1. Estatística Descritiva

Tabelas de Frequências

a. Dados qualitativos ou quantitativos quando os valores se repetem

Frequência Absoluta Simples (Fj) – Número de vezes em que cada valor distinto *j* da variável observada se repete (*j=1,2,…k*). Tem-se que:

$$\sum_{j=1}^{k} F_j = n$$

Frequência relativa simples (fj) – Proporção de vezes que cada valor distinto *j da variável observada se repete. Obtém-se* $f_j = \frac{F_j}{n}$ com

$$\sum_{j=1}^{k} f_j = \sum_{j=1}^{k} \frac{F_j}{n} = \frac{n}{n} = 1$$

b. Dados agrupados em classes ou intervalos:

Amplitude do intervalo: $ai = Li - li$

Ponto médio do intervalo: $C_i = \frac{L_i + l_i}{2}$

Densidade de frequência: $Di = fi/ai$

Medidas de Localização

Média Aritmética (ou simplesmente média)

Simples	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Ponderada	$\bar{X}_p = \frac{\sum_{i=1}^{n} w_i * x_i}{\sum_{i=1}^{n} w_i} = \sum_{i=1}^{n} \alpha_i * x_i \quad \text{com} \quad \alpha_i = \frac{w_i}{\sum_{i=1}^{n} w_i} \quad \text{e} \quad \sum_{i=1}^{n} \alpha_i = 1$
Dados agrupados	$\bar{X}^* = \frac{n_1 v_1 + \ldots + n_m v_m}{n} = \sum_{g=1}^{m} f_g v_g$
Dados classificados (em intervalos de classe)	$\bar{X}^* = \frac{F_1 C_1 + \ldots + F_m C_m}{n} = \sum_{j=1}^{m} f_j C_j$

Média Geométrica

Simples	$\bar{X}_G = \sqrt[n]{x_1 . x_2 . \ldots . x_n} = \left(\prod_{i=1}^{n} x_i \right)^{1/n}$
Ponderada	$\bar{X}_G = \sqrt[n]{x_1^{w_1} . x_2^{w_2} . \ldots . x_n^{w_n}} = \left(\prod_{i=1}^{n} x_i^{w_i} \right)^{\frac{1}{\sum w_i}} = \prod_{i=1}^{n} x_i^{\alpha_i} \quad \text{com}$ $\alpha_i = \frac{w_i}{\sum_{i=1}^{n} w_i} \quad \text{e} \quad \sum_{i=1}^{n} \alpha_i = 1$

Relação entre a média geométrica e aritmética - $\boldsymbol{ln\, \bar{X}_G = ln\, \bar{X}}$

Mediana – Seja x'_n a observação de ordem *n*

Dados não classificados – Número ímpar de observações	$Me = x'_{\frac{n+1}{2}}$
Dados não classificados – Número par de observações	$Me = \dfrac{x'_{\frac{n}{2}} + x'_{\frac{n}{2}+1}}{2}$
Data classificados	$Me = l_i(Me) + \dfrac{0{,}5 - cum\, f(Me-1)}{f(Me)} a(Me)$

Quantis – *exemplo para o 1º quartil*

1º Quartil – dados não classificados	$Q_1 = x'_{\frac{n+1}{4}}$
1º Quartil – dados classificados	$Q_1 = l_i(Q_1) + \dfrac{0{,}25 - cum\, f(Q_1 - 1)}{f(Q_1)} a(Q_1)$

Moda

Dados classificados	$Mo = l(Mo) + \frac{f(Mo)-f(Mo-1)}{(f(Mo)-f(Mo-1))+(f(Mo)-f(Mo+1))}(L(Mo) - l(Mo))$ $Mo = l_i(Mo) + \dfrac{f(Mo+1)}{f(Mo-1)+f(Mo+1)} a(Mo)$

Medidas de Dispersão

Amplitude do intervalo Inter-Quartis	$IQ = Q_3 - Q_1$
Desvio Absoluto Médio	*Dados não agrupados:* $D_x = \dfrac{\sum_{i=1}^{n} \lvert x_i - \bar{x}\rvert}{n}$ *Dados agrupados/classificados:* $D_x = \dfrac{\sum_{i=1}^{n} n_i \lvert v_i - \bar{x}\rvert}{n} = \sum_{i=1}^{n} f_i \lvert v_i - \bar{x}\rvert$
Desvio Padrão	*Dados não agrupados:* $s_x = \sqrt{\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$ *Dados agrupados/classificados:* $s_x = \sqrt{\dfrac{\sum_{i=1}^{n} n_i (v_i - \bar{x})^2}{n}} = \sqrt{\sum_{i=1}^{n} f_i (v_i - \bar{x})^2}$
Variância	$s_x^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$ ou $s_x^2 = \dfrac{\sum_{i=1}^{n} n_i (v_i - \bar{x})^2}{n} = \sum_{i=1}^{n} f_i (v_i - \bar{x})^2$

***Nota**: Para dados em classes, o símbolo v_i é substituído por C_i.*

Coeficiente de Variação	$CV_x = \frac{s_x}{\bar{\bar{x}}} = \frac{\sigma_x}{\mu}$

Simetria / Assimetria da Distribuição

Distribuição simétrica: $\bar{x} = Me = Mo$

Distribuição assimétrica à esquerda (assimetria positiva): $\bar{x} > Me > Mo$

Distribuição assimétrica à direita (assimetria negative): $\bar{x} < Me < Mo$

Medidas de Concentração

Índice de Gini	$IG = \frac{\sum_{i=1}^{n-1}(\text{cum}\, f_i - \text{cum}\, y_i)}{\sum_{i=1}^{n-1}\text{cum}\, f_i} = \frac{\sum_{i=1}^{n-1}(p_i - q_i)}{\sum_{i=1}^{n-1} p_i}$

2. Taxas de Variação e Índices

Variação absoluta

De um período	$\Delta x_{t+1,t} = x_{t+1} - x_t$
De k períodos	$\Delta x_{t+k,t} = x_{t+k} - x_t$

Variação absoluta média

k períodos	$\Delta_m x_{t+k,t} = \frac{\Delta x_{t+k,t}}{k} = \frac{x_{t+k} - x_t}{k}$

Variação relativa – *Taxa de crescimento. Taxa de variação*

De um período	$r_{t+1,t} = \frac{x_{t+1} - x_t}{x_t}$ ou $r_{t+1,t} = \frac{x_{t+1,t}}{x_t} - 1$
De k períodos	$\delta_{t+k,t} = \frac{x_{t+k} - x_t}{x_t}$

Variação relativa média – *Taxa média de crescimento em k períodos* - $r_{t+k,t}$: é a variação proporcional média por unidade de tempo da variável x entre t e $t+k$

$$r_{t+k,t} = \left(1 + \delta_{t+k,t}\right)^{\frac{1}{k}} - 1$$

$$r_{t+k,t} = \left[\left(1 + r_{t+1,t}\right)\left(1 + r_{t+2,t+1}\right)\ldots\left(1 + r_{t+k,t+k-1}\right)\right]^{\frac{1}{k}} - 1$$

$$r_{t+k,t} = \left(\frac{x_{t+k}}{x_t}\right)^{\frac{1}{k}} - 1$$

Taxa de variação homóloga

$$h_{t,s} = \frac{x_{t,s} - x_{t-1,s}}{x_{t-1,s}}$$

t: período
s: subperíodo

Índices simples

Considerando uma série de valores: x_0 , x_1, x_2, … , x_t:

- ***A série de índices de base móvel*** $i_{1,0}$, $i_{2,1}$, … , $i_{t,t-1}$ obtém-se:

$$i_{1,0} = \frac{x_1}{x_0},\; i_{2,1} = \frac{x_2}{x_1},\; \ldots\;,\; i_{t,t-1} = \frac{x_t}{x_{t-1}}$$

- ***A série de índices de base fixa*** $i_{1,0}$, $i_{2,0}$, … , $i_{t,0}$ obtém-se:

$$i_{1,0} = \frac{x_1}{x_0},\; i_{2,0} = \frac{x_2}{x_0},\; \ldots\;,\; i_{t,0} = \frac{x_t}{x_0}$$

Relação com taxas de variação

Índice de base móvel e taxa de variação no período: $\boldsymbol{i_{t,t-1} = 1 + r_{t,t-1}}$

Índice de base fixa e taxa de variação em "multi-períodos": $\boldsymbol{i_{t+k,t} = 1 + \delta_{t+k,t}}$

Circularidade dos índices

Passagem de base fixa para base móvel	$\frac{\boldsymbol{i_{t,0}}}{\boldsymbol{i_{t-1,0}}} = \boldsymbol{i_{t,t-1}}$ porque $\frac{\frac{x_t}{x_0}}{\frac{x_{t-1}}{x_0}} = \frac{x_t}{x_{t-1}}$
Passagem de base móvel para base fixa	$\boldsymbol{i_{1,0} * i_{2,1} = i_{2,0}}$ porque $\frac{\boldsymbol{x_1}}{\boldsymbol{x_0}} * \frac{\boldsymbol{x_2}}{\boldsymbol{x_1}} = \frac{\boldsymbol{x_2}}{\boldsymbol{x_0}}$

Reversibilidade dos índices

$$\boldsymbol{i_{t,0} = \frac{1}{i_{0,t}}} \quad \text{porque} \quad \boldsymbol{\frac{x_t}{x_0} = \frac{1}{\frac{x_0}{x_t}}}$$

Mudança de base

$$\boldsymbol{\frac{i_{t,0}}{i_{b,0}} = i_{t,b}} \quad \text{porque} \quad \boldsymbol{\frac{\frac{x_t}{x_0}}{\frac{x_b}{x_0}} = \frac{x_t}{x_b}}$$

Índices agregativos

Índice de uma variável produto (quociente) de outras variáveis

Se $A = B.C$, *então:* $I_A = I_B.I_C$
E, sendo $I_k = 1 + \delta_k$, com $k = A, B, C$,
$\delta_A = \delta_B + \delta_C + \delta_B\delta_C$ e, para valores pequenos, pode fazer-se $\delta_A \approx \delta_B + \delta_C$

Em geral, com índices simples, verifica-se:

$$\text{Índice de valor} = \text{Índice de preços} \times \text{Índice de quantidades}$$

Com índices agregativos a relação só é verdadeira para casos particulares, sendo: $I_{VALOR} = P^P \times Q^L = P^L \times Q^P$

Índice de valor:

$$I_{VALOR} = \frac{\sum_{j=1}^{m} p_t^j \times q_t^j}{\sum_{j=1}^{m} p_0^j \times q_0^j}$$

Índices de

	Laspeyres	**Paasche**
Preços	$P^L = \frac{\sum_{j=1}^{m} p_t^j \times q_0^j}{\sum_{j=1}^{m} p_0^j \times q_0^j}$	$P^P = \frac{\sum_{j=1}^{m} p_t^j \times q_t^j}{\sum_{j=1}^{m} p_0^j \times q_t^j}$
Quantidades	$Q^L = \frac{\sum_{j=1}^{m} p_0^j \times q_t^j}{\sum_{j=1}^{m} p_0^j \times q_0^j}$	$Q^P = \frac{\sum_{j=1}^{m} p_t^j \times q_t^j}{\sum_{j=1}^{m} p_t^j \times q_0^j}$

Índice Laspeyres como média ponderada de índices

$$P^L = \sum_{j=1}^{m} \alpha_j \frac{p_1^j}{p_0^j} \quad \text{com} \quad \alpha_j = \frac{p_0^j q_0^j}{\sum_{j=1}^{m} p_0^j q_0^j}$$

sendo α_j o coeficiente orçamental no ano base.

3. Associação e Relação entre Variáveis

Covariância entre as variáveis *x* e *y*

$$S_{YX} = \frac{\sum_{j=1}^{N}(x_j - \bar{X})(y_j - \bar{Y})}{n}$$

Coeficiente de correlação linear entre as variáveis *x* e *y*

$$r_{YX} = \frac{S_{YX}}{S_X \times S_Y} = \frac{\sum_{j=1}^{N}(x_j - \bar{X})(y_j - \bar{Y})}{\sqrt{\sum_{j=1}^{N}(x_j - \bar{X})^2 \times \sum_{j=1}^{N}(y_j - \bar{Y})^2}}$$

A recta de regressão:

$$Y_j = b_0 + b_1 X_j + \varepsilon_j$$

Os parâmetros da recta:

$$b_0 = \bar{Y} - b_1 \bar{X}$$

$$b_1 = \frac{S_{YX}}{S_X^2}$$