CC-0288 - Inferência Estatística I

Exemplo 1.3.7 - 17/03/2023

Prof. Maurício Mota

1. Vamos comentar o Exemplo 1.3.7 da página 12 do livro do Heleno e da Mônica.
   Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de $X \sim B(\theta)$.
   Sabemos que :

$$\mu = E(X) = \theta \quad e \quad \sigma^2 = \theta(1-\theta).$$

$$S = \sum_{i=1}^{n} X_i \sim Bin(n, \theta).$$

$$E(S) = n\theta \quad e \quad Var(S) = n\theta(1-\theta)$$

Consideremos os estimadores:

$$\hat{\theta}_1 = \frac{S}{n} = \bar{X} \quad e \quad \hat{\theta}_2 = \frac{S + \sqrt{n}/2}{n + \sqrt{n}}.$$

Note que:

$$E\left[\hat{\theta}_1\right] = E\left[\bar{X}\right] = \theta,$$

que é não viciado para θ.

$$Var\left[\hat{\theta}_1\right] = E\left[\bar{X}\right] = \frac{\sigma^2}{n} = \frac{\theta(1-\theta)}{n} = EQM\left[\hat{\theta}_1\right].$$

$$E\left[\hat{\theta}_2\right] = \frac{E(S) + \sqrt{n}/2}{n + \sqrt{n}} = \frac{n\theta + \sqrt{n}/2}{n + \sqrt{n}}.$$

$$E\left[\hat{\theta}_2\right] = \frac{n}{n + \sqrt{n}}\,\theta + \frac{\sqrt{n}}{2(n + \sqrt{n})}.$$

Vamos calcular o vício de $\hat{\theta}_2$

$$B\left[\hat{\theta}_2\right] = E\left[\hat{\theta}_2\right] - \theta$$

$$B\left[\hat{\theta}_2\right] = \left(\frac{n}{n + \sqrt{n}} - 1\right)\theta + \frac{\sqrt{n}}{2(n + \sqrt{n})}.$$

$$B\left[\hat{\theta}_2\right] = -\frac{\sqrt{n}}{n + \sqrt{n}}\,\theta + \frac{\sqrt{n}}{2(n + \sqrt{n})}.$$

Vemos que o vício de $\hat{\theta}_2$ é uma função afim de θ.

O quadrado do viés é dado por:

$$B^2\left[\hat{\theta}_2\right] = \frac{n}{(n+\sqrt{n})^2}\,\theta^2 - \frac{n}{(n+\sqrt{n})^2}\,\theta + \frac{n}{4(n+\sqrt{n})^2}.$$

$$B^2\left[\hat{\theta}_2\right] = \frac{n}{4(n+\sqrt{n})^2}\,\left[\theta^2 - \theta + \frac{1}{4}\right].$$

A variância de $\hat{\theta}_2$ é dada por:

$$Var\left[\hat{\theta}_2\right] = Var\left[\frac{S+\sqrt{n}/2}{n+\sqrt{n}}\right] = \frac{Var(S)}{(n+\sqrt{n})^2} = \frac{n\theta(1-\theta)}{(n+\sqrt{n})^2}.$$

O erro quadrático médio de $\hat{\theta}_2$ é dado por:

$$EQM\left[\hat{\theta}_2\right] = \frac{n\theta(1-\theta)}{(n+\sqrt{n})^2} + \frac{n}{4(n+\sqrt{n})^2}\,\left[\theta^2 - \theta + \frac{1}{4}\right]$$

$$EQM\left[\hat{\theta}_2\right] = \frac{n}{(n+\sqrt{n})^2}\left[\theta - \theta^2 + \theta^2 - \theta + \frac{1}{4}\right]$$

$$EQM\left[\hat{\theta}_2\right] = \frac{n}{4(n+\sqrt{n})^2},$$

que independe de θ.

$$EQM\left[\hat{\theta}_2\right] = \frac{9}{4(9+\sqrt{9})^2} = \frac{9}{4\times 144} = \frac{1}{64}.$$

$$EQM\left[\hat{\theta}_1\right] = \frac{\theta(1-\theta)}{9}.$$

O valor máximo do EQM de $\hat{\theta}_1$ ocorre quando $\theta = \frac{1}{2}$. Este valor máximo é dado por:

$$EQM_{max} = \frac{1/4}{9} = \frac{1}{36}.$$

Quando

$$EQM\left[\hat{\theta}_1\right] = EQM\left[\hat{\theta}_2\right]?$$

$$\frac{\theta(1-\theta)}{9} = \frac{1}{64}.$$

$$\theta(1-\theta) = \frac{9}{64}$$

$$\theta^2 - \theta + \frac{9}{64} = 0$$

Assim

$$\Delta = (-1)^2 - 4\times 1\times\frac{9}{64} = 1 - \frac{36}{64} = 1 - \frac{9}{16} = \frac{7}{16}.$$

$$\sqrt{\Delta} = \frac{\sqrt{7}}{4}.$$

As raízes são dadas por:

$$c_1 = \frac{1 - \frac{\sqrt{7}}{4}}{2} = \frac{1}{2} - \frac{\sqrt{7}}{8} = 0,1692$$

$$c_2 = \frac{1 + \frac{\sqrt{7}}{4}}{2} = \frac{1}{2} + \frac{\sqrt{7}}{8} = 0,8307.$$

Vamos achar as raízes usando o **R**:

```
> sqrt(7)
[1] 2.645751
> c_1=1/2 - sqrt(7)/8;c_1
[1] 0.1692811
> c_2=1/2 + sqrt(7)/8;c_2
[1] 0.8307189
> a=1;b=-1;c=9/64
> aux=polyroot(c(c,b,a))
> aux=polyroot(c(c,b,a));aux
[1] 0.1692811+0i 0.8307189-0i
> abs(mod)
Erro: objeto 'mod' não encontrado
> abs(aux)
[1] 0.1692811 0.8307189
> r_1=abs(aux)[1];r_1;c_1
[1] 0.1692811
[1] 0.1692811
> r_2=abs(aux)[2];r_2;c_2
[1] 0.8307189
[1] 0.8307189
>
```

Represente graficamente o erro quadrático médio dos dois estimadores para $n = 9$.

Vamos fazer o gráfico usando o **R**:

E assim lembrando que a parábola representa $EQM[\theta_1]$ e a reta representa $EQM[\theta_2]$ temos que :

Para $0 < \theta < c_1$ usamos $\hat{\theta}_2]$.

Para $c_1 < \theta < c_2$ usamos $\hat{[}\theta_1]$.

Para $c_2 < \theta < 1$ usamos $\hat{\theta}_1$.

Para $\theta = c_1$ ou $\theta = c_2$ fica indiferente.

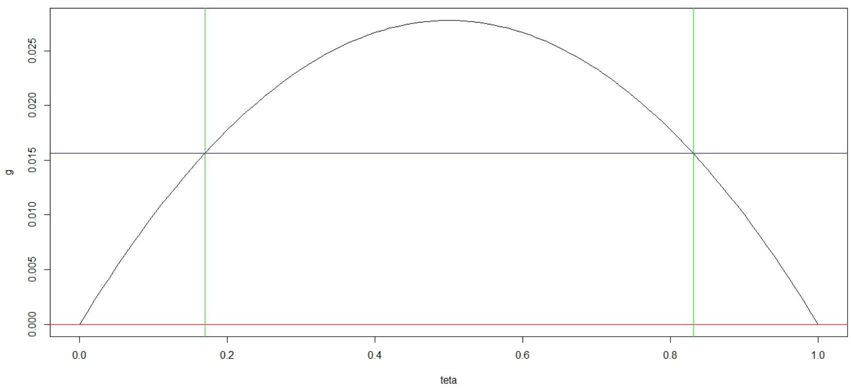

Figura 1: