# JACKKNIFE

**OBJETIVO:** estimar o viés e o erro padrão de um estimador.

**DEFINIÇÃO:** suponha que temos uma amostra $\underline{x} = (x_1, \cdots, x_n)$ e um estimador $\hat{\theta} = s(\underline{x})$. Queremos estimar o viés e o erro padrão de $\hat{\theta}$.

Então

$$\underline{x}_{(i)} = (x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n), \quad i = 1, \cdots, n \tag{1}$$

(vetor de x sem i-ésimo elemento) são denominados amostras **jackknife.**

Assim, a i-ésima amostra jackknife consiste dos dados com a i-ésima observação removida.

Seja,

$$\hat{\theta}_{(i)} = s(\underline{x}_{(i)}) \tag{2}$$

a i-ésima **réplica jackknife** de $\hat{\theta}$.

**Exemplo:**

$$\hat{\theta} = \bar{x} = \frac{x_1 + \cdots + x_n}{n}$$

$$\hat{\theta}_{(i)} = \bar{x}_{(i)} = \frac{x_1 + \cdots x_{i-1} + x_{i+1} + \cdots + x_n}{n-1} = \frac{n\hat{x}}{n-1} - \frac{x_i}{n-1} = \frac{n\bar{x} - x_i}{n-1}$$

$$\hat{\theta}_{(\bullet)} = \frac{\sum \hat{\theta}_{(i)}}{n} = \sum_{i=1}^{n} \bar{x}_{(i)} / n = \bar{x} \tag{3}$$

## ESTIMADOR JACKKNIFE DO VIÉS:

$$vi\hat{é}s_{jack} = (n-1)(\hat{\theta}_{(\bullet)} - \hat{\theta}) \qquad (4)$$

## ESTIMADOR JACKKNIFE DO ERRO PADRÃO:

$$e\hat{p}_{jack} = \left[\frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\theta}_{(\bullet)}{}^{2}\right]^{\frac{1}{2}} \qquad (5)$$

## Exemplo:

$$\hat{\theta} = \bar{x} = \frac{x_1 + \cdots + x_n}{n}$$

$$e\hat{p}(\bar{x}) = \sqrt{\frac{s^2}{n}}, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$e\hat{p}(\bar{x}) = \left[\frac{1}{n(n-1)}\sum(x_i - \bar{x})^2\right]^{\frac{1}{2}}$$

$$\bar{x}_{(i)} = \frac{n\bar{x}}{n-1} - \frac{x_i}{n-1} = \frac{1}{n-1}\left[n\bar{x} - x_i\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i - x_i\right]$$

$$\bar{x}_{(\bullet)} = \frac{1}{n}\sum_{i=1}^{n}\bar{x}_{(i)} = \bar{x}$$

Portanto,

$$e\hat{p}_{jack} = \left[\frac{n-1}{n}\sum_{i=1}^{n}(\bar{x}_{(i)} - \bar{x}_{(\bullet)})^2\right]^{\frac{1}{2}} = \left[\frac{n-1}{n}\sum_{i=1}^{n}(\frac{n\bar{x}}{n-1} - \frac{x_i}{n-1} - \bar{x})^2\right]^{\frac{1}{2}} = \left[\frac{n-1}{n}\sum(\frac{\bar{x}}{n-1} - \frac{x_i}{n-1})^2\right]^{\frac{1}{2}}$$

$$= \left[\frac{1}{n(n-1)}\sum(x_i - \bar{x})^2\right]^{\frac{1}{2}} = e\hat{p}(\bar{x})$$

Portanto, $e\hat{p}(\bar{x}) = e\hat{p}_{jack}$

O fator (n-1)/n é exatamente o que é necessário para tornar o $e\hat{p}_{jack}$ igual ao estimador não viciado do erro padrão da média.

De modo análogo, o estimador jackknife do viés dado pela expressão (4), é um múltiplo da média dos desvios jackknife

$$\hat{\theta}_{(i)} - \hat{\theta}, \quad i = 1,2,\cdots,n \qquad (6)$$

que são chamados valores de influencia jackknife.

$$\frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\theta}) = \sum_{i=1}^{n}\frac{\hat{\theta}_{(i)}}{n} - \hat{\theta} = \hat{\theta}_{(\bullet)} - \hat{\theta}$$

**Exemplo:**

$$\hat{\theta} = \bar{x} = \frac{x_1 + \cdots + x_n}{n}$$

$$\hat{\theta}_{(\bullet)} - \hat{\theta} = \bar{x} - \bar{x} = 0$$

Portanto, o estimador jackknife do vicio é zero.

**OBSERVAÇÃO:** $vi\acute{e}\hat{s}_{jack} = (n-1)(\hat{\theta}_{(\bullet)} - \hat{\theta})$

**PERGUNTA:** porque o fator (n-1)?

Considere $\hat{\sigma}^2 = \hat{\theta} = \frac{1}{n}\sum(x_i - \bar{x})^2$

$$E(\hat{\theta} - \sigma^2) = ?$$

$$\hat{\theta} = \frac{1}{n}\sum(x_i - \mu + \mu - \bar{x})^2 = \frac{1}{n}\left[\sum(x_i - \mu)^2 + n(\bar{x} - \mu)^2\right]$$

$$E(\hat{\theta}) = \frac{1}{n}\left[n\sigma^2 - n\frac{\sigma^2}{n}\right] = \frac{n-1}{n}\sigma^2 = \sigma^2 - \frac{\sigma^2}{n}$$

Portanto, $E(\hat{\theta} - \sigma^2) = -\frac{1}{n}\sigma^2$: viés do estimador da variância amostral

Pode-se mostrar que

$$vi\acute{e}\hat{s}_{jack} = (n-1)(\hat{\theta}_{(\bullet)} - \hat{\theta}) = -\frac{s^2}{n} = vici\hat{o}(\hat{\theta}), \quad onde \quad s^2 = \frac{\sum (x_i - \overline{x})^2}{n-1}$$