

AMOSTRAGEM

Unidade 6

Estimação de Razões Sob AAS

Razões Populacionais como Parâmetros de Interesse

Muitas vezes, o parâmetro populacional de interesse é uma razão entre dois totais ou médias.

Por exemplo, a densidade de pessoas por cômodo em domicílios é dada por:

$$\frac{\text{Número de pessoas por domicílio}}{\text{Número de cômodos por domicílio}} = \frac{\sum_{i \in U} y_i}{\sum_{i \in U} x_i} = \frac{Y}{X} = \frac{\bar{Y}}{\bar{X}} = R$$

Exercício 6.1: Identifique e defina duas razões que você já tenha utilizado em seu trabalho.

Razão de Médias vs Média de Razões

Um cuidado importante aqui é entender a diferença de parâmetros de interesse.

O parâmetro R não é a média $\bar{R} = \sum_{i \in U} R_i / N$ das razões calculadas por unidade ($R_i = y_i / x_i$).

Exemplo 6.1: Considere as densidades de habitantes por unidade da federação conforme o Censo 2010.

Neste exemplo, vemos a discrepância grande que pode ocorrer entre as duas quantidades, quando os valores de R_i são muito dispersos.

Unidade da Federação	Densidade (Hab/km ²)
Rondônia	6,6
Acre	4,5
Amazonas	2,2
Roraima	2,0
Pará	6,1
Amapá	4,7
Tocantins	5,0
Maranhão	19,8
Piauí	12,4
Ceará	56,8
Rio Grande do Norte	60,0
Paraíba	66,7
Pernambuco	89,6
Alagoas	112,3
Sergipe	94,4
Bahia	24,8
Minas Gerais	33,4
Espírito Santo	76,3
Rio de Janeiro	365,2
São Paulo	166,3
Paraná	52,4
Santa Catarina	65,3
Rio Grande do Sul	39,8
Mato Grosso do Sul	6,9
Mato Grosso	3,4
Goiás	17,7
Distrito Federal	444,1
Brasil (Média das razões)	68,1
Brasil (razão de médias)	22,4

Fonte: IBGE, Censo Demográfico 2010.

Estimação de Razões

Para estimar razões como R , o estimador “natural” é a **razão dos estimadores de total**:

$$\hat{R} = \frac{\sum_{i \in s} w_i y_i}{\sum_{i \in s} w_i x_i} = \frac{\hat{Y}_{HT}}{\hat{X}_{HT}}$$

Sob AAS, este estimador simplifica para $\hat{R} = \frac{\sum_{i \in s} y_i}{\sum_{i \in s} x_i} = \frac{\bar{y}}{\bar{x}}$.

Notas

1. Tanto o numerador como o denominador do estimador \hat{R} da razão R podem variar com a amostra selecionada, s .
2. Apesar de termos estimadores não viciados para as médias populacionais \bar{Y} e \bar{X} , em geral, $E(Z/W) \neq E(Z)/E(W)$, e portanto $E(\hat{R}) = E(\bar{y}/\bar{x}) \neq E(\bar{y})/E(\bar{x}) = \bar{Y}/\bar{X} = R$.
3. Logo \hat{R} é um estimador viciado de R .
4. O vício de \hat{R} é pequeno e chamado de “vício técnico”, para distingui-lo de outros vícios tais como não resposta e outros. Para n grande, este vício é desprezível.

Analisando o Vício do Estimador da Razão sob AAS

Sejam $\Delta y = \bar{y} - \bar{Y}$, e $\Delta x = \bar{x} - \bar{X}$ os erros das estimativas de médias populacionais para as variáveis x e y .

Lembrando que $DP_{AAS}(\bar{y}) = (1-f) S_y / \sqrt{n}$, segue-se que Δy e Δx são de ordem $n^{-1/2}$ (a ordem de grandeza do desvio padrão).

Analisando o Vício do Estimador da Razão sob AAS

Portanto

$$\begin{aligned}
 \frac{\bar{y}}{\bar{x}} &= \frac{\bar{Y} + \Delta y}{\bar{X} + \Delta x} \\
 &= \frac{\bar{Y}}{\bar{X}} \left(1 + \frac{\Delta y}{\bar{Y}} \right) \left(1 + \frac{\Delta x}{\bar{X}} \right)^{-1} \\
 &= \frac{\bar{Y}}{\bar{X}} \left(1 + \frac{\Delta y}{\bar{Y}} \right) \left[1 - \frac{\Delta x}{\bar{X}} + \left(\frac{\Delta x}{\bar{X}} \right)^2 + \dots \right] \\
 &= \frac{\bar{Y}}{\bar{X}} \left(1 + \frac{\Delta y}{\bar{Y}} - \frac{\Delta x}{\bar{X}} + \text{termos em } \Delta^2 \right) .
 \end{aligned}$$

Analisando o Vício do Estimador da Razão sob AAS

$$E(\hat{R}) = E(\bar{y}/\bar{x})$$

$$\begin{aligned} \text{Logo} \quad &= R \left[1 + \frac{E(\Delta y)}{\bar{Y}} - \frac{E(\Delta x)}{\bar{X}} + \text{termos em } E(\Delta^2) \right] \\ &= R + \text{termos de ordem } n^{-1} \end{aligned}$$

porque $E(\Delta y) = E(\Delta x) = 0$.

➔ O vício é de ordem n^{-1} e fica menor à medida que o tamanho da amostra cresce, daí o nome de *vício técnico*.

Vícios desse tipo são geralmente desprezados na prática, a menos que se tenha amostras **muito pequenas**.

O EQM de \hat{R} sob AAS

Como \hat{R} tem um pequeno vício técnico, avaliamos seu Erro Quadrático Médio (EQM) sob AAS, ao invés da variância:

$$\begin{aligned} EQM_{AAS}(\hat{R}) &= E_{AAS}(\hat{R} - R)^2 \\ &= E_{AAS} \left[R^2 \left(\frac{\Delta y}{\bar{Y}} - \frac{\Delta x}{\bar{X}} \right)^2 \right] + \text{termos em } \Delta^3 \\ &\cong R^2 E_{AAS} \left(\frac{\bar{y} - \bar{Y}}{\bar{Y}} - \frac{\bar{x} - \bar{X}}{\bar{X}} \right)^2 \\ &= \frac{R^2}{\bar{Y}^2} E_{AAS}(\bar{y} - R\bar{x})^2 \\ &= \frac{(1-f)}{n \bar{X}^2} \sum_{i \in U} (y_i - R x_i)^2 / (N-1) . \end{aligned}$$

O EQM de \hat{R} sob AAS

Um estimador desse EQM que tem vício de ordem n^{-1} é:

$$\hat{V}_{AAS}(\hat{R}) = \frac{1-f}{n\bar{X}^2} \frac{1}{n-1} \sum_{i \in s} (y_i - \hat{R}x_i)^2$$

Veja Cochran (1977).

Resumo da Estimação de R sob AAS

Portanto, sob AAS temos a seguinte estrutura para estimação de razões:

(1) \hat{R} estima R e $E_{AAS}(\hat{R}) = R + O(n^{-1})$;

$$(2) EQM_{AAS}(\hat{R}) = \frac{1-f}{n\bar{X}^2} \frac{\sum_{i \in U} (y_i - Rx_i)^2}{N-1} + O(n^{-3/2})$$

$$(3) \hat{V}_{AAS}(\hat{R}) = \frac{1-f}{n\bar{X}^2} \frac{1}{n-1} \sum_{i \in s} (y_i - \hat{R}x_i)^2$$

(4) $(\hat{R} - R) / \sqrt{\hat{V}_{AAS}(\hat{R})} \approx N(0;1)$ para n grande.