**CC0288 - Inferência Estatística I**

**Gabarito da Segunda Verificação de Aprendizagem - 03/04/2023.**

**Segunda Chamada da Primeira Verificação de Aprendizagem .**

**Prof. Maurício**

1. (Valor 6 escores ) Seja $X$ uma variável aleatória que depende de um único parâmetro $\theta$ com função densidade de probabilidade ou função de probabilidade $f(x|\theta)$ , suporte $A$ e espaço paramétrico $\Theta$. Seja $X_1, X_2, \ldots X_n$ uma amostra aleatória de tamanho $n$. Seja

$$T = h(X_1, X_2, \ldots, X_n)$$

um estimador de $\theta$ com $E(T)$ e $V(T)$ finitos.

a. (Valor 1 escore ) Quando $T$ é um estimador não viciado de $\theta$?

**Solução:**
Dizemos que $T$ é um estimador não viciado de $\theta$ quando

$$E(T) = \theta, \ \forall \ \theta \ \in \ \Theta.$$

b. (Valor 2 escores ) Defina viés $B(T)$ de $T$. Dê uma outra definição de estimador não viciado baseada no viés.

**Solução:**
O vício de $T$, $B(t)$, é definido por:

$$B(T) = E(T) - \theta.$$

Dizemos que $T$ é um estimador não viciado de $\theta$ quando

$$B(T) = 0, \ \forall \ \theta \ \in \ \Theta.$$

c. (Valor 3 escores ) Defina erro quadrático médio de $T$. Mostre que ele pode ser colocado na forma:

$$EQM(T) = Var(T) + B^2(T).$$

**Solução:**
O erro quadrático médio de um estimador $T$ do parâmetro $\theta$ é dado por:

$$EQM(T) = E\left[(T - \theta)^2\right].$$

Note que:

$$T - \theta = T - E(T) + E(T) - \theta$$

Elevando ao quadrado temos:

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

$$(T - \theta)^2 = (T - E(T))^2 + (E(T) - \theta)^2 + 2 \times (E(T) - \theta)\,(T - E(T))\,.$$

Aplicando o operador esperança temos:

$$E\left((T - \theta)^2\right) = E\left((T - E(T))^2\right) + E\left((E(T) - \theta)^2\right) + 2 \times (E(T) - \theta) \times E\,(T - E(T))\,.$$

Mas

$$E\,(T - E(T)) = E(T) - E(T) = 0.$$

$$E\left((E(T) - \theta)^2\right) = (E(T) - \theta)^2 = B^2(T).$$

Logo,

$$EQM(T) = Var(T) + B^2(T).$$

2. (Valor 6 escores) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho $n$ de $X$ com $E(X) = \mu$ e $Var(X) = \sigma^2$.

   Sejam

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad e \quad T = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}.$$

   a. (Valor 2 escores ) Mostre que

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\,\bar{X}^2.$$

   **Solução:**

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} \left(X_i^2 - 2\bar{X}\,X_i + \bar{X}^2\right).$$

$$= \sum_{i=1}^{n} X_i^2 - 2\bar{X} \sum_{i=1}^{n} X_i + n\,\bar{X}^2$$

   Como

$$\sum_{i=1}^{n} X_i = n\,\bar{X},$$

   temos:

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2,$$

   Finalmente temos:

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\,\bar{X}^2.$$

   b. (Valor 3 escores ) Mostre que

$$E(T) = \frac{n-1}{n}\sigma^2.$$

   Assim $T$ um estimador viciado para $\sigma^2$.

   **Solução:** De acordo com o enunciado temos:

$$nT = \sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\,\bar{X}^2.$$

   Aplicando o operador esperança temos:

$$nE(T) = \sum_{i=1}^{n} E(X_i^2) - n\,E(\bar{X}^2).$$

   Mas

$$E(X_i^2) = Var(X_i) + E^2(X_i) = \sigma^2 + \mu^2.$$

$$E(\bar{X}^2) = Var(\bar{X}) + E^2(\bar{X}) = \frac{\sigma^2}{n} + \mu^2.$$

Assim,

$$nE(T) = \sum_{i=1}^{n} (\sigma^2 + \mu^2) - n\,(\frac{\sigma^2}{n} + \mu^2).$$

$$nE(T) = n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2 = (n-1)\sigma^2.$$

$$E(T) = \frac{n-1}{n}\sigma^2.$$

c. (Valor 1 escore ) Ele é assintoticamente não viciado para $\sigma^2$?

**Solução:** Devemos mostrar que

$$\lim_{n\to\infty} E(T) = \sigma^2, \quad \forall\ \sigma^2\ \in\ \Theta$$

ou

$$\lim_{n\to\infty} B(T) = 0, \quad \forall\ \sigma^2\ \in\ \Theta$$

Note que

$$\lim_{n\to\infty} \frac{n-1}{n} = \lim_{n\to\infty} \frac{1/n-1}{1} = 1.$$

Logo

$$\lim_{n\to\infty} E(T) = \sigma^2, \quad \forall\ \sigma^2\ \in\ \Theta$$

O viés de $T$ é dado por:

$$B(T) = E(T) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = (\frac{n-1}{n} - 1)\sigma^2 = -\frac{1}{n}\sigma^2.$$

Mas

$$\lim_{n\to\infty} \frac{1}{n} = 0.$$

$$\lim_{n\to\infty} B(T) = 0, \quad \forall\ \sigma^2\ \in\ \Theta$$

3. (Valor 13 escores) Seja $X \sim Bin(m = 5, p)$, $p$ desconhecido.

a. (Valor 3 escores ) Mostre que

$$f(x|p) = \binom{5}{x} p^x (1-p)^{5-x} \;\; I_{\{0,1,2,3,4,5\}}(x), \; 0 \leq p \leq 1.$$

Identifique seu suporte, espaço paramétrico, esperança , variância e função geradora de probabilidade.

**Solução:** O suporte é dado por:

$$A = \{0, 1, 2, 3, 4, 5\}.$$

O espaço paramétrico é dado por:

$$\Theta = [0, 1].$$

Além disso

$$E(X) = \mu = 5p \;\;,\;\; Var(X) = \sigma^2 = 5p(1-p) \;\;\;,\;\;\; G(t) = (pt + q)^5, q = 1 - p.$$

b. (Valor 2 escores )Mostre que

$$\log(f(x|p)) = [c(p)T(x) + d(p) + b(x)] \; I_A(x),$$

e $A$ não depende de $p$. Identifique cada componente.

**Solução:**
Note que o suporte não depende de $p$.
Vamos colocar a f.p. de $X$ na seguinte forma:

$$f(x|p) = \binom{5}{x} p^x (1-p)^{5-x} = \left[\frac{p}{1-p}\right]^x \; (1-p)^5 \binom{5}{x}$$

Aplicando logaritmo neperiano temos:

$$\log(f(x|p)) = \log\left(\frac{p}{1-p}\right) \; x + 5\log(1-p) + \log\left(\binom{5}{x}\right).$$

Assim

$$c(p) = \log\left(\frac{p}{1-p}\right) = \log(p) - \log(1-p) \;\;\;,\;\;\; d(p) = 5\log(1-p).$$

$$T(x) = x \;\; e \;\; b(x) = \log\left(\binom{5}{x}\right),$$

logo pertence à família exponencial.

c (Valor 1 escore ) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho $n$ de $X$. Qual a distribuição conjunta da amostra?

**Solução:**
Sabemos que:

$$f(x_1, x_2, \ldots, x_n|p) = \prod_{i=1}^{n} f(x_i|p)$$

$$= \prod_{i=1}^{n} \left[\frac{p}{1-p}\right]^{x_i} (1-p)^5 \binom{5}{x_i} I_A(x_i).$$

$$= \left[\frac{p}{1-p}\right]^{s} (1-p)^{5n} \prod_{i=1}^{n} \binom{5}{x_i} I_A(x_i),$$

$$= \left[\frac{p}{1-p}\right]^{s} (1-p)^{5n} \prod_{i=1}^{n} \binom{5}{x_i} I_{A^n}(x_1, x_2, \ldots, x_n),$$

com $s = \sum_{i=1}^{n} x_i$.

d. (Valor 1 escore ) Qual a função de verossimilhança de $p$?

**Solução:**

e. (Valor 4 escores ) Qual a lei de $S = \sum_{i=1}^{n} X_i$ ?
Baseado em $S$ proponha um estimador não viciado $T$ para $g(p) = p$.
Qual a variância de $T$?

**Solução:** A função geradora de probabilidade de $S$ é dada por:

$$G_S(t) = [G_X(t)]^n = (pt+q)^{5n}.$$

assim

$$S \sim Bin(5n, p).$$

Sabemos que

$$E(S) = 5np \quad e \quad Var(S) = 5np(1-p).$$

Fazendo

$$T = \frac{S}{5n} = \frac{\bar{X}}{5}$$

é o nosso estimador não viciado para $p$.

$$E(T) = E\left[\frac{S}{5n}\right] = \frac{1}{5n} E(S) = \frac{1}{5n} 5np = p.$$

A variância de $T$ é dada por:

$$Var(T) = Var\left(\frac{S}{5n}\right) = \frac{1}{25n^2}\, Var(S) = \frac{1}{25n^2}\, 5np(1-p) = \frac{p(1-p)}{5n}.$$

f. (Valor 2 escores ) Uma amostra aleatória de tamanho $n = 100$ e apresentou o seguinte resultado:

| x | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| f | 1 | 7 | 29 | 32 | 24 | 7 |

Calcule uma estimativa pontual para $p$ baseada no item **e**.

**Solução:**
Vamos entender a nossa amostra de tamanho $n = 100$.
Temos que apareceu um único valor zero, 7 vezes o valor um,29 vezes o valor dois ,32 vezes o valor três, 24 vezes o valor quatro e 7 vezes o valor cinco.
A soma dos valores amostrais é dada por:

$$s = \sum_{i=1}^{100} x_i = \sum_{x=0}^{5} x\ f_x.$$

$$s = 0 \times 1 + 1 \times 7 + 2 \times 29 + 3 \times 32 + 4 \times 24 + 5 \times 7 =$$

$$s = 0 + 7 + 58 + 96 + 96 + 35 = 292$$

A estimativa pontual para $p$ é dada por:

$$t = p_{est} = \frac{292}{500} = 0,584.$$

vamos começar a estudar o **R**:

```
##Gerar uma amostra aleatória m=5 e p=0,6.
>
>
>
> m=5;p=0.6;n=100
>
> set.seed(32)
> A=rbinom(n,m,p)
>
> A
[1] 3 3 2 2 4 1 2 2 3 3 3 4 3 2 4 2 2 3 3 2 2 2 3 4 3 1 3 3 3 3 3 2 2 2 3 2 3
[38] 4 2 1 3 2 4 3 2 4 3 1 2 5 2 1 2 3 5 5 2 2 3 1 2 4 2 4 4 0 2 4 1 3 2 4 3 5
[75] 4 4 2 4 3 3 3 3 3 5 4 3 3 2 4 2 5 4 4 4 4 5 4 3 4 4
>
>
> table(A)
A
0  1  2  3  4  5
1  7 29 32 24  7
>
```

```
> s=sum(A);s
[1] 292
>
>
>
> #### A estimativa pontual de p:
>
> p_est=s/(5*n);p_est #### t. compare com 0,6.
[1] 0.584
>
>
> mean(A);mean(A)/5
[1] 2.92
[1] 0.584
```