Universidade do Minho

DISTRIBUIÇÕES AMOSTRAIS

Distribuições Amostrais

- A população define um conjunto vasto, em geral, impossível de conhecer.
- A amostra constitui um subconjunto da população.
- Uma amostra aleatória é uma amostra em que a probabilidade de cada elemento ser seleccionado é conhecida.
- O objectivo é, a partir da amostra, estabelecer conclusões para o todo representado pela população.

Definições

- Amostra aleatória $x_1, x_2, \ldots, x_n$
- Uma **estatística** é uma medida numérica calculada a partir dos dados amostrais.
- Um **parâmetro** é uma medida numérica de uma população.

Definições

	População	Amostra
Média	μ	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Desvio Padrão	σ	$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$
Variância	σ^2	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$

Definições

- **Estatística Inferêncial** é o conjunto de procedimentos que permitem, a partir de uma amostra, fazer inferências para a população.

- Se $x_1, x_2, \ldots, x_n$ são variáveis aleatórias independentes e identicamente distribuídas, então constituem uma **Amostra Aleatória** de uma população infinita caracterizada pela sua distribuição comum.

Definição

- Se $x_1, x_2, \ldots, x_n$ constituem uma **amostra aleatória**, então

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

é a **média amostral** e,

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

é a **variância amostral**.

- As **estatísticas** são funções de variáveis aleatórias.

Distribuição da Média

Se $x_1, x_2, \ldots, x_n$ constituem uma amostra aleatória de uma população infinita com média e variância então

$$E[\overline{x}] = \mu \qquad\qquad Var[\overline{x}] = \frac{\sigma^2}{n}$$

$$E[\overline{x}] = E\left[\frac{\sum_i x_i}{n}\right] = \frac{1}{n}(n\mu) = \mu \qquad Var[\overline{x}] = V\left[\frac{\sum_i x_i}{n}\right] = \frac{1}{n^2}\left(n\sigma^2\right) = \frac{\sigma^2}{n}$$

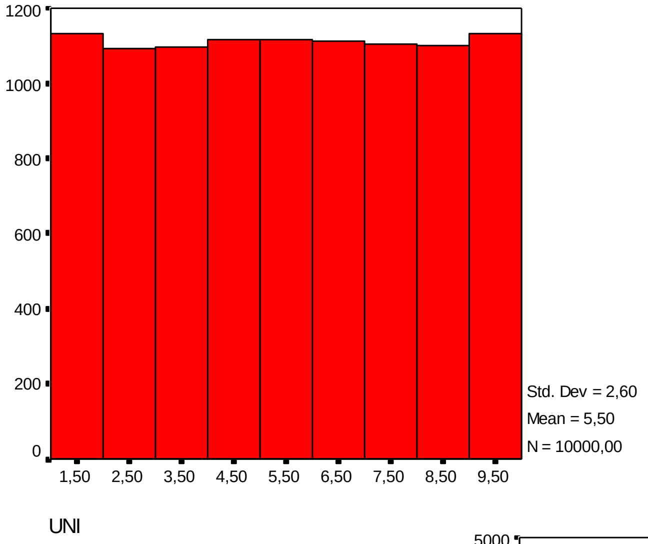
Teorema do Limite Central

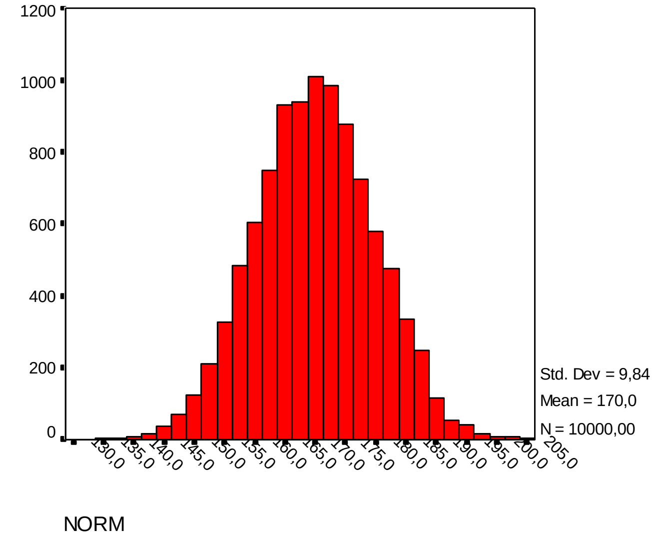
Se $x_1, x_2, \ldots, x_n$ constituem uma amostra aleatória de uma população com média μ e variância σ^2 finita, a distribuição limite de

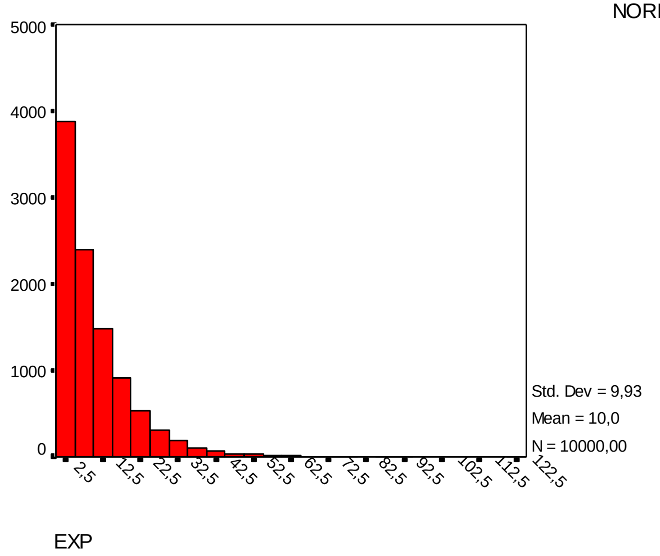
$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

à medida que $n \to \infty$ é a distribuição normal padrão.

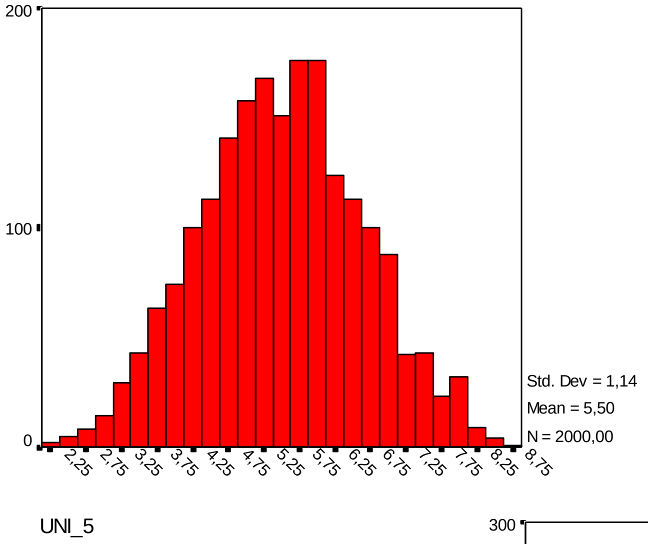
Amostras Aleatórias (N=10 000)

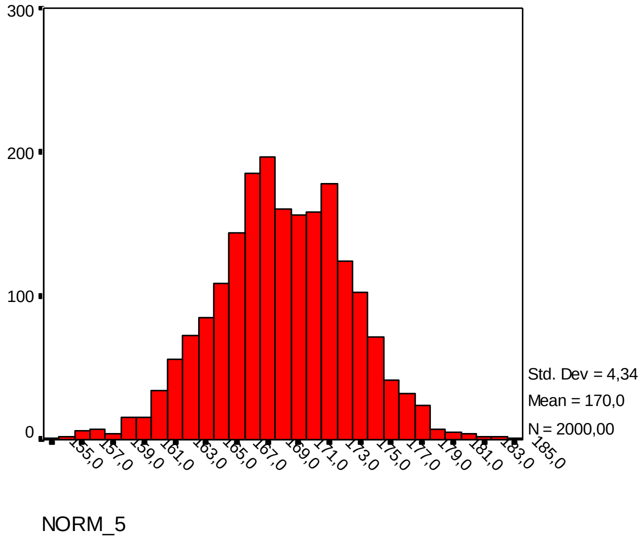


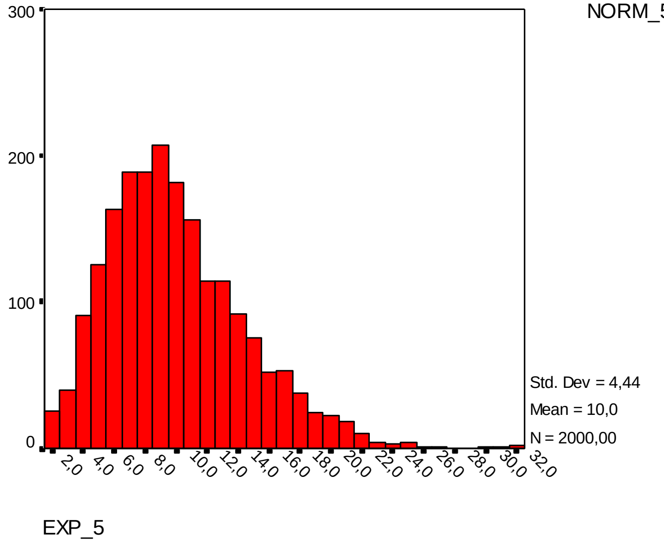




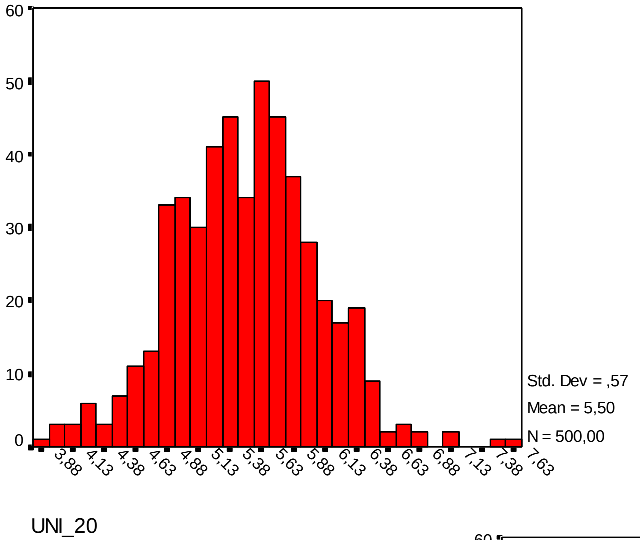
Amostras Aleatórias (N=2 000)

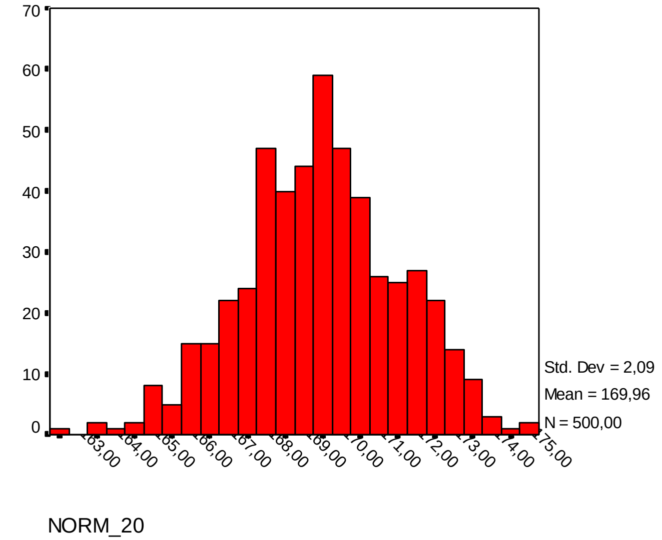


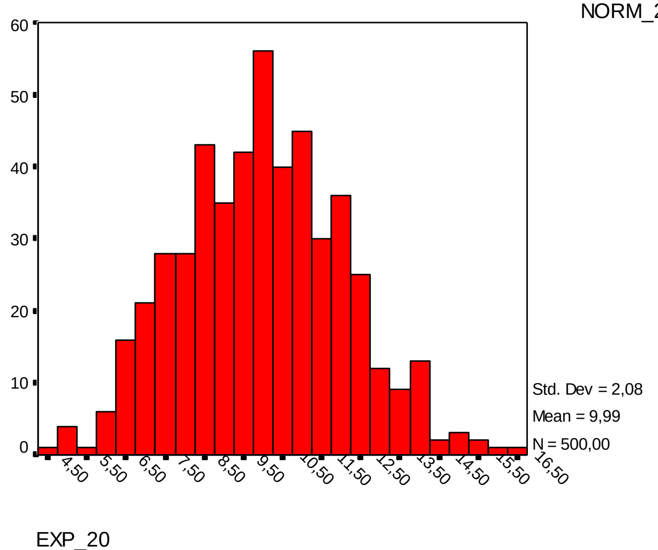




Amostras Aleatórias (N=500)







Distribuições Amostrais

Exemplo 8

Suponha que as classificações, a nível nacional, do exame de Geografia, têm uma média de 14.3, com um desvio padrão 2.1. Assumindo que a distribuição é normal, calcule:

a) a probabilidade de que um estudante, seleccionado aleatoriamente, tenha uma classificação superior a 16 valores.

b) a probabilidade de que uma amostra aleatória de 10 estudantes tenha uma média superior a 16 valores.

a)

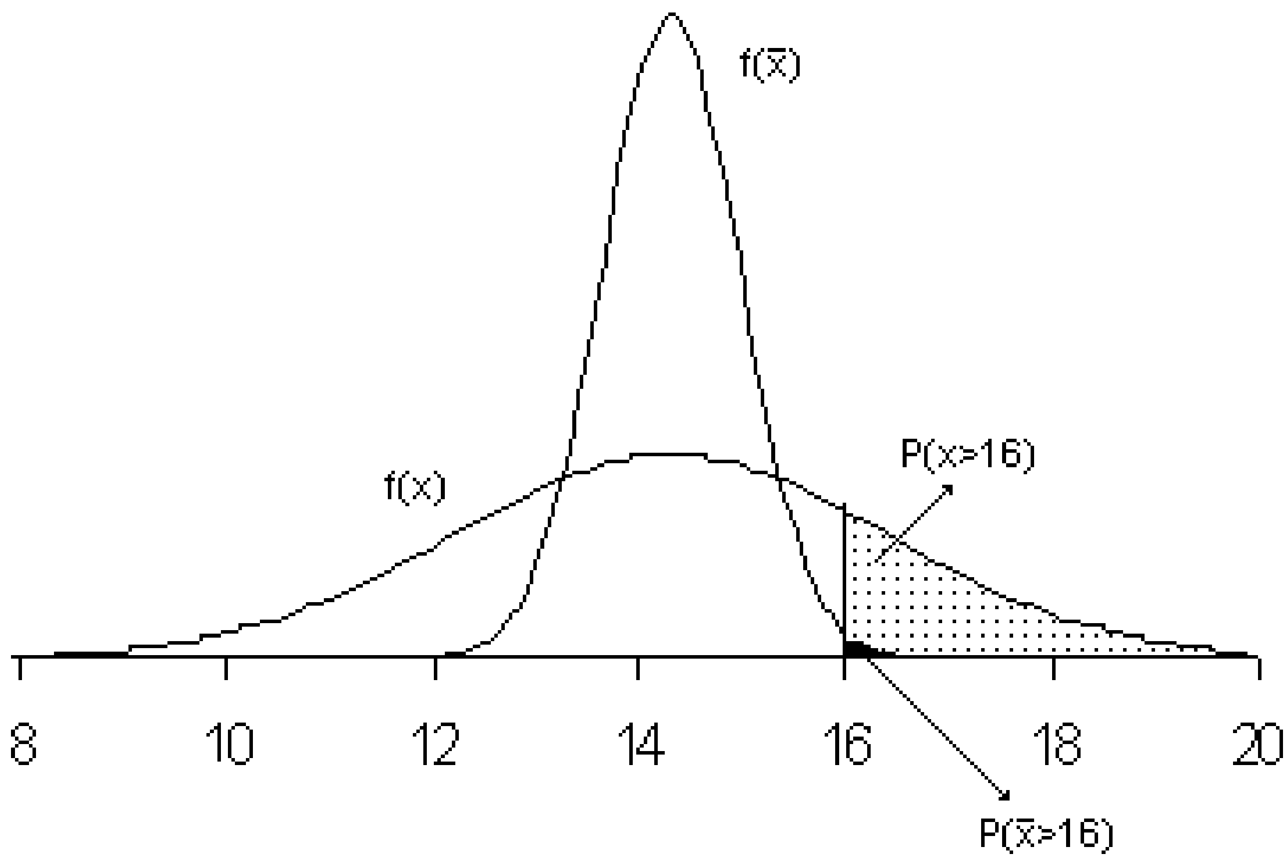
$$P(x > 16) = P\left(z > \frac{16 - 14.3}{2.1}\right) = P(z > 0.81) = 0.2090$$

b)

$$P(\bar{x} > 16) = P\left(z > \frac{16 - 14.3}{2.1/\sqrt{10}}\right) = P(z > 2.56) = 0.0052$$

Distribuições Amostrais

Exemplo 8



Distribuições Amostrais

Exemplo 9

Uma máquina de enchimento de açúcar está regulada por forma a que a quantidade em cada pacote seja de 1000 gramas, com um desvio padrão de 50 gramas.

Qual a probabilidade de que a média de uma amostra de 36 pacotes seja menor que 980 gramas?

$$\mu_{\bar{x}} = \mu_x = 1000$$

$$\sigma_{\bar{x}} = \sigma_x \Big/ \sqrt{n} = \frac{50}{\sqrt{36}} = 8.3$$

$$P(\bar{x} \leq 980) = P(z \leq \frac{980 - 1000}{8.3}) = P(z \leq -2.41) = 0.0080$$

ESTIMADORES PONTUAIS

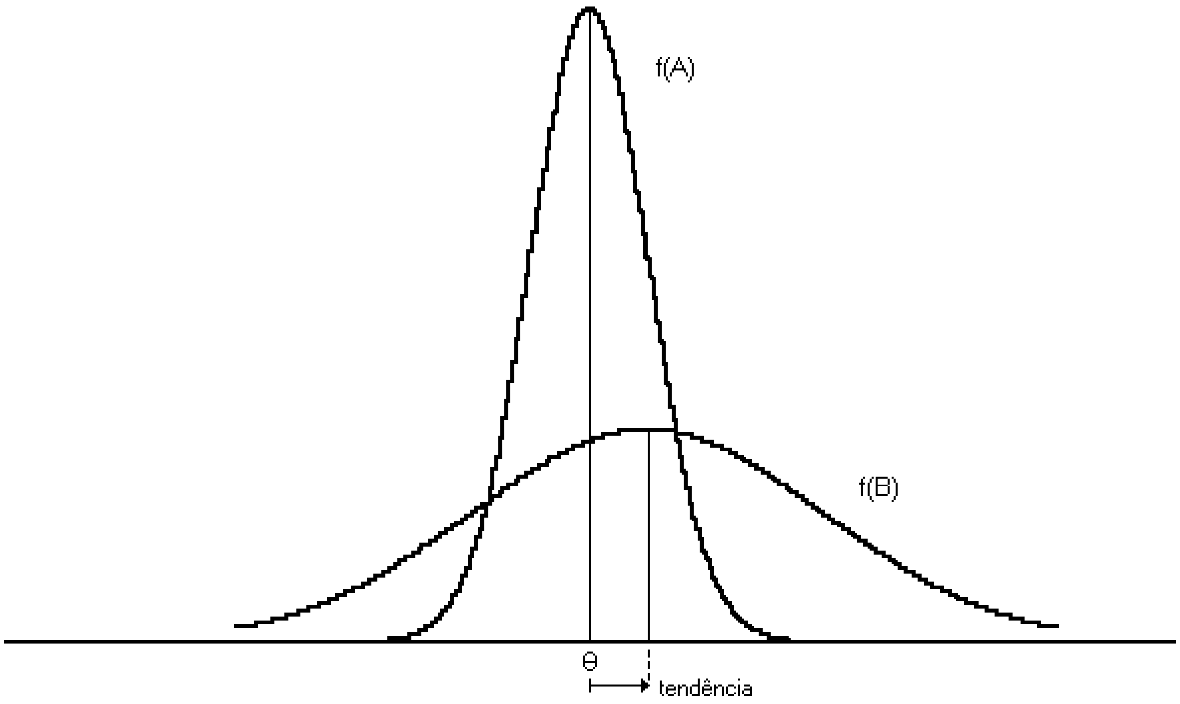
Propriedades

- Tendência
- Variância Mínima
- Eficiência
- Consistência
- Suficiência
- Robustez

Estimador não tendencioso

Uma estatística $\hat{\theta}$ é um estimador não tendencioso do parâmetro θ se e só se

$$E\left[\hat{\theta}\right]=\theta$$



Tendência

Exemplo 1

Mostre que X/*n* é um estimador não tendencioso do parâmetro *p* da distribuição binomial.

$$E[X] = np$$

$$E\left[\frac{X}{n}\right] = \frac{1}{n}E[X] = p$$

Tendência

Exemplo 2

Sejam uma amostra aleatória de uma população dada por

$$f(x) = \begin{cases} e^{-(x-\delta)} & x > \delta \\ 0 & \text{outros valores} \end{cases}$$

Mostre que $\bar{X}$ é um estimador tendencioso de δ.

$$\mu = E[X] = \int_{\delta}^{\infty} x e^{-(x-\delta)} dx = -x e^{-(x-\delta)} \Big|_{\delta}^{\infty} - \int_{\delta}^{\infty} -e^{-(x-\delta)} dx = 1 + \delta$$

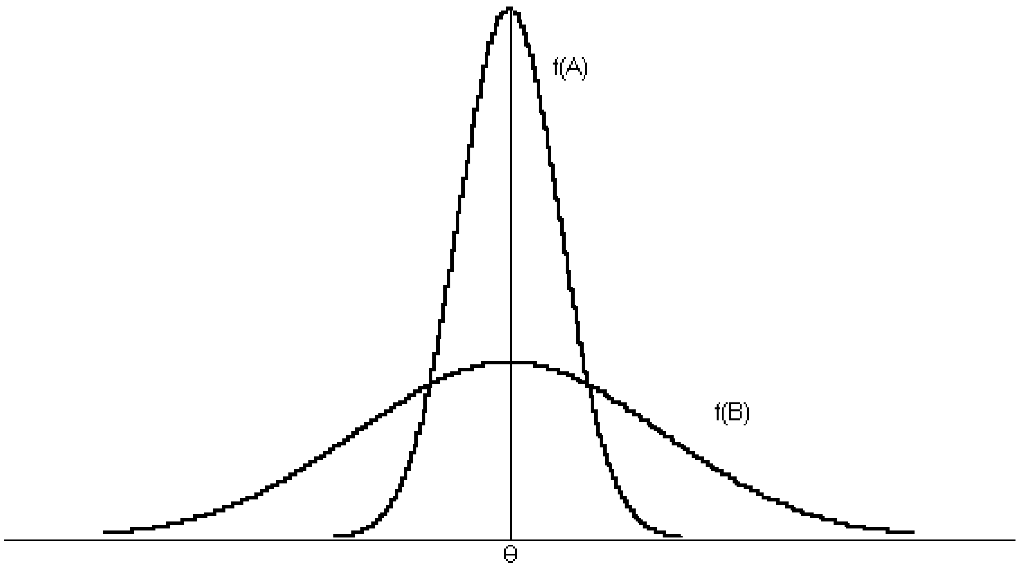
$$E\left[\bar{X}\right] = 1 + \delta \neq \delta$$

$$E\left[\bar{X} - 1\right] = \delta \Rightarrow \bar{X} - 1$$ estimador não tendencioso de δ

Eficiência

Se $\hat{\theta}_1$ e $\hat{\theta}_2$ são dois estimadores não tendenciosos do parâmetro θ de uma dada população e se a variância de $\hat{\theta}_1$ é menor que a variância de $\hat{\theta}_2$, diz-se que $\hat{\theta}_1$ é relativamente mais eficiente que $\hat{\theta}_2$.

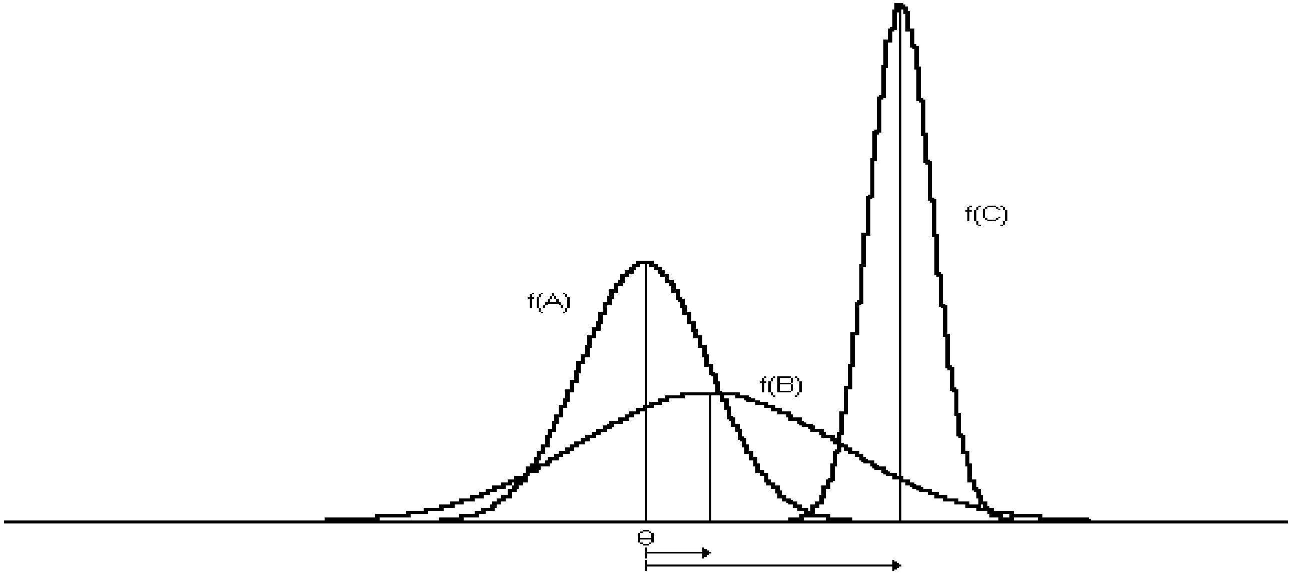
$$\frac{Var\left[\hat{\theta}_1\right]}{Var\left[\hat{\theta}_2\right]}$$



Erro Quadrático Médio

Se $\hat{\theta}$ não é um estimador não tendencioso de um dado parâmetro θ, as comparações devem ser feitas com base no Erro Quadrático Médio em vez da variância.

$$E\left[\left(\hat{\theta}-\theta\right)^2\right]=Var\left[\hat{\theta}\right]+\left[E\left[\hat{\theta}\right]-\theta\right]^2$$



Erro Quadrático Médio

$Var\left[\hat{\theta}\right]$ pequena | $Var\left[\hat{\theta}\right]$ grande

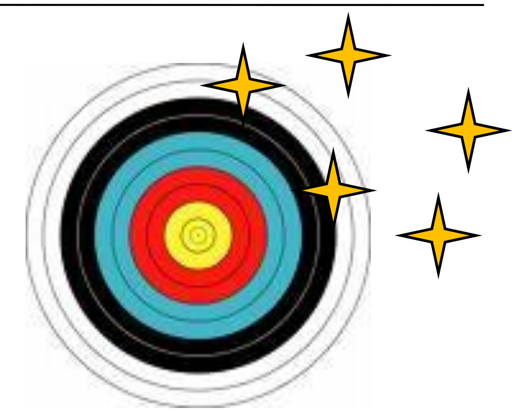
$E\left[\hat{\theta}\right]=\theta$

Exacto e preciso…

Exacto e pouco preciso…

$E\left[\hat{\theta}\right]\neq\theta$

Pouco exacto e preciso…



Pouco exacto e pouco preciso…

Consistência

- A estatística $\hat{\theta}$ é um estimador consistente do parâmetro θ se e só se para cada $c > 0$

$$\lim_{n \to \infty} P\left(\left| \hat{\theta} - \theta \right| < c \right) = 1$$

A consistência é uma propriedade assimptótica.

- Se $\hat{\theta}$ é um estimador não tendencioso de um dado parâmetro θ e $Var\left[\hat{\theta}\right] \to 0$ à medida que $n \to \infty$, então $\hat{\theta}$ é um estimador consistente de θ .

Consistência

Exemplo 3

Mostre que a média aritmética $\overline{x}$ é um estimador consistente da média μ .

$$Var\left[\overline{x}\right]=\frac{\sigma^2}{n}\rightarrow 0 \qquad\qquad n\rightarrow\infty$$

Suficiência

- Um estimador é suficiente se usa toda a informação da amostra relevante para a estimação de θ; isto é, se todo o conhecimento acerca de θ que pode ser ganho a partir dos valores amostrais individuais e da sua ordem, pode também ser ganho pelo valor de $\hat{\theta}$ por si só.
- A estatística $\hat{\theta}$ é um estimador suficiente do parâmetro θ se e só se para cada valor de $\hat{\theta}$ a probabilidade condicional da amostra aleatória $X_1, X_2, \ldots, X_n$ dado $\hat{\theta} = \hat{\theta}_1$ é independente de θ.

Consistência e Suficiência

$$Var\left[\hat{\theta}\right] \qquad\qquad Var\left[\hat{\theta}\right] \to 0$$

Consistência

$$n \to \infty$$

Suficiência

Suficiente… Não suficiente…

Máxima Verosimilhança

- Os estimadores gerados por este método são, em geral, suficientes, não tendenciosos e assimptoticamente de variância mínima.

- Com base nos valores observados na amostra aleatória, é escolhido um valor para a estimativa que maximiza a probabilidade de obter aqueles dados.

Máxima Verosimilhança

Exemplo 4

Uma urna contém um grande número de bolas vermelhas e negras, na proporção de 3:1. Contudo, não se sabe qual das cores está presente em maioria, se a vermelha se a negra.

Para o efeito, uma amostra de 3 bolas é retirada dessa urna. Assim, os resultados possíveis são:

(V,N): (3,0); (2,1); (1,2); (0,3).

– Para um grande número de bolas dentro da urna, as probabilidades podem ser descritas por uma distribuição binomial.

– No entanto, as probabilidades associadas a cada um dos eventos dependem de qual a cor presente em maioria.

Máxima Verosimilhança

Exemplo 4

Nº de bolas vermelhas	0	1	2	3
(V,N)=(3,1) $p = 3/4$	1/64	9/64	27/64	27/64
(V,N)=(1,3) $p = 1/4$	27/64	27/64	9/64	1/64

Tendo saído o resultado: (2,1)
Qual cor é que será mais provável estar em maioria?

Vermelha

Tendo saído o resultado: (0,3)
Qual cor é que será mais provável estar em maioria?

Negra

Função de Máxima Verosimilhança

- No caso discreto, uma amostra aleatória de n observações, $x_1, x_2, \ldots, x_n$ com uma função de probabilidade dependente de um parâmetro θ então a probabilidade de observar estes valores independentes é dada por,

$$P(x_1, x_2, ..., x_n) = P(x_1)P(x_2)...P(x_n) = f(x_1, x_2, ..., x_n; \theta)$$

que corresponde à distribuição de probabilidade conjunta das variáveis aleatórias no ponto amostral $(x_1, x_2, \ldots, x_n)$

Função de Máxima Verosimilhança

Se $x_1, x_2, \ldots, x_n$ são os valores de uma amostra aleatória de uma população com parâmetro θ , a função de verosimilhança é dada por,

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

para valores de θ no domínio dado.

$f(x_1, x_2, ..., x_n; \theta)$ é o valor da função de probabilidade conjunta ou da função de densidade conjunta das variáveis aleatórias $x_1, x_2, \ldots, x_n$ observadas.

Função de Máxima Verosimilhança

Exemplo 5

Considere uma variável aleatória de Bernoulli. A função de probabilidade é dada por

$$f(x;p)=\begin{cases} p^{x}(1-p)^{1-x} & x=0,1 \\ 0 & \textit{outros valores} \end{cases}$$

onde p é o parâmetro a ser estimado.

$$L(p)=p^{x_1}(1-p)^{1-x_1}\,p^{x_2}(1-p)^{1-x_2}\ldots p^{x_n}(1-p)^{1-x_n}=$$

$$=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}$$

O máximo de $L(p)$ é também o máximo de $\ln L(p)$

Função de Máxima Verosimilhança

Exemplo 5

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

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

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Função de Máxima Verosimilhança

Exemplo 6

Considere uma variável aleatória exponencial.

Calcule o estimador de máxima verosimilhança para o parâmetro θ com base numa amostra de tamanho n. Considere, em seguida, uma amostra de dez valores respeitantes ao tempo de vida (em horas) de um componente eléctrico (8.2, 40.5, 3.9, 7.7, 7.1, 3.3, 4.3, 25.4, 5.2,1.0). Estime o valor do parâmetro θ com base nestes 10 valores.

$$L(\theta) = f(x_1)f(x_2)\ldots f(x_n) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}} = \left(\frac{1}{\theta}\right)^n e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$

Função de Máxima Verosimilhança

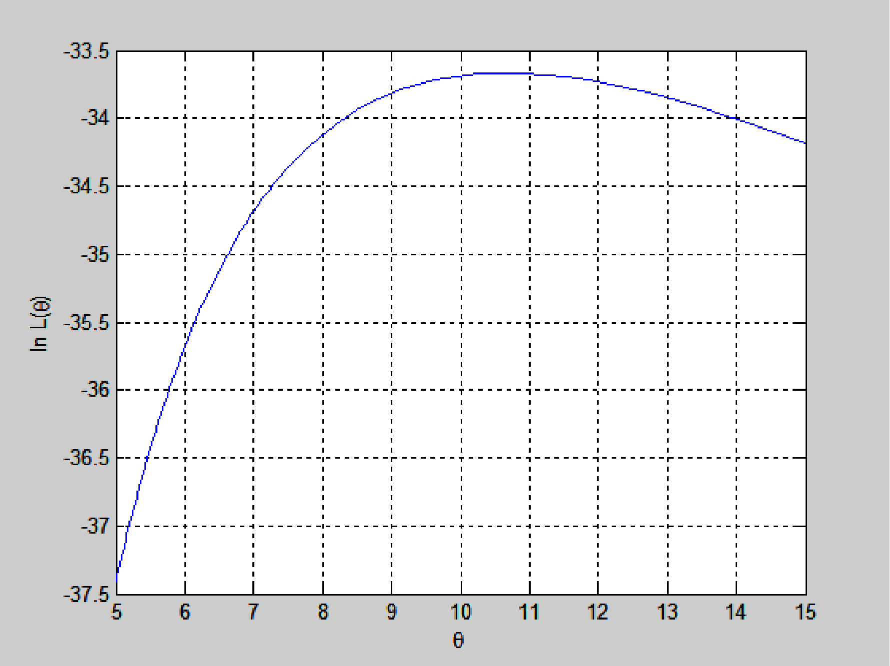
Exemplo 6

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

$$\frac{d\ln L(\theta)}{d\theta} = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i = 0$$

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$$\hat{\theta} = \bar{x} = 10.66$$



Função de Máxima Verosimilhança

Exemplo 7

Sejam $x_1, x_2, \ldots, x_n$ os valores de uma amostra de uma distribuição uniforme, com parâmetros $\alpha = 0, \beta = a$.

Encontre o estimador de máxima verosimilhança para a .

$$f(x;a) = \frac{1}{a} \qquad 0 < x < a$$

$$L(a) = \prod_{i=1}^{n} f(x_i;a) = \left(\frac{1}{a}\right)^n$$

$$\ln L(a) = -n \ln a \qquad \frac{d \ln L(a)}{da} = -\frac{n}{a} = 0$$

Função de Máxima Verosimilhança

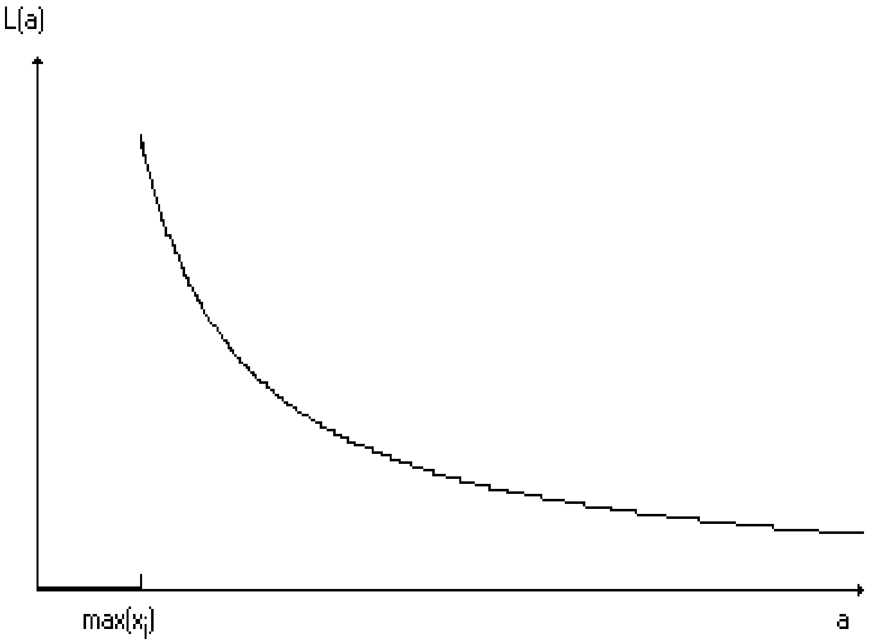
Exemplo 7

- O valor da função de verosimilhança cresce à medida que a decresce.

- Contudo, para qualquer valor observado, $0 \leq x_i \leq a$

  logo a não pode ser menor que qualquer valor da amostra.

- Assim, a função atinge o seu máximo quando a é igual ao maior dos valores da amostra.



$$\hat{a} = \max(x_i)$$