

Estadística Descritiva 1

Prof. Lorí Viali, Dr.

viali@ufrgs.br

<http://www.ufrgs.br/~viali/>



D e s t a t

Departamento de Estatística

Mat2007:

Métodos Descritivos



[Apresentação](#)

[Congressos](#)

[Contato](#)

[Currículo](#)

[Cursos](#)

[Disciplinas](#)

[Galeria](#)

[Horário](#)

[Orientações](#)

[Palestras](#)

[Projetos](#)

[Publicações](#)

[Statweb](#)



Um local de apoio ao ensino de Estatística e de Probabilidade

Se o seu experimento precisa de estatística é porque você poderia tê-lo feito melhor.

Ernest Rutherford (1871-1937)

Segunda-Feira - 23/8/2004 - 20:46



Sir Maurice George Kendall (1907 - 1983)

Melhor visualizado na resolução 1024x768

Conceitos

Básicos



- Coleção de números = estatísticas

- ✓ **O número de carros vendidos no país aumentou em 30%.**
- ✓ **A taxa de desemprego atinge, este mês, 7,5%.**
- ✓ **As ações da Telebrás subiram R\$ 1,5, hoje.**
- ✓ **Resultados do Carnaval no trânsito: 145 mortos, 2430 feridos.**



Estatística: uma definição

A ciência de coletar, organizar, apresentar, analisar e interpretar dados numéricos com o objetivo de tomar melhores decisões.



Estatística (divisão)

Descritiva

Os procedimentos usados para organizar, resumir e apresentar dados numéricos.

Indutiva

A coleção de métodos e técnicas utilizados para estudar uma população baseado em amostras probabilísticas desta população.



POPULAÇÃO



Uma coleção de todos os possíveis elementos, objetos ou medidas de interesse.



CENSO

Um levantamento efetuado sobre toda uma população é denominado de levantamento censitário ou simplesmente censo.



AMOSTRA



Uma porção ou parte de
uma população de interesse.



AMOSTRAGEM

O processo de escolha de uma amostra da população é denominado de amostragem.



PROBABILIDADE
(Matemática)

ESTATÍSTICA
(Matemática
Aplicada)

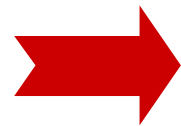
Univariada

Multivariada



P
R
O
B
A
B
I
L
I
D
A
D
E

**POPULAÇÃO
(Censo)**



Erro

Inferência

**AMOSTRA
(Amostragem)**



Estatística Descritiva

Probabilidade

Amostragem

Estatística Indutiva



Estatística x Probabilidade

Faces	Probabilidades
1	$1/6$
2	$1/6$
3	$1/6$
4	$1/6$
5	$1/6$
6	$1/6$
Total	1

Faces	Freqüências
1	15
2	18
3	23
4	25
5	22
6	17
Total	120



Arredondamento

Todo arredondamento é um erro.

O erro deve ser evitado ou então minimizado.



Arredondamento

Regra básica:

Arredondar sempre para
o mais próximo.



Exemplos:

1,456 \rightarrow 1,46 1,454 \rightarrow 1,45

1,475 $\xrightarrow{\text{É ímpar}}$ 1,48
Aumenta

1,485 $\xrightarrow{\text{É par}}$ 1,48
Não aumenta



V
A
R
I
Á
V
E
I
S

QUALITATIVAS

QUANTITATIVAS

NOMINAL

ORDINAL

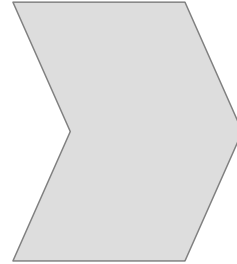
DISCRETA

CONTÍNUA



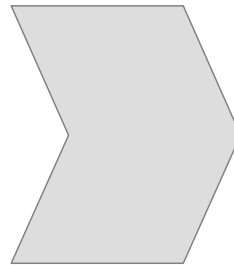
Variável Qualitativa

NOMINAL



Sexo
Religião
Estado civil
Curso

ORDINAL

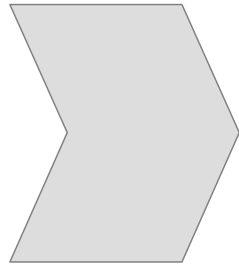


Conceito
Grau de Instrução
Mês
Dia da semana



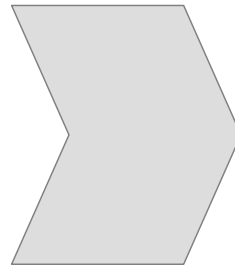
Variável Qualitativa

DISCRETA



Número de faltas
Número de irmãos
Número de acertos

CONTÍNUA



Altura
Área
Peso
Volume



Análise de Dados

Pequenos Conjuntos



ESTATÍSTICA DESCRITIVA

- **Organização;**
- **Resumo;**
- **Apresentação.**

Conjunto de dados:

↳ **Amostra**

ou

↳ **População**



Um conjunto de dados é resumido de acordo com as seguintes características:

**Amostra
ou
População**

- **Tendência central**
- **Dispersão ou variabilidade**
- **Assimetria (distorção)**
- **Achatamento ou curtose**



Tendência ou Posição Central

(a) As
médias

S
i
m
p
l
e
s

- Aritmética
- Geométrica
- Harmônica
- Quadrática
- Interna



A média Aritmética (*mean*)

$$\begin{aligned}\bar{X} &= \frac{x_1 + x_2 + \dots + x_n}{n} = \\ &= \frac{1}{n} \sum x_i = \frac{\sum x_i}{n}\end{aligned}$$



A média Geométrica

$$m_g = \sqrt[n]{X_1 \cdot X_2 \cdot \dots \cdot X_n} =$$
$$= \sqrt[n]{\prod X_i}$$



A média Harmônica

$$\begin{aligned} m_h &= \frac{1}{\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_n}} = \\ &= \frac{n}{\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_n}} = \frac{n}{\sum \frac{1}{X_i}} \end{aligned}$$



A média Quadrática

$$m_q = \frac{x_1^2 + x_2^2 + \dots + x_n^2}{n} =$$
$$= \frac{\sum x_i^2}{n}$$



A média Interna (*trimmed mean*)

É a mesma média aritmética só que aplicada sobre o conjunto onde uma parte dos dados (extremos) é descartada.



Exemplo

Médias

Conjuntos

4

6

5

4,9

4,8

1

9

5

3

1,8



Relação entre as médias

Dado um conjunto de dados qualquer, as médias aritmética, geométrica e harmônica mantêm a seguinte relação:

$$\bar{x} \geq m_g \geq m_h$$



Tendência ou Posição Central

(a) As
médias

P
o
n
d
e
r
a
d
a
s

■ Aritmética

■ Geométrica

■ Harmônica

■ Quadrática



A média Aritmética Ponderada

$$m_{ap} = \frac{X_1 \cdot W_1 + X_2 \cdot W_2 + \dots + X_k \cdot W_k}{W_1 + W_2 + \dots + W_k} =$$
$$= \frac{\sum X_i \cdot W_i}{\sum W_i}$$



A média Geométrica Ponderada

$$m_{gp} = \sum w_i \sqrt[w_i]{x_1^{w_1} \cdot x_2^{w_2} \cdot \dots \cdot x_k^{w_k}} =$$
$$= \sum w_i \sqrt[w_i]{\prod x_i^{w_i}}$$



A média Harmônica Ponderada

$$\begin{aligned} m_{hP} &= \frac{w_1 + w_2 + \dots + w_k}{\frac{w_1}{x_1} + \frac{w_2}{x_2} + \dots + \frac{w_k}{x_k}} = \\ &= \frac{\sum w_i}{\sum \frac{w_i}{x_i}} \end{aligned}$$



A média Quadrática Ponderada

$$m_{qp} = \frac{w_1 x_1^2 + w_2 x_2^2 + \dots + w_k x_k^2}{w_1 + w_2 + \dots + w_k} = \frac{\sum w_i x_i^2}{\sum w_i}$$



Exemplo

Produtos	p_{01}	p_{02}	q
Carne	4,80	5,52	5 kg
Cana	5,20	4,94	1 l
Ceva	0,80	0,92	12 lt
Pão	1,50	2,10	2 u
Total	--	--	--



P	p_{01}	p_{02}	α	$p(0,t)$
1	4,80	5,52	0,58	1,15
2	5,20	4,94	0,12	0,95
3	0,80	0,92	0,23	1,15
4	1,50	2,10	0,07	1,40
Total	--	--	1,00	--



Média aritmética ponderada dos relativos (aumentos) será:

$$m_{ap} = \frac{1,15 \cdot 0,58 + 0,95 \cdot 0,12 + 1,15 \cdot 0,23 + 1,40 \cdot 0,07}{0,57 + 0,12 + 0,23 + 0,07} = 1,1431 = 114,31\%$$

Por este critério o aumento foi de 14,31%.



Média geométrica ponderada dos relativos (aumentos) será:

$$\begin{aligned} m_{gp} &= \sqrt[1]{1,15^{0,58} 0,95^{0,12} 1,15^{0,23} 1,40^{0,07}} = \\ &= 1,15^{0,58} 0,95^{0,12} 1,15^{0,23} 1,40^{0,07} = \\ &= 1,1390 = 113,90\% \end{aligned}$$

Por este critério o aumento foi de
13,90%.



Média harmônica ponderada dos relativos (aumentos) será:

$$m_{hP} = \frac{1}{\frac{0,58}{1,15} + \frac{0,12}{0,95} + \frac{0,23}{1,15} + \frac{0,07}{1,40}} = 1,1348 = 113,48\%$$

Por este critério o aumento foi de 13,48%.



Tendência ou Posição Central

(b) A mediana (*median*)

É o valor que separa o conjunto em dois subconjuntos do mesmo tamanho.

$$m_e = [x_{(n/2)} + x_{(n/2)+1}]/2 \text{ se "n" é par}$$

$$m_e = x_{(n+1)/2} \text{ se "n" é ímpar}$$



Separatrizes

A idéia de repartir o conjunto de dados pode ser levada adiante. Se ele for repartido em 4 partes tem-se os **QUARTIS**, se em 10 os **DECIS** e se em 100 os **PERCENTIS**.



Exemplo

Considere o seguinte conjunto:

1 -1 0 4 2 5 3

Como $n = 7$ (ímpar), então $x_{(n+1)/2} = x_4$

Ordenando o conjunto, tem-se:

-1 0 1 2 4 3 5

Então: $m_e = x_4 = 2$



Se o conjunto for:

1 -1 0 4 2 5 3 -2

Tem-se: $n = 8$ (par)

Então $m_e = [x_{n/2} + x_{n/2+1}]/2 = (x_4 + x_5)/2$

Ordenando o conjunto, tem-se:

-2 -1 0 1 2 3 4 5

$$m_e = (x_4 + x_5)/2 = (1 + 2)/2 = 1,50$$



(c) A moda (*mode*)

É o(s) valor(es) do conjunto que mais se repete(m).



Exemplo

Considere o conjunto

0 1 1 2 2 2 3 5

Então: $m_0 = 2$

Pois, o **dois** é o que mais se repete
(**três** vezes).



Considere o conjunto:

0 1 1 2 2 3 5

Então: $m_0 = 1$ e $m_0 = 2$

Conjunto bimodal



Considere o conjunto:

0 1 2 3 4 5 7

Este conjunto é **amodal**, pois todos os valores apresentam a mesma frequência.



Dispersão ou Variabilidade

- (a) A amplitude (h)
- (b) O Desvio Médio (dma)
- (c) A Variância (s^2)
- (d) O Desvio Padrão (s)
- (e) A Variância Relativa (g^2)
- (f) O Coeficiente de Variação (s)



A Amplitude (*range*)

$$h = X_{\text{máx}} - X_{\text{mín}}$$

Considere o conjunto:

-2 -1 0 3 5

$$h = 5 - (-2) = 7$$



O dma (*average deviation*)

Considere o conjunto:

-2 -1 0 3 5

A média é:

$$\bar{x} = \frac{-1 - 2 + 0 + 3 + 5}{5} = \frac{5}{5} = 1$$



Calculando os desvios: $x_i - \bar{x}$

Tem-se: $d_1 = -2 - 1 = -3$

$$d_2 = -1 - 1 = -2$$

$$d_3 = 0 - 1 = -1$$

$$d_4 = 3 - 1 = 2$$

$$d_5 = 5 - 1 = 4$$



Como pode ser visto a soma é igual a zero. Tomando o módulo vem:

$$\begin{aligned} \text{dma} &= \frac{\sum |x_i - \bar{x}|}{n} = \\ &= \frac{|-3| + |-2| + |-1| + |2| + |4|}{5} = \\ &= \frac{12}{5} = 2,40 \end{aligned}$$



A variância (*variance*)

Se ao invés de tomar o módulo, elevarmos ao quadrado, tem-se:

$$\begin{aligned} s^2 &= \frac{\sum (x_i - \bar{x})^2}{n} = \\ &= \frac{(-3)^2 + (-2)^2 + (-1)^2 + 2^2 + 4^2}{5} = \\ &= \frac{9 + 4 + 1 + 4 + 16}{5} = \frac{34}{5} = 6,80 \end{aligned}$$



A variância de um conjunto de dados será:

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} =$$

$$= \frac{\sum (x_i - \bar{x})^2}{n}$$

$$s^2 = \frac{\sum x_i^2}{n} - \bar{x}^2$$



O Desvio Padrão (*standard deviation*)

É a raiz quadrada da variância

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}$$



Se extrairmos a raiz quadrada
teremos do resultado anterior
teremos o desvio padrão:

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} = \sqrt{6,80} = 2,61$$



A Variância Relativa

$$g^2 = s^2 / \bar{X}^2$$

O Coeficiente de Variação

$$g = s / \bar{X}$$



O coeficiente de variação do exemplo anterior, será:

$$g = \frac{s}{\bar{X}} = \frac{2,6077}{1} = 260,77 \%$$

