Em Estatística, a palavra viés tem um significado preciso e tem a ver com a esperança da distribuição de um estimador.

Definição 32 (Estimador não-viesado)

Um estimador $\delta(\boldsymbol{X})$ de uma função $g(\theta)$ é dito ***não-viesado*** *se $E_\theta[\delta(\boldsymbol{X})] = g(\theta)$ para todo $\theta \in \Omega$. Um estimador que não atende a essa condição é dito viesado. O* ***viés*** *de δ é definido como $B_\delta(\theta) := E_\theta[\delta(\boldsymbol{X})] - g(\theta)$.*

Exemplo 12 (Tempos de falha de lâmpadas)

Lembremos do exemplo das lâmpadas da fábrica de Astolfo. Neste caso, não é difícil mostrar que $E[\hat{\theta}_{EMV}] = \frac{n}{n-1}\theta = 3\theta/2$. Desta forma, o viés do EMV é $B_{\hat{\theta}_{EMV}}(\theta) = 3\theta/2 - \theta = \theta/2$. É possível encontrar $\delta(\boldsymbol{X})$ não-viesado? Esse estimador é bom?

Estimadores não-viesados sempre?

Quando avaliamos estimadores, o erro quadrático médio e o viés são alguns *aspectos* a serem considerados, mas há um compromisso (*trade-off*) entre eles, de certa forma.

Observação 13 (Erro quadrático, variância e viés)

$$R(\theta, \delta) = \text{Var}_\theta(\delta) + [B_\delta(\theta)]^2 .$$

No exemplo das lâmpadas, é possível mostrar que $\delta_2(\boldsymbol{X}) = 1/S$ tem o menor EQM, mas tem viés $B_{\delta_2}(\theta) = \frac{n-2}{n-1}\theta = \theta/2$, assim como o EMV.

Estimador não-viesado da variância

A variância amostral como a temos definido até aqui é viesada. Uma pequena modificação leva a um estimador não viesado da variância.

Teorema 18 (Estimador não-viesado da variância)

Seja $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ *uma amostra aleatória, com* $E[X_1] = m$ *e* $\operatorname{Var}(X_1) = v < \infty$. *Então*

$$\delta_1(\boldsymbol{X}) = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$$

é um estimador não-viesado de v.

Prova: usar a igualdade

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

e usar a linearidade da esperança e o fato de que temos uma amostra aleatória.

Nem tudo são flores

Não-viesamento é uma característica desejável, mas nem sempre um estimador não-viesado (i) existe ou (ii) é um bom estimador.

- Não existência. Exemplo: $X_1, X_2, \ldots, X_n \sim \text{Bernoulli}(p)$, estimador para $\sqrt{p}$?
- Estimador não-viesado ruim: $X \sim \text{Geometrica}(p)$. Quais as propriedades do estimador não viesado $\delta(X)$?