



Prof. MSc. Marcos Alexandruk

E-mail: alexandruk@uni9.pro.br

<https://github.com/alexandruk/analisededados>

Estatística

Estatística

A palavra estatística tem origem no latim "**status**" e relaciona-se com "**estado**".

No início, a palavra era usada para se referir ao "**cidadão político**".

Posteriormente, passou a ser utilizada em alemão com o sentido de "**conjunto de dados do Estado**", de onde decorre o seu significado desde o século XIX.

BATISTA, Carolina. Estatística. Toda Matéria, 2021. Disponível em: <https://www.todamateria.com.br/estatistica-conceito-fases-metodo/>. Acesso em: 23/02/2021.

Estatística

“Estatística é uma ciência exata que estuda a coleta, a organização, a análise e registro de dados por amostras.

Utilizada desde a Antiguidade, quando se registravam os nascimentos e as mortes das pessoas, é um método de pesquisa fundamental para tomar decisões. Isso porque fundamenta suas conclusões nos estudos realizados.”

BATISTA, Carolina. Estatística. Toda Matéria, 2021. Disponível em: <https://www.todamateria.com.br/estatistica-conceito-fases-metodo/>. Acesso em: 23/02/2021.

Estatística

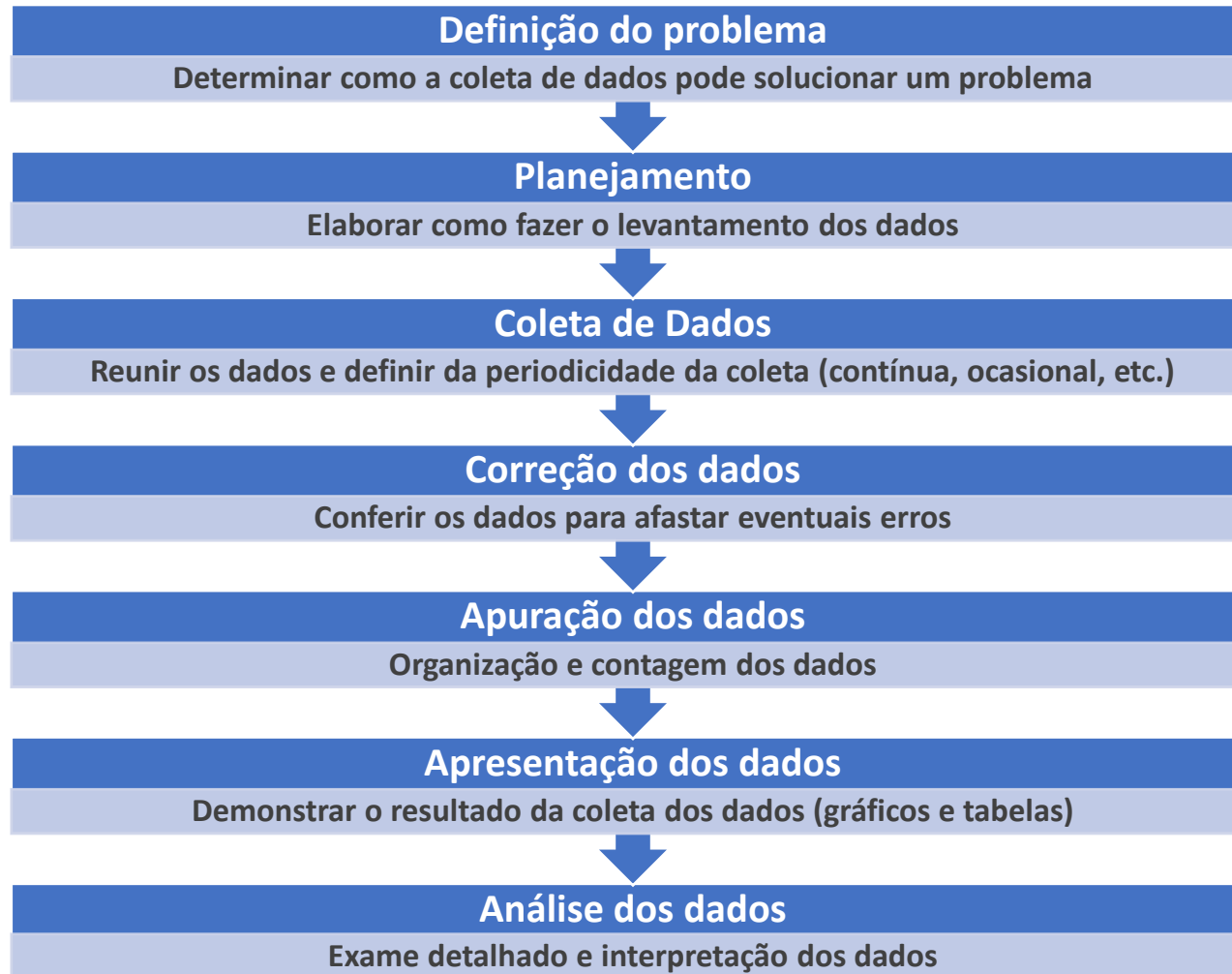
Estatística é ciência que tem por fim a **pesquisa e a comparação dos fatos gerais e particulares** verificados no movimento das sociedades.

Objetivo geral da estatística

O objetivo da estatística é a análise e interpretação dos fenômenos sociais de qualquer natureza, para planejamento de ações.

Análise de Dados

Fases do Método Estatístico



BATISTA, Carolina. Estatística. Toda Matéria, 2021. Disponível em: <https://www.todamateria.com.br/estatistica-conceito-fases-metodo/>. Acesso em: 23/02/2021. (adaptado)

Importância da Estatística na Engenharia

A probabilidade e estatística pode contribuir para mitigação dos erros e favorecer a análise de um projeto em construção, considerando as mais diversas situações, de forma que dados estatísticos podem auxiliar os testes de desempenho e o controle de qualidade.

Análise Descritiva

A análise descritiva dos dados se limita a calcular algumas medidas de posição e variabilidade, como a média e variância, por exemplo.

Inferência

Inferência estatística é um ramo da Estatística cujo objetivo é fazer afirmações a partir de um conjunto de valores representativo (amostra) sobre um universo.

Em geral, a inferência estatística está associada à coleta, à redução, à análise e à modelagem dos dados.

Tipos de Variáveis

Variáveis qualitativas: apresentam algum tipo de atributo do elemento pesquisado. (educação, estado civil, sexo, etc.)

Variáveis quantitativas: apontam para um impacto no elemento pesquisado e contribuem na análise.

Variáveis quantitativas discretas: Quando podemos expressar as variáveis por um número inteiro em certa contagem, chamamos de variável quantitativa discreta. (número de filhos, quantidade de veículos, etc.)

Variáveis quantitativas contínuas: Quando destacamos uma variável por intermédio de uma medida, chamamos de variável quantitativa contínua. (tempo, temperatura, pressão, etc.)

Distribuições de frequências

No estudo de uma variável, devemos dispor um maior interesse em conhecer a distribuição dessa variável por meio das possíveis realizações dela e dispor seus valores, de modo que se tenha uma boa ideia global dessa distribuição.

Distribuições de frequências

Frequência de porcentagens de 20 empregados segundo o grau de instrução:

Grau de instrução	Contagem	Frequência	Proporção	Porcentagem
1º grau	8	8	0,4	40%
2º grau	7	7	0,35	35%
Superior	5	5	0,25	25%
Total	20	20	1,00	100%

Distribuições de frequências

Frequência absoluta acumulada e frequência relativa acumulada:

Grau de instrução	Frequência Absoluta	Frequência Relativa	Frequência Absoluta Acumulada	Frequência Relativa Acumulada
1º grau	8	40%	8	40%
2º grau	7	35%	$8 + 7 = 15$	$40\% + 35\% = 75\%$
Superior	5	25%	$15 + 5 = 20$	$75\% + 25\% = 100\%$
Total	20	100%	20	100%

As frequências acumuladas são extremamente úteis quando o objetivo é saber a quantidade ou a porcentagem até determinada característica.

Amplitude total

Alturas de 32 crianças de 1 a 4 anos:

73,93	71,51	66,83	64,17	66,16	65,7	64,78	65,81
63,15	62,56	61,88	60,94	60,3	60,15	56,57	55,86
71,47	70,09	64,44	63,27	66,06	65,09	64,73	64,16
62,69	61,91	61,49	60,73	60,24	59,37	56,03	55,77

Amplitude total = Valor Máximo – Valor Mínimo

Amplitude total = 73,93 – 55,77

Amplitude total = 18,16

Números de classes

Alturas de 32 crianças de 1 a 4 anos:

73,93	71,51	66,83	64,17	66,16	65,7	64,78	65,81
63,15	62,56	61,88	60,94	60,3	60,15	56,57	55,86
71,47	70,09	64,44	63,27	66,06	65,09	64,73	64,16
62,69	61,91	61,49	60,73	60,24	59,37	56,03	55,77

Número de classes = SQRT (n)

Número de classes = SQRT (32)

Número de classes = 5,65 (aproximado para 6)

SQRT => Raiz Quadrada

Amplitude do intervalo

Amplitude do intervalo = Amplitude total / número de classes

Amplitude do intervalo = 18,16 / 6

Amplitude do intervalo = 3,02

Classes	fi	Fi	fr	Fr
55 ┤	4	4	12,50%	12,50%
58 ┤	6	10	18,75%	31,25%
61 ┤	7	17	21,88%	53,13%
64 ┤	11	28	34,38%	87,50%
67 ┤	1	29	3,13%	90,63%
70 ┤ [ERRO]	3	32	9,38%	100%
Total	32		100%	

Amplitude do intervalo

Amplitude do intervalo = Amplitude total / número de classes

Amplitude do intervalo = 18,16 / 6

Amplitude do intervalo = 3,02

Classes	fi	Fi	fr	Fr
55 ┤ 58	4	4	12,50%	12,50%
58 ┤ 61	6	10	18,75%	31,25%
61 ┤ 64	7	17	21,88%	53,13%
64 ┤ 67	11	28	34,38%	87,50%
67 ┤ 70	1	29	3,13%	90,63%
70 ┤ 73 [ERRO]	3	32	9,38%	100%
Total	32		100%	

[ERRO] O maior valor é 73,93 (está acima de 73)

Amplitude do intervalo

Amplitude do intervalo = Amplitude total / número de classes

Amplitude do intervalo = $18,16 / 6$

Amplitude do intervalo = 3,02 (arredondar para 4)

Classes	fi	Fi	fr	Fr
55 † 59				
59 † 63				
63 † 67				
67 † 71				
71 † 75				
75 † 79				
Total	32		100%	

Regra de Sturges

$$k = 1 + 3,3 * \text{LOG}(n)$$

k = Número de classes

n = Total de dados

$$A_{\text{Total}} = \text{Valor}_{\text{Max}} - \text{Valor}_{\text{Min}}$$

$$h = A_{\text{Total}}/k$$

h = Amplitude do Intervalo

$$k = 1 + 3,3 * \text{LOG}(20)$$

$$k = 5,293399$$

$$k \approx 5$$

$$A_{\text{Total}} = 42 - 15$$

$$A_T = 27$$

$$h = 27/5$$

$$h = 5,4$$

$$h \approx 6$$

(Arredondar para cima)

Pesquisa: Idade				
17	18	16	24	23
42	40	36	15	18
26	23	23	24	28
41	16	18	20	27

IDADE	fi
15 - 21	8
15 - 27	6
27 - 33	2
33 - 39	1
39 - 45	3

Exercício

SALÁRIOS					
20,50	9,50	15,30	17,20	24,10	19,90
15,40	12,70	7,40	16,50	15,30	26,20
14,90	7,80	23,30	15,90	11,80	18,40
13,40	14,30	16,20	16,70	9,20	16,80
9,80	20,10	17,80	17,10	12,60	15,90

Classes	fi

Amplitude Total (A_{Total}) =	
-----------------------------------	--

Total de dados (n) =	
----------------------	--

Número de classes (k) =	
-------------------------	--

Amplitude do intervalo (h) =	
------------------------------	--

Referências

DACHS J. N. W. Análise de dados e regressão. São Paulo: IME USP, 1978.

LEVIN J. Estatística aplicada a Ciências Humanas. São Paulo: Harper e Row do Brasil, 1978.

MORETTIN P. A. Introdução a estatística para ciências exatas. São Paulo: Atual Editora, 1981.