

MAT02025 - Amostragem 1

O erro quadrático médio

Rodrigo Citton P. dos Reis
citton.padilha@ufrgs.br

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA

Porto Alegre, 2021



O erro quadrático médio

O erro quadrático médio

- ▶ Para comparar um estimador enviesado com um estimador imparcial, ou dois estimadores com diferentes valores de enviesamento, um critério útil é o **erro quadrático médio (EQM)** da estimativa, medido a partir do valor da população que está sendo estimado.
- ▶ Formalmente

$$\begin{aligned}
 EQM(\hat{\theta}) &= E(\hat{\theta} - \theta)^2 \\
 &= E\{[\hat{\theta} - E(\hat{\theta})] + [E(\hat{\theta}) - \theta]\}^2 \\
 &= E\{[\hat{\theta} - E(\hat{\theta})]^2 + 2[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta] + [E(\hat{\theta}) - \theta]^2\} \\
 &= E\{[\hat{\theta} - E(\hat{\theta})]^2\} + 2[E(\hat{\theta}) - \theta]E[\hat{\theta} - E(\hat{\theta})] + E[E(\hat{\theta}) - \theta]^2 \\
 &= \text{Var}(\hat{\theta}) + 2[E(\hat{\theta}) - \theta][E(\hat{\theta}) - E(\hat{\theta})] + [B(\hat{\theta})]^2 \\
 &= \text{Var}(\hat{\theta}) + [B(\hat{\theta})]^2.
 \end{aligned}$$

O erro quadrático médio

- Note que se um estimador $\hat{\theta}$ é não enviesado para θ , então

$$EQM(\hat{\theta}) = \text{Var}(\hat{\theta}) + [0]^2 = \text{Var}(\hat{\theta}) = \sigma_{\hat{\theta}}^2.$$

- No exemplo da aula passada, considerando o parâmetro populacional de interesse como a média, μ , temos $EQM(\hat{\mu}) = \sigma_{\hat{\mu}}^2 + B^2$, em que $B = m - \mu$.

O erro quadrático médio

- ▶ O uso do EQM como critério de precisão de um estimador equivale a considerar equivalentes duas estimativas que têm o mesmo erro quadrático médio.
- ▶ Isso não é inteiramente verdadeiro, porque a distribuição de frequência de erros $(\hat{\mu} - \mu)$ não será a mesma para dois estimadores, caso eles apresentem viéses de valores diferentes.
- ▶ Entretanto, **Hansen, Hurwitz e Madow (1953)**¹ mostraram que se B/σ for menor que cerca de 0,5, as duas distribuições de frequência são quase idênticas em relação aos erros **absolutos** $|\hat{\mu} - \mu|$ de tamanhos diferentes.

¹Hansen, M. H., Hurwitz, W. N. e Madow, W. G. (1953) **Sample Survey methods and theory**, John Wiley & Sons, Nova York, Vol. I, pg. 58.

O erro quadrático médio

- ▶ Mais uma vez suponha que $\hat{\mu}$ tem uma distribuição aproximadamente normal com média $m = E(\hat{\mu})$ e desvio padrão $\sigma = \sigma_{\hat{\mu}}$. Ainda, denote $EQM = EQM(\hat{\mu}) = \sigma^2 + B^2$.
- ▶ Então

$$\begin{aligned} \Pr\left(|\hat{\mu} - \mu| \geq k\sqrt{EQM}\right) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu+k\sqrt{EQM}}^{\infty} e^{-(\hat{\mu}-m)^2/2\sigma^2} d\hat{\mu} \\ &\quad + \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu-k\sqrt{EQM}} e^{-(\hat{\mu}-m)^2/2\sigma^2} d\hat{\mu}. \end{aligned}$$

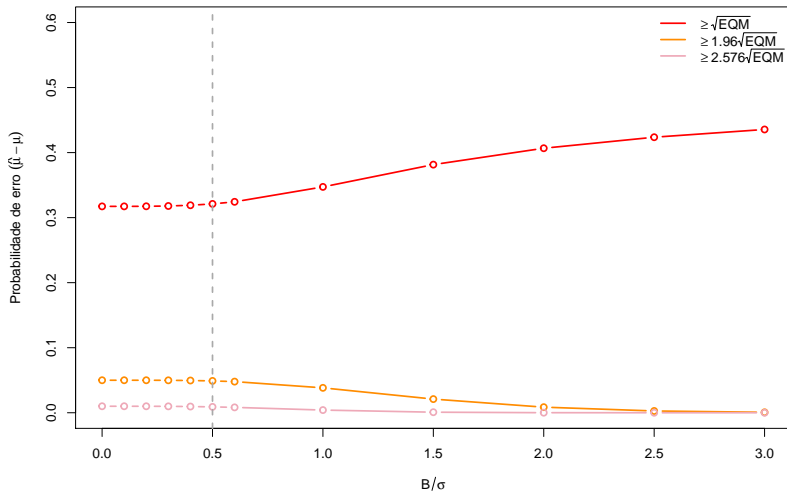
O erro quadrático médio

Table 1: Proporção de casos em que o valor verdadeiro, μ , não está incluído no intervalo $\hat{\mu} \pm k\sqrt{EQM}$, para diferentes níveis de viés em $\hat{\mu}$

B/σ	$\geq \sqrt{EQM}$	$\geq 1,96\sqrt{EQM}$	$\geq 2,576\sqrt{EQM}$
0.0	0.32	0.050	0.0100
0.1	0.32	0.050	0.0100
0.2	0.32	0.050	0.0100
0.3	0.32	0.050	0.0098
0.4	0.32	0.050	0.0095
0.5	0.32	0.049	0.0090
0.6	0.32	0.048	0.0083
1.0	0.35	0.038	0.0041
1.5	0.38	0.021	0.0008
2.0	0.41	0.009	0.0001
2.5	0.42	0.003	-
3.0	0.44	0.001	-

O erro quadrático médio

Efeito do viés sobre a probabilidade de um erro maior que $k\sqrt{EQM}$



Comentários

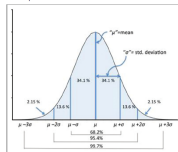
- ▶ Esses resultados, para muitos propósitos práticos, concordam com as interpretações baseadas nos múltiplos correspondentes do desvio padrão quando uma estimativa não enviesada é usada.
 - ▶ Ou seja, quando $\sqrt{EQM} = \sqrt{\sigma^2 + B^2} = \sqrt{\sigma^2 + 0} = \sigma$.
- ▶ Da aula passada, temos

MAT02025 - Amostragem 1
└ O uso da distribuição normal

O uso da distribuição normal

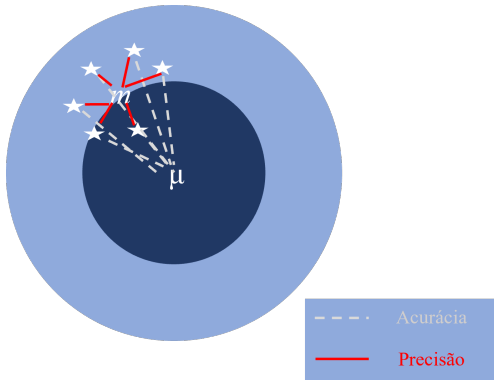
... a partir das **propriedades da curva normal**, as chances são

- ▶ 0,32 (cerca de 1 em 3) que o erro absoluto $|\hat{\theta} - \theta|$ excede $\sigma_{\hat{\theta}}$.
- ▶ 0,05 (1 em 20) que o erro absoluto $|\hat{\theta} - \theta|$ excede $1,96\sigma_{\hat{\theta}} \approx 2\sigma_{\hat{\theta}}$.
- ▶ 0,01 (1 em 100) que o erro absoluto $|\hat{\theta} - \theta|$ excede $2,58\sigma_{\hat{\theta}}$.

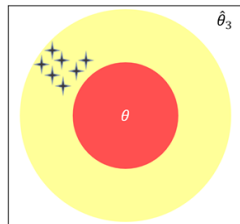
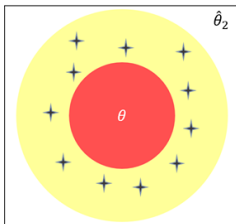
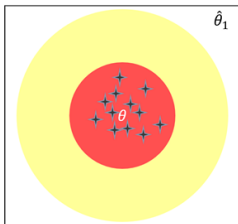


Comentários

- ▶ Devido à dificuldade de garantir que nenhum viés insuspeitado entre nas estimativas, geralmente falaremos da **precisão** de uma estimativa em vez de sua **acurácia** (exatidão).
- ▶ A acurácia se refere ao tamanho dos desvios da verdadeira média μ , enquanto a precisão se refere ao tamanho dos desvios da média m obtida pela aplicação repetida do procedimento de amostragem.



Comentários



► **Pergunta:** o que podemos concluir sobre $\hat{\theta}_1$, $\hat{\theta}_2$ e $\hat{\theta}_3$?

Exercícios

Para casa

- ▶ **Atividade de avaliação I.**

Próxima aula

- ▶ Amostragem aleatória simples;
- ▶ Definições e notação.

Por hoje é só!

Bons estudos!

