

7

Distribuições Amostrais e Estimação Pontual de Parâmetros

ESQUEMA DO CAPÍTULO

7.1 INTRODUÇÃO

7.2 DISTRIBUIÇÕES AMOSTRAIS E TEOREMA DO LIMITE CENTRAL

7.3 CONCEITOS GERAIS DE ESTIMAÇÃO PONTUAL

7.3.1 Estimadores não-tendenciosos

7.3.2 Variância de um Estimador Pontual

7.3.3 Erro-Padrão: Reportando uma Estimador

7.3.5 Erro Quadrático Médio de um Estimador

7.4 MÉTODOS DE ESTIMAÇÃO PONTUAL

7.4.1 Método dos Momentos

7.4.2 Método da Máxima Verossimilhança

7.4.3 Estimação Bayesiana de Parâmetros

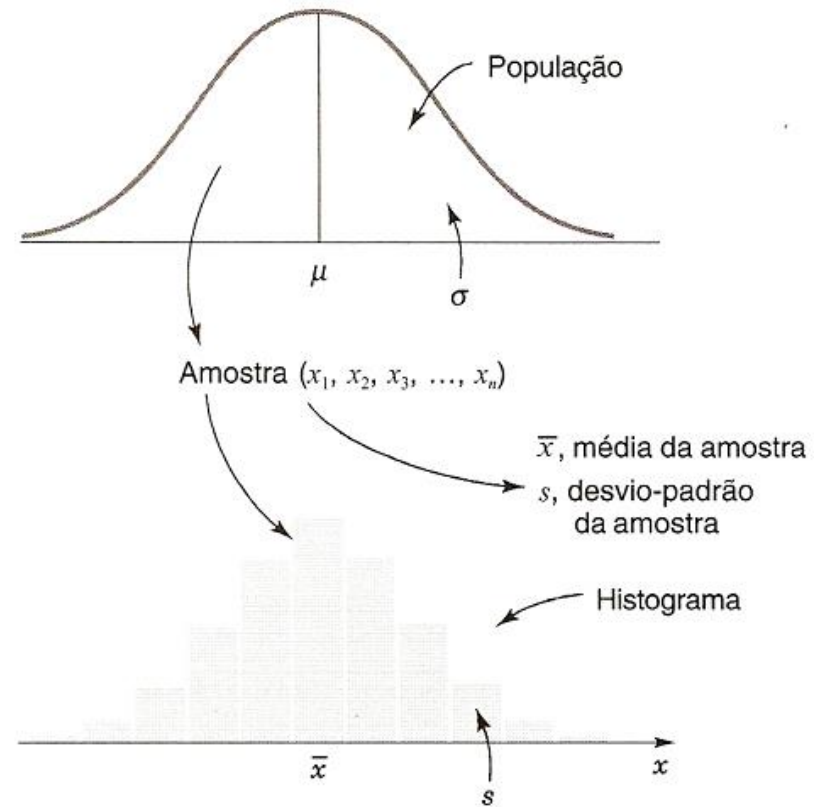
Objetivos de Aprendizagem

Após estudo cuidadoso deste capítulo você deverá ser capaz de:

1. Explicar os conceitos gerais da estimação de parâmetros de uma população ou distribuição de probabilidade;
2. Explicar o importante papel da distribuição normal com uma distribuição amostral;
3. Compreender o Teorema Central do Limite;
4. Explicar propriedades importantes dos estimadores pontuais, inclusive tendência, variância e erro quadrático médio;
5. Saber como construir estimadores pontuais pelo método do momentos e pelo método da máxima verossimilhança;
6. Saber como calcular e explicar a precisão com a qual um parâmetro é estimado;

7.1 Introdução

- O campo da inferência estatística consiste nos métodos utilizados para tomar decisões ou tirar conclusões sobre uma **população**;
- Tais métodos utilizam a informação contida em uma **amostra** da população para tirar conclusões (ver figura);



7.1 Introdução

- A inferência estatística pode ser dividida em **duas** grandes áreas:
 - **Estimação de parâmetros - Exemplo:** Um engenheiro está interessado na estimação da resistência média à tração dos componentes usados em um chassi de automóvel e usará dados de uma amostra para calcular uma estimativa da média verdadeira;
 - **Testes de hipóteses – Exemplo:** Um engenheiro conjectura que duas temperaturas diferentes de reação, t_1 e t_2 , resultam em rendimentos diferentes. A hipótese de interesse, a ser testada, seria que o rendimento médio usando a temperatura t_1 é maior do que o rendimento médio usando a temperatura t_2 . Note que não há ênfase na estimação dos rendimentos.

7.1 Introdução

- Definições:

Definição

Uma **população** consiste na totalidade das observações em que estamos interessados.

Definição

Uma **amostra** é um subconjunto de observações selecionadas a partir de uma população.

7.1 Introdução

- Definições:

Definição

As variáveis aleatórias (X_1, X_2, \dots, X_n) são uma amostra aleatória de tamanho n , se (a) os X_i 's forem variáveis aleatórias independentes e (b) cada X_i tiver a mesma distribuição de probabilidades.

Definição

Uma **estatística** é qualquer função das observações em uma amostra aleatória.

Definição

Uma **estimativa** pontual de algum parâmetro θ da população é um único valor numérico $\hat{\theta}$ de uma estatística $\hat{\Theta}$.

7.1 Introdução

- Exemplo:

Como um exemplo, suponha que a variável aleatória X seja normalmente distribuída com uma média desconhecida μ . A média da amostra é um estimador da média desconhecida μ da população. Isto é, $\hat{\mu} = \bar{X}$. Depois da amostra ter sido selecionada, o valor numérico \bar{x} é a estimativa de μ . Assim, se $x_1 = 25$, $x_2 = 30$, $x_3 = 29$ e $x_4 = 31$, então a estimativa de μ é

$$\bar{x} = \frac{25 + 30 + 29 + 31}{4} = 28,75$$

7.1 Introdução

- Problemas de *estimação* ocorrem frequentemente em engenharia;
- Geralmente necessitamos *estimar*:
 - A *média* μ de uma única população;
 - A *variância* σ^2 (ou desvio-padrão σ) de uma única população;
 - A *proporção* p de itens em uma população que pertence a uma classe de interesse;
 - A *diferença nas médias de duas populações*, $\mu_1 - \mu_2$;
 - A *diferença nas proporções de duas populações*, $p_1 - p_2$.

7.2 Distribuições Amostrais e Teorema do Limite Central

- Estimativas *razoáveis* desses parâmetros são:
 - Para a *média populacional*, a estimativa é a média aritmética da amostra, $\hat{\mu} = \bar{x}$;
 - Para a *variância populacional*, a estimativa é a variância da amostra, $\hat{\sigma}^2 = s^2$
 - Para a *proporção populacional*, a estimativa é a proporção da amostra, $\hat{p} = x/n$
 - Para a *diferença de duas médias populacionais*, a estimativa é a diferença entre as médias de duas amostras aleatórias independentes, $\hat{\mu}_1 - \hat{\mu}_2 = \bar{x}_1 - \bar{x}_2$;
 - Para a *diferença de duas proporções populacionais*, a estimativa é a diferença entre duas proporções amostrais, calculadas a partir de duas amostras aleatórias independentes, $\hat{p}_1 - \hat{p}_2$.

7.2 Distribuições Amostrais e Teorema do Limite Central

- O Teorema Central do Limite:

Teorema 7.2: O Teorema Central do Limite

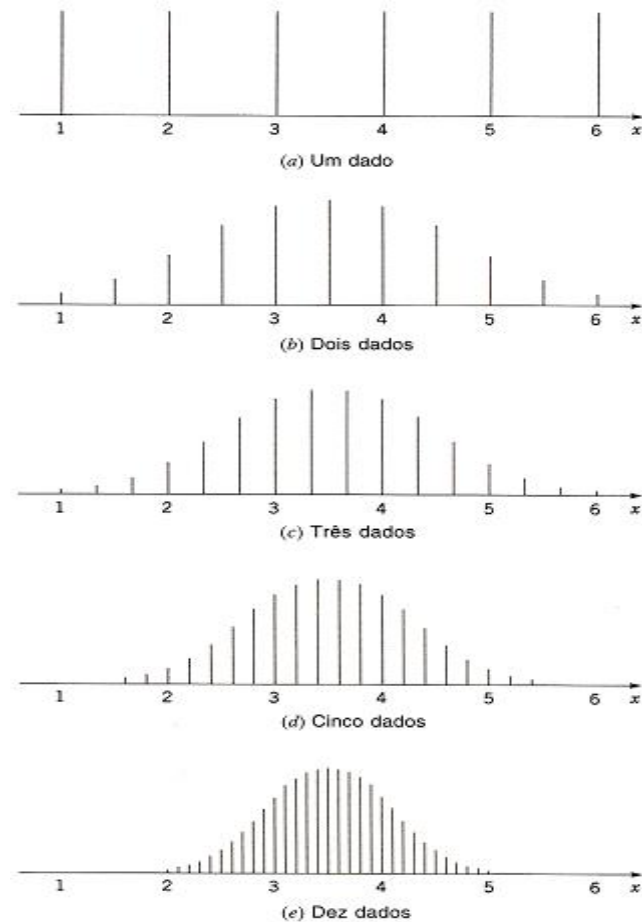
Se X_1, X_2, \dots, X_n for uma amostra aleatória de tamanho n , retirada de uma população (finita ou infinita), com média μ e variância finita σ^2 , e se \bar{X} for a média da amostra, então a forma limite da distribuição de

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (7.7)$$

quando $n \rightarrow \infty$ é a distribuição normal padrão.

7.2 Distribuições Amostrais e Teorema do Limite Central

Fig. 7.6
Distribuições das pontuações médias obtidas quando arremessamos dados.



7.2 Distribuições Amostrais e Teorema do Limite Central

- **Exemplo:** Uma companhia eletrônica fabrica resistores que têm uma resistência média de 100Ω e um desvio-padrão de 10Ω . A distribuição de resistências é normal. Encontre a probabilidade de uma amostra aleatória de $n = 25$ resistores ter uma resistência média menor que 95Ω .

Note que a distribuição amostral \bar{X} de é normal, com média $\mu_{\bar{X}} = 100 \Omega$ e um desvio-padrão de

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$$

Conseqüentemente, a probabilidade desejada corresponde à área sombreada na Fig. 7.7. Padronizando o ponto $\bar{X} = 95$ na Fig. 7.7, encontramos que

$$z = \frac{95 - 100}{2} = -2,5$$

e desse modo,

$$\begin{aligned} P(\bar{X} < 95) &= P(Z < -2,5) \\ &= 0,0062 \end{aligned}$$

7.2 Distribuições Amostrais e Teorema do Limite Central

- Exemplo:

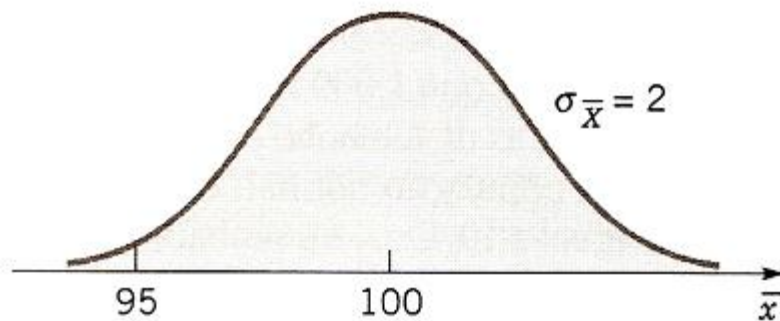


Fig. 7.7 Probabilidade para o exemplo.

Uma companhia eletrônica fabrica resistores que têm uma resistência média de 100Ω e um desvio-padrão de 10Ω . A distribuição de resistências é normal. Encontre a probabilidade de uma amostra aleatória de $n = 25$ resistores ter uma resistência média menor que 95Ω .

Note que a distribuição amostral \bar{X} de X é normal, com média $\mu_{\bar{X}} = 100 \Omega$ e um desvio-padrão de

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$$

Conseqüentemente, a probabilidade desejada corresponde à área sombreada na Fig. 7.7. Padronizando o ponto $\bar{X} = 95$ na Fig. 7.7, encontramos que

$$z = \frac{95 - 100}{2} = -2,5$$

e desse modo,

$$\begin{aligned} P(\bar{X} < 95) &= P(Z < -2,5) \\ &= 0,0062 \end{aligned}$$

7.2 Distribuições Amostrais e Teorema do Limite Central

- Exemplo II:

Suponha que uma variável aleatória X tenha uma distribuição contínua uniforme

$$f(x) = \begin{cases} 1/2, & 4 \leq x \leq 6 \\ 0, & \text{caso contrário} \end{cases}$$

Encontre a distribuição amostral da média de uma amostra aleatória de tamanho $n = 40$.

A média e a variância de X são $\mu = 5$ e $\sigma^2 = (6 - 4)^2/12 = 1/3$. O teorema central do limite indica que a distribuição de \bar{X} é aproximadamente normal, com média $\mu_{\bar{X}} = 5$ e variância $\sigma_{\bar{X}}^2 = \sigma^2/n = 1/[3(40)] = 1/120$. As distribuições de X e \bar{X} são mostradas na Fig. 7.8.

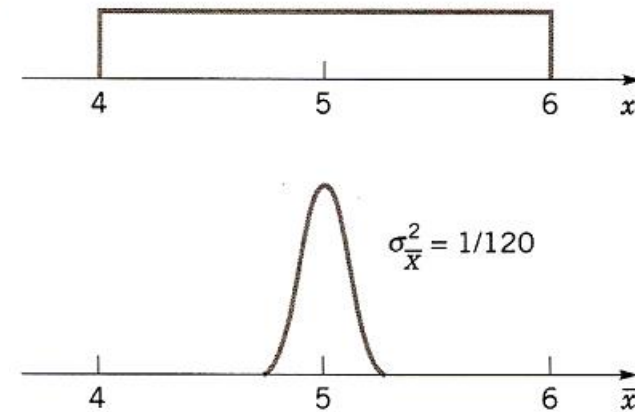


Fig. 7.8 As distribuições de X e \bar{X} , para o exemplo.

7.3 Conceitos Gerais de Estimação Pontual

7.3.1 Estimadores Não Tendenciosos

- Definição:

Definição

O estimador $\hat{\Theta}$ é um **estimador não tendencioso** para o parâmetro θ , se

$$E(\hat{\Theta}) = \theta \quad (7.1)$$

Se o estimador for tendencioso, então a diferença

$$E(\hat{\Theta}) - \theta \quad (7.2)$$

é chamada de **tendência** do estimador $\hat{\Theta}$.

7.3 Conceitos Gerais de Estimação Pontual

7.3.1 Estimadores Não Tendenciosos

- Exemplo:

Suponha que X seja uma variável aleatória com média μ e variância σ^2 . Faça X_1, X_2, \dots, X_n ser uma amostra aleatória de tamanho n , de uma população representada por X . Mostre que a média da amostra \bar{X} e a variância da amostra S^2 são estimadores não tendenciosos de μ e σ^2 , respectivamente.

Considere primeiro a média da amostra. Na Eq. 5.28, no Cap. 5, mostramos que $E(\bar{X}) = \mu$. Conseqüentemente, a média da amostra \bar{X} é um estimador não tendencioso da média da população.

7.3 Conceitos Gerais de Estimaco Pontual

7.3.1 Estimadores No Tendenciosos

- Exemplo (continuao):

Considere agora a varincia da amostra. Temos

$$\begin{aligned} E(S^2) &= E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right] = \frac{1}{n-1} E\sum_{i=1}^n (X_i - \bar{X})^2 \\ &= \frac{1}{n-1} E\sum_{i=1}^n (X_i^2 + \bar{X}^2 - 2\bar{X}X_i) \\ &= \frac{1}{n-1} E\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n E(X_i^2) - nE(\bar{X}^2)\right] \end{aligned}$$

7.3 Conceitos Gerais de Estimação Pontual

7.3.1 Estimadores Não Tendenciosos

- Exemplo (final):

A última igualdade vem da Eq. 5.25 no Cap. 5. Entretanto, uma vez que $E(X_1^2) = \mu^2 + \sigma^2$ e $E(\bar{X}^2) = \mu^2 + \sigma^2/n$, temos

$$\begin{aligned} E(S^2) &= \frac{1}{n-1} \left[\sum_{i=1}^n (\mu^2 + \sigma^2) - n(\mu^2 + \sigma^2/n) \right] \\ &= \frac{1}{n-1} (n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2) \\ &= \sigma^2 \end{aligned}$$

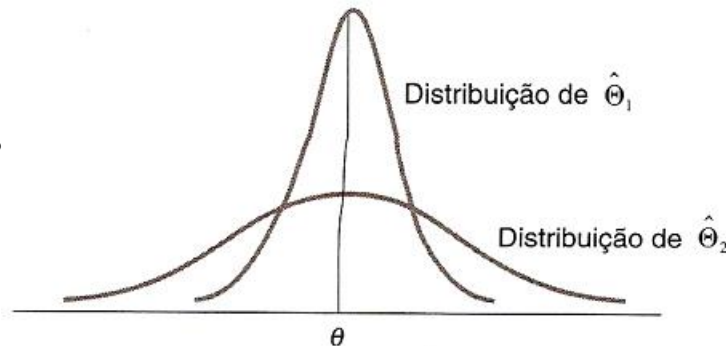
Logo, a variância da amostra S^2 é um estimador não tendencioso da variância σ^2 da população. No entanto, podemos mostrar que o desvio-padrão da amostra é um estimador tendencioso do desvio-padrão da população. Para amostras grandes, essa tendência é negligenciável.

7.3 Conceitos Gerais de Estimação Pontual

7.3.2 Variância de um Estimador Pontual

- Como não há um estimador não tendencioso **único**, não podemos confiar **apenas** na propriedade de não tendenciosidade para selecionar um estimador (ver figura);

Fig. 7.2 As distribuições amostrais de dois estimadores não tendenciosos $\hat{\Theta}_1$ e $\hat{\Theta}_2$.



- **Definição:**

Definição

Se considerarmos todos os estimadores não tendenciosos de θ , aquele com a menor variância será chamado de **estimador não tendencioso de variância mínima (ENTVM)**.

7.3 Conceitos Gerais de Estimação Pontual

7.3.2 Variância de um Estimador Pontual (cont.)

- Teorema:

Teorema 7.1

Se X_1, X_2, \dots, X_n for uma amostra aleatória de tamanho n , proveniente de uma distribuição normal com média μ e variância σ^2 , então a média da amostra, \bar{X} , será o ENTVM para μ .

7.3 Conceitos Gerais de Estimação Pontual

7.3.3 Erro-Padrão: Reportando uma Estimativa Pontual

- Definição:

Definição

O **erro-padrão** de um estimador $\hat{\Theta}$ é o seu desvio-padrão, dado por $\hat{\sigma}_{\hat{\Theta}} = \sqrt{V(\hat{\Theta})}$. Se o erro-padrão envolver parâmetros desconhecidos que possam ser estimados, então a substituição daqueles valores em $\hat{\sigma}_{\hat{\Theta}}$ produzirá um **erro-padrão estimado**, denotado por $\hat{\sigma}_{\hat{\Theta}}$.

7.3 Conceitos Gerais de Estimativa Pontual

7.3.3 Erro-Padrão: Reportando uma Estimativa Pontual (cont.)

- **Exemplo:**

Suponha que estejamos amostrando a partir de uma distribuição normal com média μ e variância σ^2 . Agora, a distribuição de \bar{X} é normal, com média μ e variância σ^2/n ; assim, o erro-padrão de \bar{X} é

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Se não conhecermos σ , mas substituirmos o desvio-padrão S da amostra na equação anterior, então o erro-padrão estimado de \bar{X} será

$$\hat{\sigma}_{\bar{X}} = \frac{S}{\sqrt{n}}$$

7.3 Conceitos Gerais de Estimativa Pontual

7.3.3 Erro-Padrão: Reportando uma Estimativa Pontual (cont.)

- Exemplo II:

Um artigo no *Journal of Heat Transfer* (Trans. ASME, Sec. C, 96, 1974, p. 59) descreveu um novo método de medir a condutividade térmica de ferro Armco. Usando uma temperatura de 100°F e uma potência de 550 W, as 10 medidas seguintes de condutividade térmica (em BTU/h·ft·°F) foram obtidas:

41,60; 41,48; 42,34; 41,95; 41,86;
42,18; 41,72; 42,26; 41,81; 42,04

Uma estimativa da condutividade térmica média a 100°F e 550 W é a média amostral ou

$$\bar{x} = 41,924 \text{ Btu/hr-ft-}^\circ\text{F}$$

7.3 Conceitos Gerais de Estimativa Pontual

7.3.3 Erro-Padrão: Reportando uma Estimativa Pontual (cont.)

- Exemplo II (final):

O erro-padrão da média amostral é $\sigma_{\bar{x}} = \sigma / \sqrt{n}$ e sendo σ desconhecido, podemos trocá-lo pelo desvio-padrão da amostra $s = 0,284$, de modo a obter o erro-padrão estimado de \bar{X} como

$$\hat{\sigma}_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{0,284}{\sqrt{10}} = 0,0898$$

Note que o erro-padrão é cerca de 0,2% da média amostral, implicando que obtivemos uma estimativa relativamente precisa da condutividade térmica. Se pudermos considerar que a condutividade térmica seja normalmente distribuída, então duas vezes o erro-padrão é $2\hat{\sigma}_{\bar{x}} = 2(0,0898) = 0,1796$ e estamos altamente confiantes de que a condutividade térmica média está no intervalo $41,924 \pm 0,1756$ ou entre 41,744 e 42,104.

7.3 Conceitos Gerais de Estimação Pontual

7.3.4 Erro Quadrático Médio de um Estimador

- Definição:

Definição

O **erro quadrático médio** de um estimador $\hat{\Theta}$ do parâmetro θ é definido como

$$\text{EQM}(\hat{\Theta}) = E(\hat{\Theta} - \theta)^2 \quad (7.4)$$

7.3 Conceitos Gerais de Estimação Pontual

7.3.4 Erro Quadrático Médio de um Estimador (cont.)

- Definição (cont.):

O erro quadrático médio pode ser reescrito como segue:

$$\begin{aligned}\text{EQM}(\hat{\Theta}) &= E[\hat{\Theta} - E(\hat{\Theta})]^2 + [\theta - E(\hat{\Theta})]^2 \\ &= V(\hat{\Theta}) + (\text{tendência})^2\end{aligned}$$

Ou seja, o erro quadrático médio de $\hat{\Theta}$ é igual à variância do estimador mais o quadrado da tendência. Se $\hat{\Theta}$ for um estimador não tendencioso de θ , o erro quadrático médio de $\hat{\Theta}$ será igual à variância de $\hat{\Theta}$.

7.3 Conceitos Gerais de Estimação Pontual

7.3.4 Erro Quadrático Médio de um Estimador (cont.)

- **Definição (final):**

O erro quadrático médio é um critério importante para comparar dois estimadores. Sejam $\hat{\Theta}_1$ e $\hat{\Theta}_2$ dois estimadores do parâmetro θ e sejam $\text{EQM}(\hat{\Theta}_1)$ e $\text{EQM}(\hat{\Theta}_2)$ os erros quadráticos médios de $\hat{\Theta}_1$ e $\hat{\Theta}_2$. Então, a **eficiência relativa** de $\hat{\Theta}_2$ para $\hat{\Theta}_1$ é definida como

$$\frac{\text{EQM}(\hat{\Theta}_1)}{\text{EQM}(\hat{\Theta}_2)} \quad (7.5)$$

Se essa eficiência relativa for menor que um, concluiremos que $\hat{\Theta}_1$ é um estimador mais eficiente de θ do que $\hat{\Theta}_2$, pelo fato dele ter menor erro quadrático médio.

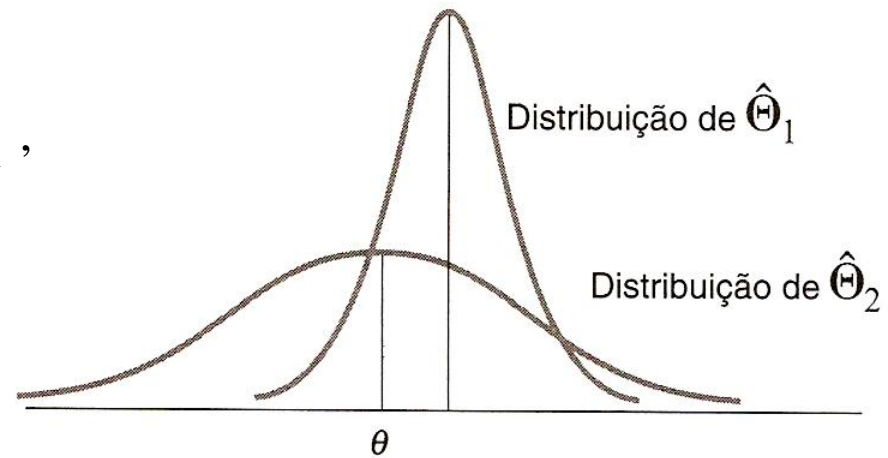
7.3 Conceitos Gerais de Estimação Pontual

7.3.4 Erro Quadrático Médio de um Estimador (cont.)

- **Observação:**

Algumas vezes, encontramos que estimadores tendenciosos são preferíveis aos estimadores não tendenciosos, porque eles têm menor erro quadrático médio (ver figura).

Fig. 7.3 Estimador tendencioso, $\hat{\Theta}_1$, que tem variância menor do que o estimador não tendencioso, $\hat{\Theta}_2$.



7.4 Métodos de Estimação Pontual

7.4.1 Método dos Momentos

- **Definição:**

A ideia por trás do método dos momentos é igualar os **momentos da população**, que são definidos em termos de valores esperados, $E(X)$, $E(X^2)$, etc., aos correspondentes **momentos da amostra**;

- **Exemplo:**

Estimadores de Momento da Distribuição Normal Suponha que X_1, X_2, \dots, X_n seja uma amostra aleatória proveniente de uma distribuição normal, com parâmetros μ e σ^2 . Para a distribuição normal $E(X) = \mu$ e $E(X^2) = \mu^2 + \sigma^2$. Igualando $E(X)$ a \bar{X} e $E(X^2)$ a $\frac{1}{n} \sum_{i=1}^n X_i^2$ fornece

$$\mu = \bar{X}, \quad \mu^2 + \sigma^2 = \frac{1}{n} \sum_{i=1}^n X_i^2$$

A solução dessas equações fornece os estimadores de momento

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

Conclusão Prática: Note que o estimador de momento de σ^2 não é um estimador não tendencioso.

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança

- Definição:

Definição

Suponha que X seja uma variável aleatória com distribuição de probabilidades $f(x; \theta)$, em que θ é um único parâmetro desconhecido. Faça x_1, x_2, \dots, x_n serem os valores observados na amostra aleatória de tamanho n . Então a **função verossimilhança** da amostra é

$$L(\theta) = f(x_1; \theta) \cdot f(x_2; \theta) \cdot \dots \cdot f(x_n; \theta) \quad (7.6)$$

Note que a função verossimilhança é agora uma função somente do parâmetro desconhecido θ . O **estimador de máxima verossimilhança** de θ é o valor de θ que maximiza a função verossimilhança $L(\theta)$.

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- **Comentário:**

No caso de uma variável aleatória discreta, a interpretação da função verossimilhança é clara. A função verossimilhança da amostra $L(\theta)$ é apenas a probabilidade

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

Ou seja, $L(\theta)$ é apenas a probabilidade de obter os valores amostrais x_1, x_2, \dots, x_n . Logo, no caso discreto, o estimador de máxima verossimilhança é aquele que maximiza a probabilidade de ocorrência dos valores da amostra.

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- Exemplo:

Seja X uma variável aleatória de Bernoulli. A função de distribuição de probabilidades é

$$f(x; p) = \begin{cases} p^x(1 - p)^{1-x} & x = 0, 1 \\ 0 & \text{caso contrário} \end{cases}$$

sendo p o parâmetro a ser estimado. A função verossimilhança de uma amostra de tamanho n é

$$\begin{aligned} L(p) &= p^{x_1}(1 - p)^{1-x_1} p^{x_2}(1 - p)^{1-x_2} \dots p^{x_n}(1 - p)^{1-x_n} \\ &= \prod_{i=1}^n p^{x_i} (1 - p)^{1-x_i} \\ &= p^{\sum_{i=1}^n x_i} (1 - p)^{n - \sum_{i=1}^n x_i} \end{aligned}$$

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- Exemplo (final):

Observamos que se \hat{p} maximiza $L(p)$, então \hat{p} também maximiza $\ln L(p)$. Assim,

$$\ln L(p) = \left(\sum_{i=1}^n x_i \right) \ln p + \left(n - \sum_{i=1}^n x_i \right) \ln(1 - p)$$

Agora,

$$\frac{d \ln L(p)}{dp} = \frac{\sum_{i=1}^n x_i}{p} - \frac{\left(n - \sum_{i=1}^n x_i \right)}{1 - p}$$

Igualando isso a zero e resolvendo para p resulta

$$\hat{p} = (1/n) \sum_{i=1}^n x_i.$$

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- Exemplo II:

Seja X normalmente distribuída, com média μ desconhecida e variância σ^2 conhecida. A função verossimilhança de uma amostra aleatória de tamanho n , isto é, X_1, X_2, \dots, X_n , é

$$\begin{aligned} L(\mu) &= \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i - \mu)^2 / (2\sigma^2)} \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-(1/2\sigma^2) \sum_{i=1}^n (x_i - \mu)^2} \end{aligned}$$

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- Exemplo II (final):

Agora

$$\ln L(\mu) = -(n/2) \ln(2\pi\sigma^2) - (2\sigma^2)^{-1} \sum_{i=1}^n (x_i - \mu)^2$$

e

$$\frac{d \ln L(\mu)}{d\mu} = (\sigma^2)^{-1} \sum_{i=1}^n (x_i - \mu)$$

Igualando esse último resultado a zero e resolvendo para μ , tem-se

$$\hat{\mu} = \frac{\sum_{i=1}^n X_i}{n} = \bar{X}$$

7.4 Métodos de Estimaco Pontual

7.4.2 Mtodo da Mxima Verossimilhana (cont.)

- Exemplo III:

Seja X exponencialmente distribuda com parmetro λ . A funo mxima verossimilhana de uma amostra aleatria de tamanho n , X_1, X_2, \dots, X_n , 

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^n \lambda e^{-\lambda x_i} \\ &= \lambda^n e^{-\lambda \sum_{i=1}^n x_i} \end{aligned}$$

O logaritmo da verossimilhana 

$$\ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^n x_i$$

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- Exemplo III (final):

Agora

$$\frac{d \ln L(\lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i$$

e igualando esse último resultado a zero, obtemos

$$\hat{\lambda} = n / \sum_{i=1}^n X_i = 1/\bar{X}$$

Assim, o estimador de máxima verossimilhança de λ é a recíproca da média da amostra.

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- **Propriedades:**

Propriedades do Estimador de Máxima Verossimilhança

Sob condições muito gerais e não restritivas, quando uma amostra de tamanho n for grande e se $\hat{\Theta}$ for um estimador de máxima verossimilhança do parâmetro θ , então

- (1) $\hat{\Theta}$ é um estimador aproximadamente não tendencioso para θ [$E(\hat{\Theta}) \cong \theta$],
- (2) a variância de $\hat{\Theta}$ é aproximadamente tão pequena quanto a variância que poderia ser obtida com qualquer outro estimador, e
- (3) $\hat{\Theta}$ tem uma distribuição normal aproximada.

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- Invariância:

A Propriedade da Invariância

Sejam $\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_k$ estimadores de máxima verossimilhança dos parâmetros $\theta_1, \theta_2, \dots, \theta_k$. Então o estimador de máxima verossimilhança de qualquer função $h(\theta_1, \theta_2, \dots, \theta_k)$ desses parâmetros é a mesma função $h(\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_k)$ dos estimadores $\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_k$.

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- Exemplo:

No caso da distribuição normal, os estimadores de máxima verossimilhança de μ e σ^2 foram $\hat{\mu} = \bar{X}$ e $\hat{\sigma}^2 = \sum_{i=1}^n (X_i - \bar{X})^2/n$. Para obter o estimador de máxima verossimilhança da função $h(\mu, \sigma^2) = \sqrt{\sigma^2} = \sigma$, substitua os estimadores $\hat{\mu}$ e $\hat{\sigma}^2$ na função h , resultando

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \left[\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right]^{1/2}$$

Logo, o estimador de máxima verossimilhança do desvio-padrão σ *não* é o desvio-padrão S da amostra.

7.4 Métodos de Estimação Pontual

7.4.2 Método da Máxima Verossimilhança (cont.)

- **Comentários:**

- Embora o método da máxima verossimilhança seja uma excelente técnica, algumas vezes complicações aparecem durante o seu uso;
- Por exemplo, nem sempre é fácil maximizar a função verossimilhança, pois as equações obtidas de

$$dL(\theta)/d(\theta) = 0$$

podem ser difíceis de resolver;

- Além disso, pode não ser sempre possível usar diretamente métodos de cálculo para determinar

$$\max L(\theta).$$

7.5 Sumário

- Estimativas *razoáveis* para os seguintes parâmetros populacionais são:
 - Para a *média populacional* μ , a estimativa é a *média aritmética da amostra*,
$$\hat{\mu} = \bar{x};$$
 - Para a *variância populacional* σ^2 , a estimativa é a *variância da amostra*,
$$\hat{\sigma}^2 = s^2$$
 - Para a *proporção populacional* p , a estimativa é a *proporção da amostra*,
$$\hat{p} = x/n$$

TERMOS E CONCEITOS IMPORTANTES

Tendência em estimação de parâmetros	Estimador de máxima Verossimilhança	Distribuição normal como uma distribuição amostral da média	Erro padrão e erro padrão estimado de um estimador
Teorema Central do Limite	Erro quadrático médio de um estimador	Estimador pontual	Estatística
Estimador <i>versus</i> estimativa	Estimador não tendencioso de variância mínima	Estimador pontual	Inferência estatística
Função Verossimilhança		Distribuição amostral	Estimador não tendencioso
