Estimação de parâmetros

ESTIMAÇÃO

- ESTIMAÇÃO PONTUAL
- ESTIMAÇÃO POR INTERVALOS

Objetivo da estimação pontual

Consiste em tentar encontrar a "estatística", cujo valor numérico, obtido através dos dados da amostra, esteja próximo do parâmetro da população, que é constante mas desconhecido.

$\theta \rightarrow$ parâmetro da população

$\hat{\theta} \rightarrow$ estimdor pontual para θ

PROPRIEDADES DE UM ESTIMADOR

- TENDÊNCIA NULA (NÃO TENDENCIOSO, CENTRADO, NÃO ENVIESADO)
- MÉDIA QUADRÁTICA DO ERRO MÍNIMA
- EFICIENTE
- CONSISTENTE
- SUFICIENTE
- ROBUSTO

TENDÊNCIA $t_T(\theta)$

$$\boxed{t_T(\theta) = E[T] - \theta}$$

Diz-se que uma estatística T é um estimador não tendenciosos (ou centrado) em relação ao parâmetro θ, se e só se:

$$\boxed{t_T(\theta) = 0 \Leftrightarrow E[T] = \theta}$$

Exemplo:

$X \sim Bin(n,\pi)$. **Mostrar que $\frac{X}{n}$ é um estimador não tendencioso de π .**

Resolução:

$$E[X] = n.\pi$$

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

$$\therefore T = \frac{X}{n} \text{ é um estimador não tendencioso para } \pi$$

Exemplo:

Se $X_1, X_2, \cdots X_n$ constituem uma amostra aleatória duma população dada por

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

Mostre que $T = \bar{X}$ é um estimador tendencioso de θ.

RESOLUÇÃO:

$$T = \frac{\sum_{i=1}^{n} X_i}{n} \Rightarrow E[T] = E\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] = \frac{1}{n}.E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}.\sum_{i=1}^{n} E[X_i] = E[X_i]$$

$$\mu = E[X_i] = \int_{\theta}^{+\infty} x.f(x)dx = \int_{\theta}^{+\infty} x.e^{-(x-\theta)}dx = \left[-x.e^{-(x-\theta)}\right]_{\theta}^{+\infty} - \int_{\theta}^{+\infty} -e^{-(x-\theta)}dx = 1+\theta$$

$$E[T] = E[X_i] = 1+\theta$$

$$t_T(\theta) = E[T] - \theta \Leftrightarrow t_T(\theta) = 1 \neq 0 \Rightarrow T \text{ é um estimador tendencioso para } \theta.$$

SE SE CONSIDERAR $T' = \bar{X} - 1$ ENTÃO T' É NÃO TENDENCIOSO, POIS $E[T'] = E[\bar{X}] - 1 = \theta$.

MÉDIA QUADRÁTICA DO ERRO (MQE)

A medida, do desempenho de um estimador, mais utilizada é a média quadrática do erro, definida por:

$$\boxed{MQE = E\left[(T-\theta)^2\right]}$$

$$E\left[(T-\theta)^2\right] = \mathrm{var}[T] + \underbrace{\left(E[T]-\theta\right)}_{t_T(\theta)}{}^2$$

$$MQE = E[(T-\theta)^2] = E[T^2 - 2T\theta + \theta^2] = E[T^2] - 2\theta E[T] + \theta^2 =$$

$$= E[T^2] - (E[T])^2 + (E[T])^2 - 2\theta E[T] + \theta^2 =$$

$$= Var[T] + (E[T]-\theta)^2 = Var[T] + t_T^2(\theta)$$

QUANDO O ESTIMADOR É NÃO TENDENCIOSO A MQE RESUME-SE À VARIÂNCIA DO ESTIMADOR.

UM "BOM" ESTIMADOR CORRESPONDE ÀQUELE QUE POSSUIR MENOR MQE.

EFICIÊNCIA

Se T_1 e T_2 são dois estimadores não tendenciosos do parâmetro θ duma população e se $\mathrm{var}[T_1] < \mathrm{var}[T_2]$, diz-se que T_1 é relativamente mais eficiente que T_2.

$$\boxed{ef(T_1,T_2) = \frac{\mathrm{var}[T_1]}{\mathrm{var}[T_2]}}$$

Se T é um estimador tendencioso dum dado parâmetro θ, as comparações devem ser feitas com base na média quadrática do erro.

$$\boxed{ef(T_1,T_2) = \frac{MQE[T_1]}{MQE[T_2]}}$$

CONSISTÊNCIA

A estatística T é um estimador consistente do parâmetro θ se e só se para cada $c > 0$,

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

De notar que a consistência é uma propriedade assimptótica.
Se T é um estimador não tendencioso do parâmetro θ e $\mathrm{var}[T] \to 0$ à medida que $n \to \infty$, então T é um estimador consistente de θ.

O estimador é **consistente** quando suas estimativas se aproximam do valor verdadeiro que se quer estimar, à medida que a amostra cresce.

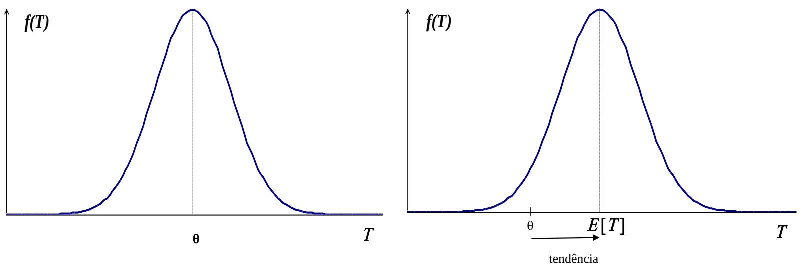
SUFICIÊNCIA

Um estimador é suficiente se toda a informação na amostra relevante para a estimação de θ, isto é, se todo o conhecimento acerca de θ que pode ser ganho a partir dos valores individuais e da sua ordem, pode também ser ganho pelo valor de T por si só.

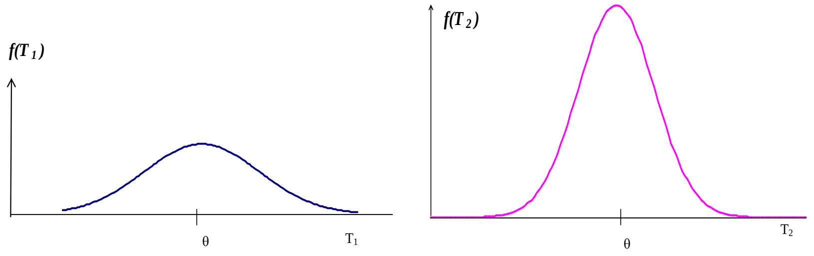
A estatística T é um estimador suficiente do parâmetro θ se e só se para cada valor de T a probabilidade condicional da amostra aleatória $X_1, X_2, \cdots X_n$ dado $T = t$ é independente de θ.

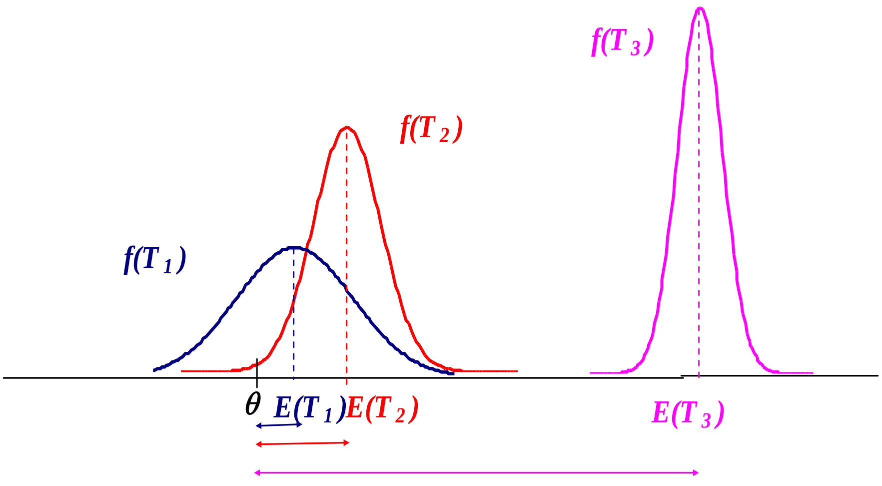
Estimadores Pontuais: Propriedades

Não Tendenciosos $t_T(\theta)=0 \Leftrightarrow E[T]=\theta$



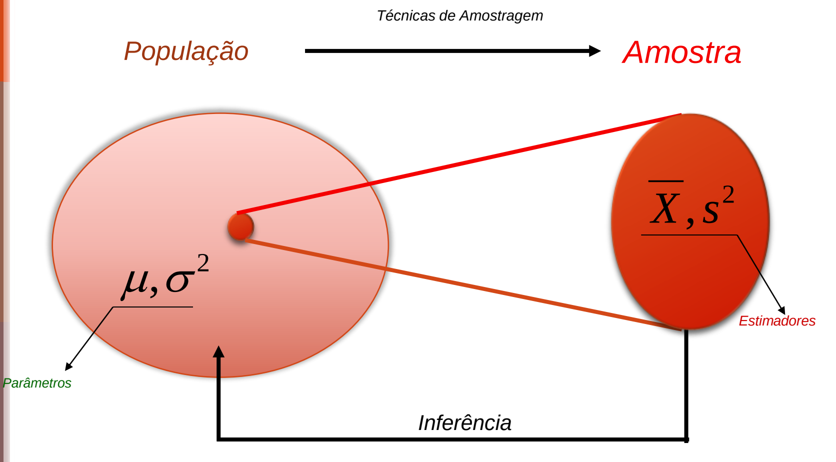
VARIÂNCIA MÍNIMA





f(T 3)
f(T 2)
f(T 1)
θ E(T 1)E(T 2)
E(T 3)

ESTIMAÇÃO



Técnicas de Amostragem
População
Amostra
$\overline{X}, s^2$
μ, σ^2
Estimadores
Parâmetros
Inferência

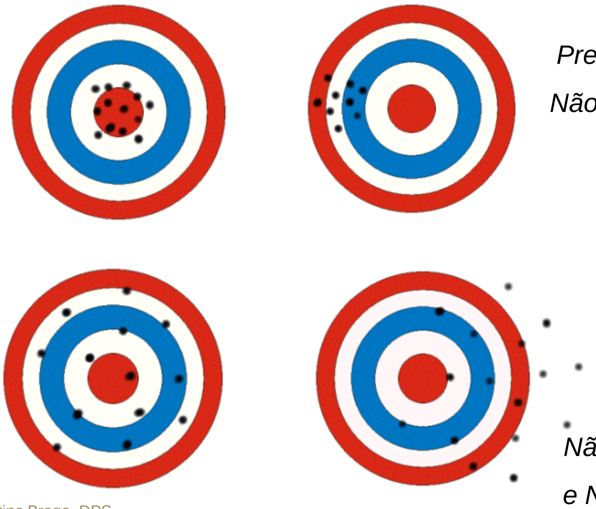
TIPOS DE ERROS

Como as unidades de uma população variam, na estimação, deve-se ter em conta essas variações e calcular o possível erro cometido.

Erros Sistemáticos: *Todas as medidas* $x_1, x_2, \ldots, x_n$ *da amostra diferem do valor verdadeiro* μ *por uma quantidade (ou sentido) constante,* δ.

Erros Aleatórios ou Estatísticos: *Todas as medidas* $x_1, x_2, \ldots, x_n$ *da amostra se distribuem de maneira aleatória em torno do valor verdadeiro* μ.

PRECISÃO *vs* EXATIDÃO



PRECISÃO *vs* EXATIDÃO

Na prática, não vemos o alvo …

… então, como avaliar a qualidade de sua estimativa?

Propriedades dos estimadores

- TENDÊNCIA NULA
- MÉDIA QUADRÁTICA DO ERRO MÍNIMA
- EFICIENTE
- CONSISTENTE
- SUFICIENTE
- ROBUSTO