. *Introductory Econometrics, a modern approach*, 6th ed, Cengage, 2016.

Avaliação

EN e ER: Exame final (100%)

Conceitos fundamentais de amostragem



- Universo ou População
    - População objetivo e população inquirida
    - Especificação de um modelo (família de modelos)
        - Modelos paramétricos e não paramétricos
        - População $X \sim f(x \mid \theta)$ com θ desconhecido
            - Espaço do parâmetro Θ - Dar alguns exemplo

Universo

Processo de amostragem $\downarrow$ $\uparrow$ **Inferência indutiva**

Amostra

- Processo de amostragem
    - Amostra casual simples $(X_1, X_2, \cdots, X_n)$
        - Observações **independentes** e **identicamente distribuídas** (i.i.d.)
        - Cada X_i tem distribuição dada por $f(x \mid \theta)$
    - Espaço-amostra
    - Distribuição da amostra

- Processo de amostragem (continua)
    - Noção de estatística
        - Definição
        - Amostra e estatística – a redução da informação
            - Exemplos de estatísticas
                - A própria amostra
                - A média da amostra $\bar{X} = (1/n)\sum_{i=1}^{n} X_i$
                - A variância da amostra $S^2 = (1/n)\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = (1/n)\sum_{i=1}^{n} X_i^2 - \bar{X}^2$
                - A variância corrigida da amostra $S'^2 = (1/(n-1))\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = (n/(n-1))\, S^2$
                - O par $\left(\bar{X}, S'^2\right)$
                - O máximo da amostra
                - O 1º valor da amostra
        - Distribuição por amostragem de uma estatística
        - Estatísticas e valores observados das estatísticas

Distribuições mais utilizadas

- **Universos normais**
  - Distribuição da média da amostra com (ou sem) variância (do universo) conhecida
  - Distribuição da variância da amostra
  - Distribuição da diferença de médias de duas amostras independentes

- **Outros universos**
  - Grandes amostras – Caso Geral – **Teorema do limite central**
    - Populações de Bernoulli
    - Populações de Poisson

    ...

ESTIMAÇÃO

- **Estimação paramétrica**

  - $(X_1, X_2, ..., X_n)$ amostra casual de uma população cuja função densidade (ou função probabilidade) pertence à família $\mathcal{F}_\theta = \{ f(x \mid \theta) : \theta \in \Theta \}$.

  - A forma funcional $f(.)$ é conhecida, apenas se desconhecendo o (verdadeiro) valor do parâmetro.

- **Problema**: Como utilizar a informação dada pela amostra para "*adivinhar*" o valor de θ?
(estimar)

- Dois aspectos a ter em, conta: a **precisão** e a **confiança.**
  - Ideia Importante $\rightarrow$ Fixada a dimensão da amostra, quanto mais precisa a resposta, menor a confiança que nela se deposita.
  - A estimação paramétrica vai então desenvolver-se ou privilegiando a precisão (**estimação por pontos**) ou privilegiando a confiança (**estimação por intervalos**).

Ter presente que o parâmetro de interesse pode ser

- **multidimensional**

  Exemplo $\rightarrow$ Suponha que a valorização de um activo financeiro tem distribuição normal com média μ e variância σ^2. Observada uma amostra casual, pretende-se estimar μ e σ (rendibilidade esperada e risco).

  Temos 2 parâmetros desconhecidos.

- uma **função do(s) parâmetro(s) da distribuição**

  Exemplo $\rightarrow$ Suponha-se que o número de sinistros originados anualmente por uma apólice de seguro automóvel tem distribuição de Poisson de parâmetro (desconhecido) λ.

  Em vez de nos interessarmos pelo parâmetro λ (que representa a média do fenómeno) podemos estar interessados numa função de λ, por exemplo, na probabilidade de não se verificar nenhum sinistro $\Pr(X=0\,|\,\lambda)=e^{-\lambda}$.
  Assim pretende-se estimar $\tau(\lambda)=e^{-\lambda}$ que traduz essa probabilidade.

ESTIMAÇÃO POR PONTOS

- **Ponto de Partida**:
  - Amostra casual $(X_1, X_2, ..., X_n)$, de uma população $F_\theta = \{f(x|\theta) : \theta \in \Theta\}$.
  - Apenas θ é desconhecido. No entanto, $\theta \in \Theta$ sendo Θ (espaço do parâmetro) conhecido.

- **Conceitos Fundamentais**:
  - **Estimador**: Processo de obtenção de estimativas.

    O **estimador** é uma variável aleatória, função da amostra casual e representa-se por $T(X_1, X_2, ..., X_n)$ - Por exemplo $T(X_1, X_2, ..., X_n) = \bar{X}$. **Um estimador é uma estatística.**
  - **Estimativa:** Valor assumido pelo estimador para a amostra que se observou. Representa-se por $T(x_1, x_2, ..., x_n)$ (por exemplo $T(x_1, x_2, ..., x_n) = \bar{x}$). A estimativa é um número, não uma variável aleatória.

- **Dois problemas em aberto**:
  - Como encontrar estimadores para determinado parâmetro?
  - Encontrado um ou mais estimadores, como avaliar a sua qualidade?

Vamos começar por apresentar 2 métodos de estimação e posteriormente abordar-se-ão as propriedades dos estimadores.

Método dos momentos

Ideia: Utilizar os momentos da amostra para estimar os correspondentes momentos da população e, a partir daí, estimar os parâmetros de interesse.

- **Formalização**:
  - $(X_1, X_2, ..., X_n)$ $\quad f(x \mid \theta_1, \theta_2, ..., \theta_k)$ $\quad k$ parâmetros desconhecidos.
  - Momentos de ordem r em relação à origem, da população: $\mu'_r = E(X^r) = \Psi_r(\theta_1, \theta_2, ..., \theta_k)$

    Para que existam estimadores dos momentos têm de existir os correspondentes momentos da população
  - Momentos de ordem r da amostra: $\Sigma_{i=1}^n X_i^r / n$ (os momentos da amostra existem sempre)
  - Constrói-se um sistema de k equações igualando os k primeiros momentos da amostra aos k primeiros momentos do universo:

$$\frac{\sum_{i=1}^n X_i^r}{n} = \Psi_r(\theta_1, \theta_2, ..., \theta_k), \quad r = 1,2,...,k$$

  - Resolve-se o sistema que se admite ter solução única:

    $\tilde{\theta}_j = \phi_j(X_1, X_2, ..., X_n), \quad j = 1,2,...,k.$

    Diz-se então que os estimadores, $\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_k$, foram obtidos pelo método dos momentos.

Exemplos:

1. (exemplo 7.4 do livro) Considere-se uma população de Bernoulli da qual se extraiu uma amostra casual de dimensão n com o objetivo de estimar θ. Como se sabe, $E(X)=\theta$ e consequentemente o estimador pelo método dos momentos é dado por $\tilde{\theta}=\overline{X}$. Em termos formais:

    - 1º momento do universo: $\mu_1' = E(X) = \theta$;
    - 1º momento da amostra: $\dfrac{\sum_{i=1}^{n} X_i}{n}$;
    - Sistema: $\dfrac{\sum_{i=1}^{n} X_i}{n} = \theta$;   (já está resolvido)
    - Solução: Estimador $\rightarrow \tilde{\theta} = \dfrac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$;

    Estimativa $\rightarrow \tilde{\theta} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$ **Cuidado com a notação!**

2. (exemplo 7.5 do livro) Seja $(X_1, X_2, ..., X_n)$ uma amostra casual de uma população normal, $X \sim N(\mu, \sigma^2)$. Suponha-se que se pretende estimar os dois parâmetros μ e σ^2.
   - Momentos do universo:
   - Momentos da amostra:
   - Sistema:
   - Solução:

3. Seja $(X_1, X_2, ..., X_n)$ uma amostra casual de uma população uniforme (-θ, θ) e pretende-se estimar θ.
   - Momentos do universo:
   - Momentos da amostra:
   - Sistema:
   - Solução:

Método da máxima verosimilhança

(mais complicado mas origina, geralmente, estimadores melhores)

Definição 7.1 – Função de verosimilhança

Se $(X_1, X_2, ..., X_n)$ é uma amostra casual de população com função densidade (função probabilidade) $f(x\,|\,\theta)$, a expressão,

$$f(x_1, x_2, \ldots, x_n \,|\, \theta) = \prod_{i=1}^{n} f(x_i \,|\, \theta), \quad (x_1, x_2, ..., x_n) \in \Re^n,$$

define a função densidade (função probabilidade) conjunta das variáveis que constituem a amostra, isto é, designa para **um dado** $\theta \in \Theta$, a densidade (probabilidade) associada com a amostra particular $(x_1, x_2, ..., x_n)$.

Fixado $(x_1, x_2, ..., x_n)$, isto é, observada uma amostra concreta, a **mesma expressão interpretada como função do parâmetro** θ **define a função de verosimilhança** e representa-se por,

$$L(\theta \,|\, x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i \,|\, \theta), \quad \theta \in \Theta,$$

ou mais simplesmente por $L(\theta)$.

Exemplo:

Considere-se novamente um universo com distribuição de Bernoulli de parâmetro θ

$$X \sim Ber(\theta) \qquad f(x\,|\,\theta) = \theta^{x}(1-\theta)^{1-x} \qquad x = 0,1 \qquad 0 < \theta < 1$$

do qual se observou uma particular amostra casual simples de dimensão n, $(x_1, x_2, ..., x_n)$.

Como obter a função de verosimilhança?

Universo: $X \sim Ber(\theta)$, $f(x \mid \theta) = \theta^x (1-\theta)^{1-x}$, $x = 0,1$, $0 < \theta < 1$

Amostra: $(X_1, X_2, ..., X_n)$ observou-se: $(x_1, x_2, ..., x_n)$

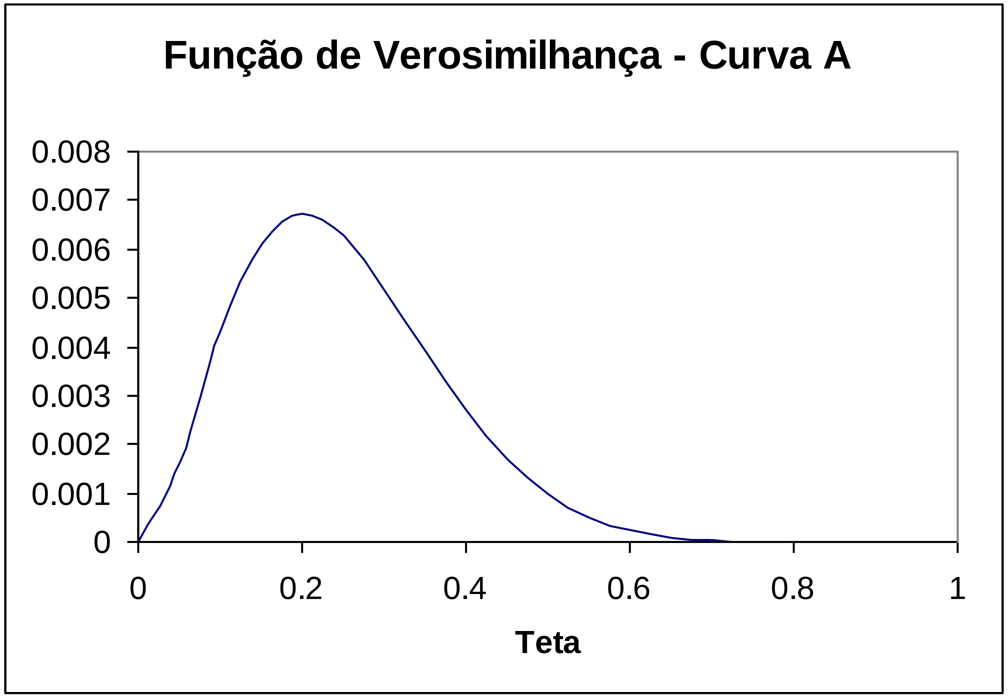
$$L(\theta \mid x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i \mid \theta) = \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i}$$

Como se pode ver, para a função de verosimilhança, não é necessário conhecer cada uma das observações da amostra, bastando conhecer o valor de uma estatística adequada. Neste caso $\sum_{i=1}^{n} x_i$ ou $\bar{x}$.

Vamos admitir que $n = 10$ e traçar a função de verosimilhança para 2 situações:

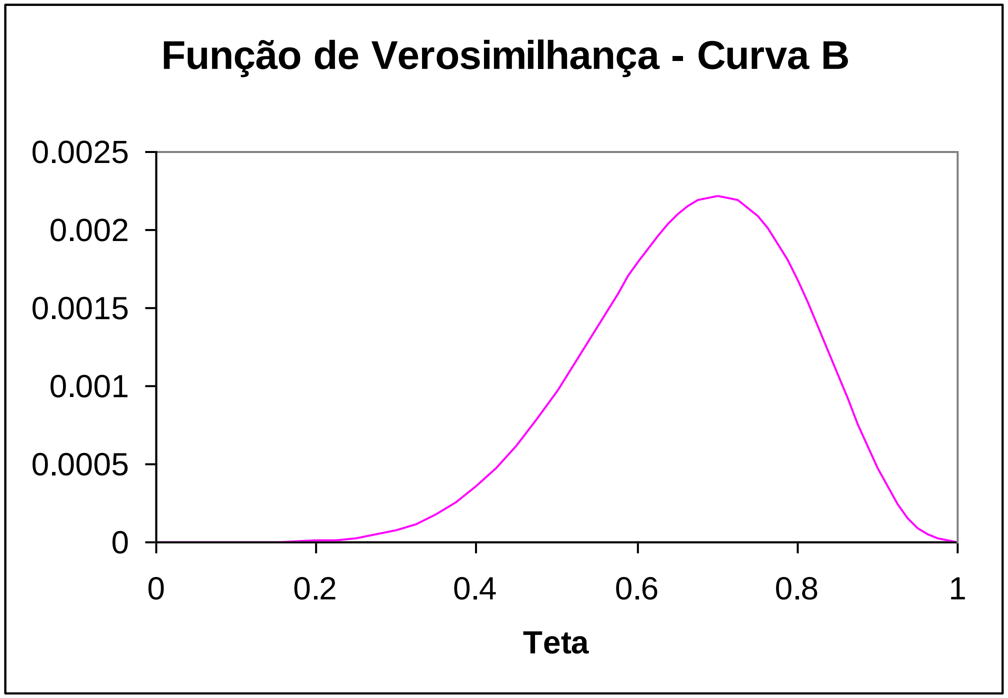
- ○ supondo que se observou $\sum_{i=1}^{10} x_i = 2$ - Curva A: $L(\theta) = \theta^2 (1-\theta)^{10-2}$
- ○ supondo que se observou $\sum_{i=1}^{10} x_i = 7$ - Curva B: $L(\theta) = \theta^7 (1-\theta)^{10-7}$

Primeiro caso, $\sum_{i=1}^{10} x_i = 2$ - Curva A - $L(\theta) = \theta^2 (1-\theta)^8$



Os valores mais "verosímeis" de θ situam-se em torno de 0.2 (note-se que $\bar{x} = 0.2$).

Segundo caso, $\sum_{i=1}^{10} x_i = 7$ - Curva B - $L(\theta) = \theta^7\ (1-\theta)^3$



Os valores mais “verosímeis” de θ situam-se agora em torno de 0.7 (note-se que $\bar{x} = 0.7$).

Método da máxima verosimilhança

- **Intuição**: Observada a amostra, escolhe-se para estimativa do parâmetro o **valor mais verosímil**.

- Assim, dado $(x_1, x_2, \ldots, x_n)$, procura-se uma estimativa, $\hat{\theta} = \hat{\theta}(x_1, x_2, \ldots, x_n)$, tal que,

$$L(\hat{\theta} \mid x_1, x_2, \ldots, x_n) \geq L(\theta \mid x_1, x_2, \ldots, x_n), \quad \forall \theta \in \Theta.$$

- A esta estimativa corresponde o estimador $\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$.

Quando não houver ambiguidade, estimador e estimativa representam-se indistintamente por $\hat{\theta}$.

- Como proceder à maximização da função de verosimilhança?

  - Recorrer aos conhecimentos de Matemática I e II.

  - Geralmente é mais fácil recorrer ao logaritmo da função de verosimilhança, $\ell(\theta) = \ln L(\theta)$. Como a função logaritmo é monótona crescente, $\ell(\theta)$ e $L(\theta)$ têm o mesmo maximizante.

  - A obtenção do maximizante segue, geralmente, os passos habituais:

$$\frac{dL(\theta)}{d\theta} = 0, \quad \frac{d^2L(\theta)}{d\theta^2} < 0 \quad \text{ou} \quad \frac{d\ell(\theta)}{d\theta} = 0, \quad \frac{d^2\ell(\theta)}{d\theta^2} < 0.$$

CUIDADO: O maximizante pode não ser um ponto interior do domínio (exemplo 7.12 livro)

- Muito embora o estimador de máxima verosimilhança seja, na maioria das situações práticas, único, nada garante esta unicidade (exemplo 7.11 livro).

Exemplos

1. (exemplo 7.9 do livro) Retome-se o exemplo anterior referente à Bernoulli. Como se viu

$$L(\theta)=\theta^{\sum_{i=1}^{n}x_i}(1-\theta)^{n-\sum_{i=1}^{n}x_i},\quad 0<\theta<1,$$
$$\ell(\theta)=\left(\sum_{i=1}^{n}x_i\right)\ln\theta+\left(n-\sum_{i=1}^{n}x_i\right)\ln(1-\theta),\quad 0<\theta<1.$$

$$\ell'(\theta)=\frac{d\ell(\theta)}{d\theta}=\frac{\sum_{i=1}^{n}x_i}{\theta}-\frac{\left(n-\sum_{i=1}^{n}x_i\right)}{1-\theta}$$

Da equação $\ell'(\theta)=0$ vem $\hat{\theta}=\dfrac{\sum_{i=1}^{n}x_i}{n}=\bar{x}$ tendo-se o cuidado de verificar que $\ell''(\theta)<0$ para $\theta=\hat{\theta}$.

A esta estimativa de máxima verosimilhança corresponde o estimador de MV

$$\hat{\theta}=\frac{\sum_{i=1}^{n}X_i}{n}$$

Se a amostra observada tivesse sido (0; 1; 0; 0; 0; 0; 1; 0; 0; 0), viria $\hat{\theta}=0.2$ (estimativa).

2. (exemplo 7.10 do livro) $(X_1, X_2, ..., X_n)$ amostra casual de uma população com função densidade $f(x \mid \theta) = \theta x^{\theta - 1}, 0 < x < 1, \theta > 0$.

Pretende-se estimar θ pelo **método dos momentos** e pelo **método da MV**.

Solução do exemplo:

Método dos momentos:

Sabe-se que $E(X)=\int_0^1 x\,\theta\,x^{\theta-1}dx=\dfrac{\theta}{\theta+1}$.

Logo, resolve-se $\overline{X}=\theta/(\theta+1)$, cuja solução é $\tilde{\theta}=\overline{X}/(1-\overline{X})$.

Método da máxima verosimilhança:

$$L(\theta\,|\,x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n} f(x_i\,|\,\theta)$$

$$\ell(\theta\,|\,x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\ln f(x_i\,|\,\theta)=\sum_{i=1}^{n}\ln\left(\theta\,x_i^{\theta-1}\right)=\sum_{i=1}^{n}\left(\ln\theta+(\theta-1)\ln x_i\right)$$

$$\ell'(\theta\,|\,x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\left(\frac{1}{\theta}+\ln x_i\right)=\frac{n}{\theta}+\sum_{i=1}^{n}\ln x_i$$

$$\ell''(\theta\,|\,x_1,x_2,\cdots,x_n)=-n/\theta^2<0$$

A estimativa de MV será então a solução da equação $\ell'(\theta\,|\,x_1,x_2,\cdots,x_n)=0$, $\hat{\theta}=-\dfrac{n}{\sum_{i=1}^{n}\ln x_i}$ a que corresponde o estimador $\hat{\theta}=-\dfrac{n}{\sum_{i=1}^{n}\ln X_i}$.

Como se pode verificar os estimadores dos momentos e de MV são diferentes.

3. **(Exemplo 7.12 que corresponde a uma situação mais invulgar) –** Se $(X_1, X_2, \dots, X_n)$ é uma amostra casual de uma população com função densidade uniforme,

$$f(x \mid \theta) = \begin{cases} \dfrac{1}{\theta} & (0 \le x \le \theta) \\ 0 & (\text{outros} \quad x) \end{cases} \qquad (\theta > 0),$$

tem-se,

$$L(\theta) = \frac{1}{\theta^n}, \quad \theta \ge x_{(n)},$$

já que a desigualdade $\theta \ge x_{(n)}$ $(x_{(n)} = \max x_i)$ é equivalente ao conjunto de desigualdades,

$$0 \le x_1 \le \theta, 0 \le x_2 \le \theta, \dots, 0 \le x_n \le \theta.$$

Assim,

$$\ell(\theta) = -n \ln \theta, \quad \frac{d\ell(\theta)}{d\theta} = -\frac{n}{\theta}, \quad \theta \ge x_{(n)},$$

e a derivada não é igual a zero para nenhum valor finito de θ.
No entanto, a estimativa de máxima verosimilhança deve ser, $\hat{\theta} = x_{(n)}$, já que não pode ser inferior à maior das observações e é errado tomar como estimativa qualquer valor superior, pois,

$$L(x_{(n)}) > L(x_{(n)} + \delta), \quad \forall \delta > 0.$$

Propriedade fundamental dos estimadores de MV: Invariância dos EMV

Se $\hat{\theta}$ é estimador MV para θ e $\tau(\theta)$ é função biunívoca de θ, então $\tau(\hat{\theta})$ é estimador MV de $\tau(\theta)$.

Exemplo – Retome-se o exemplo 7.10 do livro e suponha-se que se pretende uma estimativa não para θ mas sim para $P(X < 0.1)$.

A probabilidade é uma função de θ: $P(X < 0.1) = \int_0^{0.1} \theta x^{\theta-1}\, dx = 0.1^{\theta}$

A função é biunívoca para $\theta > 0$.

Assim $P(X \overset{\wedge}{<} 0.1) = 0.1^{\hat{\theta}}$, pela propriedade da invariância dos EMV

Se, por exemplo, $\hat{\theta} = 2.3$, $P(X \overset{\wedge}{<} 0.1) = 0.1^{2.3} \approx 0.005$ estimativa da MV

Distribuições com *k* parâmetros

Se a função densidade (função probabilidade) envolve *k* parâmetros, a função de verosimilhança escreve-se,

$$L(\theta_1, \theta_2, ..., \theta_k) = \prod_{i=1}^{n} f(x_i \mid \theta_1, \theta_2, ..., \theta_k), \quad (\theta_1, \theta_2, ..., \theta_k) \in \Theta,$$

e as estimativas de máxima verosimilhança são obtidas maximizando a função em ordem aos *k* parâmetros. O processo é em tudo semelhante (nomeadamente logaritmizando) mas, mais complicado.

Propriedades dos estimadores

- **Estimador centrado (ou estimador não enviesado)**

  - **Definição -** Um estimador $T = T(X_1, X_2, ..., X_n)$ para o parâmetro θ diz-se centrado ou não enviesado quando, $E(T) = \theta \quad \forall \theta \in \Theta$.

  - Observações:
    - O valor esperado do estimador deve ser igual ao verdadeiro valor do parâmetro a estimar, *qualquer que seja* θ pertencente ao espaço-parâmetro;
    - O conceito de estimador centrado só se aplica quando existe $E(T)$.

  - **Enviesamento**: Se $E(T) \neq \theta$ o estimador diz-se enviesado e a diferença, $\boxed{\text{Env}(T) = E(T) - \theta}$, mede o seu enviesamento.

  - **Exemplo** – Retome-se o universo de Bernoulli do qual se observou uma amostra casual de dimensão *n*.

    Como se viu $\hat{\theta} = \overline{X}$. Será este estimador centrado?

➢ **Estimador eficiente**

O conceito de estimador centrado não permite distinguir estimadores que apresentem uma distribuição por amostragem fortemente concentrada em torno do parâmetro a ser estimado de outros em que a dispersão é claramente superior. A figura 7.2 ilustra a situação

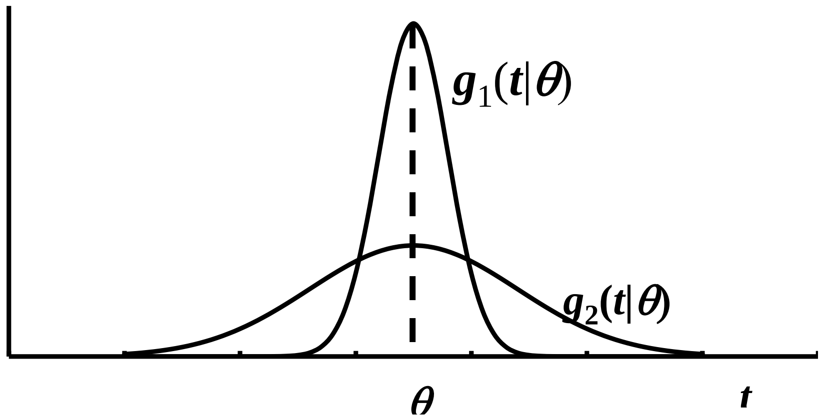


Fig. 7.2 – Comparação de dois estimadores centrados.

- Definição de **Eficiência** – Sejam T e T' dois **estimadores centrados** para θ. O estimador T é **mais eficiente do que T'** quando,

$$\mathrm{Var}(T) \le \mathrm{Var}(T'),\ \forall \theta \in \Theta.$$

O estimador T é **o mais eficiente**, quando a relação se verifica, qualquer que seja o outro estimador, T', centrado para θ.

Observações:

- A eficiência exige a existência de momentos de segunda ordem dos estimadores.
- A definição de eficiência apresenta dois conceitos diferentes:
- O primeiro estabelece uma relação entre dois estimadores centrados para θ, sendo portanto uma eficiência ***relativa***.
- O segundo é um conceito de eficiência ***absoluta*** na classe dos estimadores centrados para θ.
- Para obter estimadores mais eficientes recorre-se à **desigualdade de Fréchet-Cramér-Rao**.

Teorema 7.1 – Seja $(X_1, X_2, \ldots, X_n)$ uma amostra casual de população com função densidade (função probabilidade) $f(x \mid \theta)$, *satisfazendo certas condições de regularidade*, e seja $T = T(X_1, X_2, \ldots, X_n)$ um estimador centrado de θ. Então,

$$\mathrm{Var}(T) \geq \frac{1}{n\Im(\theta)},$$

onde,

$$\Im(\theta) = E\left\{\left[\frac{\partial \ln f(X \mid \theta)}{\partial \theta}\right]^2\right\} = -E\left\{\left[\frac{\partial^2 \ln f(X \mid \theta)}{\partial \theta^2}\right]\right\} \text{ (quantidade de informação de Fisher)}$$

Conhecido o limite inferior dado pela desigualdade de Fréchet-Cramér-Rao, compara-se a variância do estimador centrado em análise com este limite:

- caso sejam iguais → não existe nenhum outro estimador centrado de variância inferior sendo o estimador em análise o mais eficiente;
- caso contrário → o quociente, $\left[n\Im(\theta)\right]^{-1} / \operatorname{Var}(T)$, fornece uma indicação sobre a eficiência relativa do estimador T face ao **hipotético** estimador de variância igual ao limite inferior da desigualdade (nada garante que tal estimador exista).

Exemplo (7.17 do livro) – Seja $f(x\,|\,\theta) = \theta\, x^{\theta-1}$, $0 < x < 1$, $\theta > 0$. A quantidade de informação de Fisher obtém-se determinando,

$$\Im(\theta) = E\left\{\left[\frac{\partial \ln f(X\,|\,\theta)}{\partial\theta}\right]^2\right\} \quad \text{ou} \quad \Im(\theta) = -E\left\{\left[\frac{\partial^2 \ln f(X\,|\,\theta)}{\partial\theta^2}\right]\right\},$$

ficando a escolha dependente da expressão de mais fácil cálculo que é geralmente a segunda.

FAZER …

$$\ln f(X \mid \theta) = \ln\left(\theta X^{\theta-1}\right) = \ln\theta + (\theta-1)\ln X,$$

$$\frac{\partial \ln f(X \mid \theta)}{\partial \theta} = \frac{1}{\theta} + \ln X,$$

$$\frac{\partial^2 \ln f(X \mid \theta)}{\partial \theta^2} = -\frac{1}{\theta^2},$$

obtém-se, $$\Im(\theta) = -E\left[\frac{\partial^2 \ln f(X \mid \theta)}{\partial \theta^2}\right] = \frac{1}{\theta^2}.$$

Para simplificar, no quadro abaixo apresenta-se $\Im(\theta)$ para as distribuições mais utilizadas:

Distribuição	**Quantidade de informação de Fisher**
$X \sim \mathrm{B}(k;\theta)$ (k conhecido)	$\Im(\theta) = k/[\theta(1-\theta)]$
$X \sim \mathrm{Po}(\lambda)$	$\Im(\lambda) = 1/\lambda$
$X \sim \mathrm{BN}(r;\theta)$ (r conhecido)	$\Im(\theta) = r/[\theta^2(1-\theta)]$
$X \sim N(\mu,\sigma^2)$ (σ^2 conhecido)	$\Im(\mu) = 1/\sigma^2$
$X \sim N(\mu,\sigma^2)$ (μ conhecido)	$\Im(\sigma^2) = 1/(2\sigma^4)$
$X \sim G(\alpha,\lambda)$ (α conhecido)	$\Im(\lambda) = \alpha/\lambda^2$

Uma vez que a eficiência está associada ao conceito de estimador centrado, que fazer quando se quer comparar estimadores enviesados?

Erro quadrático médio

Definição: Seja $T = T(X_1, X_2, \ldots, X_n)$ um estimador para o parâmetro θ. O erro quadrático médio de T é dado por, $\mathrm{EQM}(T) = E\left((T-\theta)^2\right)$.

- **Utilização**:
  - Mostra-se sem dificuldade que $EQM(T) = E(T-\theta)^2 = \mathrm{var}(T) + \left(E(T)-\theta\right)^2$.

    O EQM pondera variância e enviesamento. Quando o estimador é centrado, EQM e variância são sinónimos.
  - O estimador T_1 é "melhor" do que T_2 se $\mathrm{EQM}(T_1) \leq \mathrm{EQM}(T_2)$, $\forall \theta \in \Theta$, pois fornece estimativas mais concentradas em torno do verdadeiro valor de θ;
  - O estimador T é "o melhor" estimador se o seu EQM é menor ou igual ao EQM de qualquer outro estimador para θ.

- Como o erro quadrático médio depende, em geral, de θ, deve-se procurar o estimador com EQM uniformemente mínimo, isto é, o estimador T tal que, sendo T' outro estimador qualquer, verifica,

$$E\left((T-\theta)^2\right) \le E\left((T'-\theta)^2\right), \ \forall \theta \in \Theta .$$

O facto da desigualdade se ter de verificar em todo o espaço-parâmetro torna esta propriedade de difícil verificação.

➢ Estimador Consistente

- Na impossibilidade de garantir um estimador com EQM uniformemente mínimo, procura-se que um "bom" estimador tenha EQM que decresça com o aumento da dimensão da amostra, pelo menos a partir de certo valor, isto é, estimadores que sejam consistentes em média quadrática.

- **Definição de consistência em média quadrática** - Um estimador $T_n = T(X_1, X_2, ..., X_n)$ diz-se consistente em média quadrática, se $\lim\limits_{n\to\infty} E\left((T_n-\theta)^2\right)=0, \quad \forall\theta\in\Theta$.

- Condição **necessária e suficiente** para que o estimador T_n seja consistente em média quadrática: $\lim\limits_{n\to+\infty} E(T_n)=\theta$ e $\lim\limits_{n\to+\infty} \mathrm{Var}(T_n)=0$.

- **Definição de estimador consistente** ou simplesmente consistente - Um estimador T_n diz-se (simplesmente) consistente quando, qualquer que seja o número real $\varepsilon>0$,

$$\lim_{n\to+\infty} P_\theta(\theta-\varepsilon<T_n<\theta+\varepsilon)=1, \ \forall\theta\in\Theta.$$

- A **consistência em média quadrática implica a consistência simples**.

- **Exemplo (7.21 do livro) –** Se $(X_1, X_2, ..., X_n)$ é uma sucessão de variáveis aleatórias independentes com distribuição de Bernoulli, $X_i \sim B(1;\theta)$, então o estimador $T_n = \sum_i X_i / n$ será um estimador consistente de θ ?

FAZER …

Propriedades dos estimadores obtidos pelo método dos momentos

- Em condições bastante gerais, são consistentes e possuem distribuição aproximadamente normal quando a dimensão da amostra é muito grande (distribuição assintótica).

Propriedades dos estimadores obtidos pelo método da máxima verosimilhança

- Os estimadores de máxima verosimilhança não são necessariamente centrados.
- Em condições muito gerais eles são consistentes.
- Demonstra-se que, se existir estimador mais eficiente (na óptica do teorema de Fréchet-Cramér-Rao) ele é solução única da equação $dL / d\theta = 0$ e portanto estimador de máxima verosimilhança.
- Verificadas certas condições de regularidade, os estimadores da máxima verosimilhança seguem assintoticamente uma distribuição normal. Caso haja apenas 1 parâmetro desconhecido, tem-se

$$\sqrt{n\Im(\theta)}(\hat{\theta} - \theta) \overset{a}{\sim} N(0,1)$$

ESTIMAÇÃO POR INTERVALOS

Em vez de propor uma estimativa isolada, $\hat{\theta}$, para θ, propõe-se um intervalo (t_1, t_2) a que se associa um "grau de confiança".

Em muitos casos, sobretudo quando θ é um parâmetro de localização, o intervalo é da forma, $\left(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon\right)$, em que o valor de ε pode ser considerado como uma **medida de precisão** ou **medida do erro** inerente à estimativa $\hat{\theta}$.

- **Intervalo aleatório** para θ

  Se $T_1 = T_1(X_1, X_2, ..., X_n)$, $T_2 = T_2(X_1, X_2, ..., X_n)$, $T_1 < T_2$, com

  $$P(T_1 < \theta < T_2) = 1 - \alpha \,, \quad \forall \theta \in \Theta, \quad 0 < \alpha < 1,$$

  (α não depende de θ), então (T_1, T_2) é um **intervalo aleatório** para θ de probabilidade $1 - \alpha$.

- **Intervalo de confiança** para θ

  $(x_1, x_2, ..., x_n) \rightarrow$ Amostra particular - realização de $(X_1, X_2, ..., X_n)$

  $t_1 = T_1(x_1, x_2, ..., x_n)$ e $t_2 = T_2(x_1, x_2, ..., x_n) \rightarrow$ valores assumidos por T_1 e T_2

  A qualquer intervalo (t_1, t_2) que seja concretização do intervalo aleatório (T_1, T_2), chama-se **intervalo de confiança** a $(1 - \alpha)100\%$ para θ

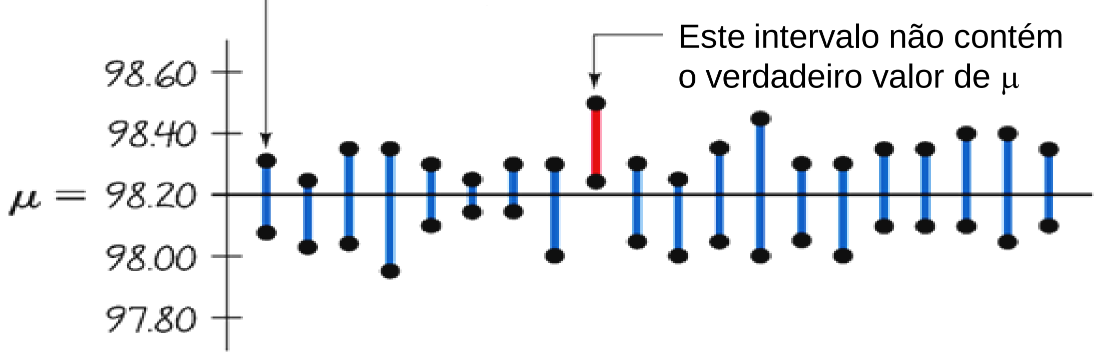
- **Comentários**

  - As definições foram apresentadas para θ e não para $\tau(\theta)$ para simplificar a notação. A generalização é imediata.
  - Um intervalo de confiança mais não é do que uma realização particular de um intervalo aleatório; trata-se de um processo semelhante à passagem de estimador para estimativa na estimação por pontos.
  - **Assim só se atribui probabilidade ao intervalo aleatório**.
  - O conceito, tal como o vimos, é válido para $\Re$. No caso de $\tau(\theta) \in \Re^k$, $k > 1$, é necessário estendê-lo para **regiões de confiança.**

Interpretação frequencista do intervalo de confiança

Seleccionadas várias amostras de idêntica dimensão, da população em estudo, e calculados os correspondentes intervalos de confiança, cerca de 100(1 - α)% dos intervalos calculados contêm o verdadeiro valor do parâmetro θ.

20 amostras $\Rightarrow$ 20 I. C. a 95% para μ



Exemplo: Considere-se uma população que segue uma distribuição $N(\mu, \sigma^2)$ com $\sigma^2 = 605$ da qual se recolheu a seguinte amostra casual simples (807.7, 790.7, 818.8, 853.4, 858.6). Pretende-se construir um intervalo de confiança a 95% para μ.

Como fazê-lo?

$\bar{x} =$ 825.84

....

Resultado … (804.28, 847.40).

Método da variável fulcral

- **Variável fulcral (“pivotal quantity”)**

$Z(X_1, X_2, ..., X_n, \theta)$, **função das observações e de θ** (não depende de mais nenhum valor desconhecido), diz-se uma **variável fulcral** se a sua **função densidade (função probabilidade) $g(z)$ é independente de θ**.

- **Obtenção de um intervalo de confiança**

1. Encontrar uma variável fulcral, Z, adequada ao problema em estudo.

2. Fixado o grau confiança desejado, $1-\alpha$, procurar dois números no domínio de Z, $z_1(\alpha)$ e $z_2(\alpha)$, tais que, $P[z_1(\alpha) < Z < z_2(\alpha)] = 1-\alpha, \quad \forall \theta \in \Theta$.

   Existem muitos (infinitos) pares de valores $(z_1 ; z_2)$ que satisfazem esta equação. A ideia é, para o mesmo grau de confiança, procurar o par de valores que minimiza a amplitude (ou o seu valor esperado se esta for aleatória) para o intervalo resultante.

   Geralmente faz-se $\Pr(Z < z_1(\alpha)) = \Pr(Z > z_2(\alpha)) = \alpha / 2$.

3. Passar de $z_1(\alpha) < Z < z_2(\alpha)$ para uma dupla desigualdade equivalente, $T_1(X_1, X_2,..., X_n) < \theta < T_2(X_1, X_2,..., X_n)$, isto é,

$$P[z_1(\alpha) < Z < z_2(\alpha)] = P[T_1(X_1, X_2,..., X_n) < \theta < T_2(X_1, X_2,..., X_n)] = 1-\alpha$$

$(T_1(X_1, X_2,..., X_n), T_2(X_1, X_2,..., X_n)) = (T_1, T_2)$ é um **intervalo aleatório** de probabilidade $1-\alpha$.

4. Um **intervalo de confiança** a $(1-\alpha)100\%$ para θ é dado por

$$(T_1(x_1, x_2,..., x_n), T_2(x_1, x_2,..., x_n)) = (t_1, t_2)$$

Exemplo anterior:

População segue uma distribuição $N(\mu,\sigma^2)$ mas agora σ^2 desconhecido. Pretende-se igualmente construir um intervalo de confiança a 95% para μ.
(807.7, 790.7, 818.8, 853.4, 858.6)

Qual a VF ?

$\rightarrow$ $\bar{x} = 825.84$ e $s'^2 = 861.553$ IC (789.394, 862.286)

POPULAÇÕES NORMAIS

A) Intervalos de confiança para populações normais: média

- **Variância conhecida** (ver exemplo anterior)
  - Variável fulcral: $Z(X_1, X_2, ..., X_n, \mu) = \dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$
  - Intervalo de confiança $\left(\bar{x} - z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}} , \bar{x} + z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}} \right)$, em que $z_{\alpha/2}$ verifica $\Phi(z_{\alpha/2}) = 1 - \dfrac{\alpha}{2}$.
  - Amplitude do intervalo (precisão da estimativa): $2 z_{\alpha/2}\, \sigma / \sqrt{n}$.

- **Variância σ^2 desconhecida (caso mais corrente)**
  - Variável fulcral: $T = \dfrac{\bar{X} - \mu}{S' / \sqrt{n}} \sim t(n-1)$
  - Intervalo de confiança: $\left(\bar{x} - t_{\alpha/2} \dfrac{s'}{\sqrt{n}} , \; \bar{x} + t_{\alpha/2} \dfrac{s'}{\sqrt{n}} \right)$
  - $t_{\alpha/2}$ verifica $P(T > t_{\alpha/2}) = \alpha / 2$ (minimizou-se o valor esperado da amplitude)

Exemplo – Retome-se o exemplo anterior (com $\sigma^2 = 605$) e determine-se n para que a amplitude do intervalo de confiança seja inferior a 20.

Solução:

Pretende-se o menor valor de n (inteiro) tal que,

$$2\, z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < 20 \Leftrightarrow n > \left(\frac{2\, z_{\alpha/2}\, \sigma}{20} \right)^2 \Leftrightarrow n > 23.24,$$

escolhendo-se portanto $n = 24$.

B) Intervalos de confiança para populações normais: variância

- Variável fulcral: $Q = \dfrac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$.
- Intervalo de confiança: $\left(\dfrac{(n-1)s'^2}{q_2} , \dfrac{(n-1)s'^2}{q_1} \right)$
- q_1 e q_2 tais que $P(Q < q_1) = P(Q > q_2) = \alpha / 2$. Utilizou-se um critério simplificador ao considerar probabilidades iguais nas 2 caudas já que a minimização da amplitude esperada envolve uma otimização "caso a caso". Esta solução fornece geralmente uma aproximação adequada.
- Amplitude do intervalo: $(n-1)s'^2 \left(\dfrac{1}{q_1} - \dfrac{1}{q_2} \right)$

Exemplo – Retome-se, uma vez mais, o exemplo anterior supondo agora que σ^2 é desconhecido. Determine-se um intervalo de confiança a 90% para σ^2.

amostra casual simples (807.7, 790.7, 818.8, 853.4, 858.6)

Resolução

Com $\alpha = 0.1$, $\alpha/2 = 0.05$ e $n-1=4$ graus de liberdade, tem-se,

$$q_1 = 0.711 \text{ e } q_2 = 9.488,$$

Por outro lado $s'^2 = 861.553$

Então

$$\frac{(n-1)\,s'^2}{q_2} = \frac{4 \times 861.553}{9.488} = 363.22, \quad \frac{(n-1)\,s'^2}{q_1} = \frac{4 \times 861.553}{0.711} = 4846.99,$$

obtendo-se assim como intervalo de confiança, (363.22, 4846.99).

Quando a amostra é pequena a estimação da variância é sempre problemática

C1) Intervalos de confiança para populações normais: diferença entre médias $\mu_1 - \mu_2$

- 2 populações normais, $N(\mu_1, \sigma_1^2)$ e $N(\mu_2, \sigma_2^2)$
- duas amostras casuais independentes:
  - uma de dimensão *m*, $(X_{11}, X_{12}, ..., X_{1m})$, da primeira população;
  - outra, de dimensão *n*, $(X_{21}, X_{22}, ..., X_{2n})$, da segunda população.
- Construção de intervalos de confiança para a diferença entre as médias das duas populações, $\mu_1 - \mu_2$;

É necessário distinguir 2 situações:

- **Variâncias conhecidas**

V. F.

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)$$

IC a 100(1-α)% para $\mu_1 - \mu_2$?

I. C.

$$\left((\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}, (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} \right)$$

- **Variâncias desconhecidas, mas iguais**

V F.

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{m} + \frac{1}{n}} \sqrt{\frac{(m-1)S_1'^2 + (n-1)S_2'^2}{m+n-2}}} \sim t(m+n-2).$$

IC a 100(1-α)% para $\mu_1 - \mu_2$?

- **Variâncias desconhecidas** (e possivelmente diferentes)

V F.

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1'^2}{m} + \frac{S_2'^2}{n}}} \overset{a}{\sim} t(r),$$

sendo r dado pelo maior inteiro contido em $\dfrac{\left(\frac{s_1'^2}{m} + \frac{s_2'^2}{n}\right)^2}{\frac{1}{m-1}\left(\frac{s_1'^2}{m}\right)^2 + \frac{1}{n-1}\left(\frac{s_2'^2}{n}\right)^2}$.

Neste caso apenas se obtém um resultado aproximado.

IC a 100(1-α)% para $\mu_1 - \mu_2$?

Exemplo (7.34 do livro) – Para testar, em condições tradicionais de cultura, dois novos tipos de milho, uma exploração seleccionou aleatoriamente 8 zonas de terreno mais ou menos homogéneo e em cada uma delas plantou sementes de ambos os tipos em talhões separados.

Os resultados da colheita (alqueires por hectare), são dados por:

$\bar{x}_1 = 81.63,\ \bar{x}_2 = 75.88,\ s_1'^2 = 232.4107,\ s_2'^2 = 102.1250.$

Supondo que as produções (alqueires por hectare) têm distribuição normal e que não há razão para admitir que os desvios padrões são diferentes, isto é, supondo $\sigma_1 = \sigma_2$, pretende construir-se um intervalo de confiança a 95% para a diferença $\mu_1 - \mu_2$.

Solução:
Com $\alpha = 0.05$ e $16 - 2 = 14$ graus de liberdade, a Tabela 7 mostra que $t_{0.025} = 2.145$.

Assim,

$$81.63 - 75.88 \pm 2.145\left(\sqrt{\frac{1}{8}+\frac{1}{8}}\right)\left(\sqrt{\frac{7(232.4107)+7(102.1250)}{8+8-2}}\right) = 5.75 \pm 13.8709,$$

conduz ao intervalo de confiança $(-8.12, 19.62)$.

Como o valor 0 pertence ao intervalo de confiança, não se pode concluir que as produções médias das populações são diferentes.

C2) Intervalos de confiança para populações normais: relação entre variâncias σ_2^2 / σ_1^2

V. F. $$F = \frac{S_1'^2}{S_2'^2} \frac{\sigma_2^2}{\sigma_1^2} \sim F(m-1, n-1)$$

IC a 100(1-α)% para σ_2^2 / σ_1^2 ?

I. C. $$\left(f_1 \frac{s_2'^2}{s_1'^2} \, , \, f_2 \frac{s_2'^2}{s_1'^2} \right)$$

f_1 e f_2 a verificar $P(F < f_1) = P(F > f_2) = \alpha / 2$. Utilizando as tabelas, f_2 é obtido "directamente" e para obter f_1 fazer:

- Ir à tabela da $F(n-1, m-1)$ e procurar o valor f que tem uma probabilidade à direita de $\alpha / 2$;
- Fazer $f_1 = 1/ f$

Exemplo – Retome-se o exemplo anterior e calcule-se um intervalo de confiança a 95% para σ_2^2 / σ_1^2.

Solução

Como a distribuição *F*-Snedcor envolvida é a $F(7,7)$, a tabela 8 dá imediatamente $f_2 \approx 4.99$.

Para obter f_1 tem-se:

- Trocar os graus de liberdade da F (sem efeito prático já que $m = n$) e ir à tabela → 4.99
- $f_1 = 1/4.99 \approx 0.2$.

O intervalo de confiança a 90% é, portanto,

$$\left(0.2\frac{s_2'^2}{s_1'^2},\ 4.99\frac{s_2'^2}{s_1'^2}\right) = \left(0.2\times\frac{102.1250}{232.4107},\ 4.99\times\frac{102.1250}{232.4107}\right) = (0.0879,\ 2.1927).$$

GRANDES AMOSTRAS

D) Intervalos de confiança para grandes amostras

Ideia fundamental: Recorrer ao **teorema do limite central** e obter intervalos assintóticos, isto é, aproximados e **com validade quando *n* é grande**.

Parâmetro: μ
(população qualquer com variância finita).

VF: $Z = \dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}} \overset{a}{\sim} N(0,1)$ **IC**: $\left(\bar{x} - z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}} ; \bar{x} + z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}} \right)$ com $z_{\alpha/2}$: $\Phi(z_{\alpha/2}) = 1 - \alpha / 2$

Quando σ é **desconhecido**, pode-se utilizar $\hat{\sigma}$, estimador consistente de σ

VF: $Z = \dfrac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{n}} \overset{a}{\sim} N(0,1)$ **IC**: $\left(\bar{x} - z_{\alpha/2} \dfrac{\hat{\sigma}}{\sqrt{n}} ; \bar{x} + z_{\alpha/2} \dfrac{\hat{\sigma}}{\sqrt{n}} \right)$

No caso geral, pode-se utilizar $\hat{\sigma} = S'$

VF: $Z = \dfrac{\bar{X} - \mu}{S' / \sqrt{n}} \overset{a}{\sim} N(0,1)$ **IC**: $\left(\bar{x} - z_{\alpha/2} \dfrac{s'}{\sqrt{n}} ; \bar{x} + z_{\alpha/2} \dfrac{s'}{\sqrt{n}} \right)$

- **Parâmetro:** $\mu_1 - \mu_2$
  (populações quaisquer independentes com variâncias finitas)

V. F.

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \overset{a}{\sim} N(0,1) \qquad \text{ou} \qquad Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1'^2}{m} + \frac{S_2'^2}{n}}} \overset{a}{\sim} N(0,1).$$

I. C.

$$(\bar{x}_1 - \bar{x}_2 - z_{\alpha/2}\sigma^*, \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}\sigma^*) \qquad \text{ou} \qquad (\bar{x}_1 - \bar{x}_2 - z_{\alpha/2}s^*, \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}s^*),$$

com $\sigma^* = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$ $\qquad$ $s^* = \sqrt{\frac{s_1'^2}{m} + \frac{s_2'^2}{n}}$

e $z_{\alpha/2}$ a verificar $\Phi(z_{\alpha/2}) = 1 - \alpha/2$.

Populações de Bernoulli

- **Parâmetro:** θ

A utilização de $\dfrac{\bar{X}-\theta}{\sqrt{\dfrac{\theta(1-\theta)}{n}}} \overset{a}{\sim} N(0,1)$, como variável fulcral é possível mas complicada e origina resultados (p. 369 livro) semelhantes a utilizar

$$Z = \frac{\bar{X}-\theta}{\sqrt{\dfrac{\bar{X}(1-\bar{X})}{n}}} \overset{a}{\sim} N(0,1)$$

I.C. $\left(\bar{x} - z_{\alpha/2}\sqrt{\dfrac{\bar{x}(1-\bar{x})}{n}}\,;\, \bar{x} + z_{\alpha/2}\sqrt{\dfrac{\bar{x}(1-\bar{x})}{n}} \right)$ com $z_{\alpha/2}$ a verificar $\Phi(z_{\alpha/2}) = 1-\alpha/2$.

Exemplo 1 (7.37 do livro) – Um candidato a deputado por determinado círculo uninominal teve conhecimento de que, numa sondagem feita junto de $n = 350$ eleitores, havia 185 que declararam dar-lhe o seu voto. Será que a proporção observada de eleitores favoráveis, $\bar{x} = 185/350 \approx 0.5286$, pode "assegurar" a eleição por maioria absoluta do candidato?

Resolução

Calcular o IC para o parâmetro θ, proporção dos apoiantes do candidato na população.

Escolher um nível de confiança, por exemplo 95%

$$\alpha = 0.05, \quad 1-\Phi(z_{0.025}) = 0.025, \quad z_{0.025} = 1.96.$$

O intervalo vem

$$0.5286 \pm 1.96\sqrt{\frac{0.5286 \times 0.4714}{350}} \Rightarrow (0.4927,\ 0.5645),$$

o que aponta para a possibilidade do candidato poder falhar a maioria absoluta.

Exemplo 2 (7.38 do livro) – Suponha-se que uma empresa pretende lançar um novo produto no mercado. O respectivo estudo de mercado tem por principal objectivo determinar, em dada população (que se supõe muito numerosa), a proporção θ de potenciais clientes para o novo produto. A questão que se põe é a seguinte: qual deve ser a dimensão da amostra casual para que, com uma confiança de 95%, o erro cometido seja, em valor absoluto, inferior a 3%?

Resolução

Pretende-se determinar o menor n (inteiro) tal que, $P(|\bar{X}-\theta|<0.03)\geq 0.95$, sendo $X_i \sim B(1;\theta)$

$$P\left(\frac{-0.03}{\sqrt{\theta(1-\theta)/n}}<\frac{\bar{X}-\theta}{\sqrt{\theta(1-\theta)/n}}<\frac{0.03}{\sqrt{\theta(1-\theta)/n}}\right)\approx 2\Phi\left(\frac{0.03}{\sqrt{\theta(1-\theta)/n}}\right)-1\geq 0.95,$$

$$\Phi\left(\frac{0.03}{\sqrt{\theta(1-\theta)/n}}\right)\geq 0.975.$$

Logo,

$$\frac{0.03}{\sqrt{\theta(1-\theta)/n}}\geq 1.96 \Rightarrow n\geq\left(\frac{1.96}{0.03}\right)^2\theta(1-\theta).$$

Embora θ seja desconhecido, sabe-se que $\theta(1-\theta)\leq 1/4$ (resultado que se deduz sem dificuldade, notando que $\theta^* = 1/2$ é maximizante de $\theta(1-\theta)$ sujeito à restrição $0<\theta<1$).

Adoptando a situação mais desfavorável, pode propor-se $n=1067$, uma vez que nesta situação tem-se $n\geq(1.96/0.03)^2(1/4)$.

Dimensionamento da amostra

Qual deverá ser a dimensão da amostra para que, com uma confiança de 100(1-α)%, se garanta um erro máximo de ε?

$$n : P(|\bar{X}-\mu| < \varepsilon) \geq 1-\alpha$$

Com n grande, aplicar o TLC $\quad Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$

$$n : P\left(\frac{|\bar{X}-\mu|}{\sigma/\sqrt{n}} < \frac{\varepsilon}{\sigma/\sqrt{n}}\right) \geq 1-\alpha \Rightarrow \frac{\varepsilon}{\sigma/\sqrt{n}} \geq z_{\alpha/2}$$

$n \geq \dfrac{z_{\alpha/2}^2\sigma^2}{\varepsilon^2}$ com $z_{\alpha/2}$ a verificar $\Phi(z_{\alpha/2}) = 1-\alpha/2$

será necessário **estimar a variância da população**, caso não seja conhecida.

Alternativamente pode utilizar-se o IC a 100(1-α)% para μ $\left(\bar{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right)$ sendo

$$z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \rightarrow erro$$

Se população de Bernoulli $n \geq \dfrac{z_{\alpha/2}^2\theta(1-\theta)}{\varepsilon^2}$ e como $\theta(1-\theta) \leq 1/4$ então utiliza-se $n \geq \dfrac{z_{\alpha/2}^2}{4e^2}$

Populações de Bernoulli (continuação)

Parâmetro: $\theta_1 - \theta_2$ (populações de Bernoulli).

V. F.

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\bar{X}_1(1-\bar{X}_1)}{m} + \frac{\bar{X}_2(1-\bar{X}_2)}{n}}} \overset{a}{\sim} N(0,1).$$

I. C.

$$(\bar{x}_1 - \bar{x}_2 - z_{\alpha/2}s^*, \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}s^*),$$

com $s^* = \sqrt{\frac{\bar{x}_1(1-\bar{x}_1)}{m} + \frac{\bar{x}_2(1-\bar{x}_2)}{n}}$ e $z_{\alpha/2}$ a verificar $\Phi(z_{\alpha/2}) = 1 - \alpha/2$.

Outras Populações

- **População de Poisson:** IC para λ.

VF: $Z = \dfrac{\overline{X} - \lambda}{\sqrt{\dfrac{\overline{X}}{n}}} \overset{a}{\sim} N(0,1)$

IC: $\left(\overline{x} - z_{\alpha/2}\sqrt{\dfrac{\overline{x}}{n}} ; \overline{x} - z_{\alpha/2}\sqrt{\dfrac{\overline{x}}{n}} \right)$

- **População exponencial: IC para o** λ mesmo em amostras de pequena dimensão?

Pequena dimensão

VF: $Q = 2n\lambda\overline{X} \sim \chi^2_{(2n)}$

IC: $\left(\dfrac{q_1}{2n\overline{x}} ; \dfrac{q_2}{2n\overline{x}} \right)$ **com** $P(Q < q_1) = P(Q > q_2) = \alpha / 2$

- **Utilização da distribuição assintótica dos estimadores da m.v. na construção de IC** …

$\sqrt{n\Im(\theta)}(\hat{\theta} - \theta) \overset{a}{\sim} N(0,1)$ ou $\sqrt{n\Im(\hat{\theta})}(\hat{\theta} - \theta) \overset{a}{\sim} N(0,1)$