Estatísticas

Funções da amostra que não dependem de $\theta \in \Theta$

Exemplo:
Seja $(X_1, \dots, X_n)$ a.a. de $X \sim f_\theta, \theta \in \Theta$.
São estatísticas:

1. $T_1(X_1, \dots, X_n) = X_1 + \dots + X_n$
2. $T_2 = \bar{X} = \frac{X_1 + \dots + X_n}{n}$
3. $T_3 = \max\{X_1, \dots, X_n\} = X_{(n)}$
4. $T_4 = \min\{X_1, \dots, X_n\} = X_{(1)}$
5. $T_5(X_1, \dots, X_n) = X_{(n)} - X_{(1)}$
6. $T_6(X_1, \dots, X_n) = X_i, \text{para algum } i = 1, \dots, n$
7. $T_7(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$
8. $T_8(X_1, \dots, X_n) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$
9. $T_9(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^{n} |X_i - \bar{X}|$
10. $T_7(X_1, \dots, X_n) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}$

    etc.

Observação: As estatísticas são variáveis aleatórias.

$$\underset{\sim}{X_n} = (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$$

$$T(\underset{\sim}{X}) = T \circ \underset{\sim}{X} : \Omega \to \mathbb{R}$$

Estimadores

São estatísticas cujo objetivo é estimar uma Quantidade de Interesse. Portanto, estimadores são também variáveis aleatórias

Estimativas

São os valores observados a partir da amostra observada dos estimadores. Portanto, *estimativas são valores numéricos*

Exemplos:

1. $\bar{X}$ é uma estatística, $\bar{X}$ é um estimador para $g(\theta) = E_\theta(X)$
2. Observando $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ é uma estimativa

   Seja (X_1, X_2) a.a de $X \sim \text{Ber}(\theta), \theta \in (0,1)$. Considere as estatísticas e suas funções de probabilidade:

$$
\begin{aligned}
&T_1(X_1, X_2) = X_1 \\
&T_2(X_1, X_2) = X_2 \\
&T_3(X_1, X_2) = X_1 + X_2 \\
&T_4(X_1, X_2) = \max\{X_1, X_2\} \\
&T_5(X_1, X_2) = \min\{X_1, X_2\} \\
&T_6(X_1, X_2) = \frac{1}{2}[(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2] = S_n^2 = \frac{1}{2} S_{n-1}^2
\end{aligned}
$$

(X_1, X_2)	$P_{\theta X_1, X_2}$	X_1	X_2	$X_1 + X_2$	$\max\{X_1, X_2\}$	$\min\{X_1, X_2\}$	$\bar{X}$	S_n^2	S[illegible]
$(0,0)$	$(1-\theta)^2$	0	0	0	0	0	0	0	[illegible]
$(1,0)$	$\theta(1-\theta)$	1	0	1	1	0	0.5	0.25	0[illegible]
$(0,1)$	$\theta(1-\theta)$	0	1	1	1	0	0.5	0.25	0[illegible]
$(1,1)$	θ^2	1	1	2	1	1	1	0	[illegible]

   Calcule para $T \in \{T_1, T_2, T_3, T_4, T_5, T_6\}$

   a-)

   $E_\theta(T)$

$E_\theta(T_1(X_1, X_2)) = E_\theta(X_1) = \theta = 0(1-\theta)^2 + 1 \cdot \theta(1-\theta) + 0\theta(1-\theta) + 1\theta^2 = \theta$

O mesmo vale para 2.

$E_\theta(T_3(X_1, X_2)) = E_\theta(X_1 + X_2) = 2\theta$

$E_\theta(T_4(X_1, X_2)) = E_\theta(\max\{X_1, X_2\}) = 0 \cdot (1-\theta)^2 + 1 \cdot [2 \cdot \theta(1-\theta) + \theta^2] = 2\theta - \theta^2$

$E_\theta(T_5(X_1, X_2)) = E_\theta(\min\{X_1, X_2\}) = \theta^2$

$E_\theta(T_6(X_1, X_2)) = 2\theta(1-\theta)$

b-)$\text{Var}_\theta(T)$

Termine com os mesmos raciocínios

c-)$P_\theta(T = 0)$

Termine com os mesmos raciocínios

Alguns resultados importantes:

$$
\begin{aligned}
E_\theta(\bar{X}) &= E_\Theta(\frac{X_1 + X_2}{2}) = \theta, \theta \in (0,1) \\
E_\theta(\bar{X}^2) &= 0^2(1-\theta)^2 + \frac{2}{4}\theta(1-\theta) + 1^2\theta^2 = \frac{1}{2}\theta + \frac{1}{2}\theta^2, \theta \in (0,1) \\
\Rightarrow \text{Var}_\theta(\bar{X}) &= \frac{\theta + \theta^2}{2} - \theta^2 = \frac{\theta(1-\theta)}{2}, \theta \in (0,1)
\end{aligned}
$$

Aplicando em nossa tabela (S_{n-1}^2):

$$E_\theta([(X_1-\bar{X})^2+(X_2-\bar{X})^2]) = \theta(1-\theta)$$
$$E_\theta([(X_1-\bar{X})^2+(X_2-\bar{X})^2]^2) = \frac{\theta(1-\theta)}{2}$$
$$\Rightarrow \mathrm{Var}_\Theta([(X_1-\bar{X})^2+(X_2-\bar{X})^2]) = \frac{1}{2}\theta(1-\theta)[1-2\cdot\theta(1-\theta)]$$

Propriedades dos estimadores para quantidades de interesse

Estimados não viciados ou enviesados

Seja $(X_1,\ldots,X_n)$ a.a. de $X \sim f_{\theta,}\theta \in \Theta$ e considere $T(X_1,\ldots,X_n) = \hat{\theta}$ um estimador para θ.
Dizemos que $\hat{\theta}$ é *não-enviesado* para θ $\Leftrightarrow$

$$E_\theta(\hat{\theta}_n) = \theta, \forall\theta \in \Theta$$

De forma geral, $T(X_1,\ldots,X_n)$ é um estimador não-viciado para $g(\theta) \Leftrightarrow$,

$$E_\theta(T(X_1,\ldots,X_n)) = g(\theta), \forall\theta \in \Theta$$

Caso contrário, dizemos que $T(X_1,\ldots,X_n)$ é viciado ou enviesado para $g(\theta)$
Dizemos que $\hat{\theta}_n$ é fracamente consistente para $\theta \Leftrightarrow$

$$\lim_{n\to\infty} P_\theta(|\hat{\theta}_n - \theta| > \epsilon) = 0, \forall\theta \in \Theta$$

e para cada $\epsilon > 0$ fixado.

Estimados não viciados assintoticamente

Dizemos que $T(X_1,\ldots,X_n)$ é um estimador assintoticamente não viciado para $g(\theta) \Leftrightarrow$

$$\lim_{n\to\infty} E_\theta(T(X_1,\ldots,X_n)) = g(\theta), \forall\theta \in \Theta$$