

Variáveis Contínuas e Classes

Matemática – 8º Ano

Dados quantitativos, intervalos de classe, frequências e histogramas

Objetivos de aprendizagem

Ao final deste curso, o estudante deverá ser capaz de:

- 1) compreender o conceito de variável estatística;
- 2) distinguir variáveis qualitativas e quantitativas;
- 3) distinguir variáveis quantitativas discretas e contínuas;
- 4) reconhecer situações em que o agrupamento em classes é conveniente;
- 5) compreender o conceito de intervalo de classe;
- 6) interpretar corretamente notações como:
 $[a, b)$;
- 7) identificar limite inferior e limite superior de uma classe;
- 8) determinar a amplitude total de um conjunto de dados;
- 9) determinar a amplitude de uma classe;
- 10) determinar o ponto médio de uma classe;
- 11) organizar dados contínuos em tabelas de distribuição de frequências;
- 12) calcular frequência absoluta;
- 13) calcular frequência relativa;
- 14) calcular frequência percentual;
- 15) calcular frequência acumulada;
- 16) verificar a soma das frequências;
- 17) construir histogramas com classes de mesma amplitude;
- 18) distinguir rigorosamente histogramas de gráficos de barras;
- 19) interpretar distribuições a partir de histogramas;
- 20) identificar classes modais;
- 21) reconhecer limitações decorrentes do agrupamento dos dados;
- 22) compreender, como aprofundamento, o papel da densidade de frequência quando as classes possuem larguras diferentes.

Sumário

1	Introdução	1
2	Variável estatística	1
3	Classificação das variáveis	2
4	Variáveis qualitativas	2
5	Variáveis quantitativas	3
6	Variável quantitativa discreta	3
7	Variável quantitativa contínua	4
8	Discreto versus contínuo	4
9	Tabela comparativa	5
10	Precisão de medida	5
11	Por que agrupar dados em classes?	5
12	Classe estatística	6
13	Notação de intervalos	6
14	Exemplo	7
15	Por que usar intervalos semiabertos?	7
16	Outra convenção possível	8
17	Limite inferior	8
18	Limite superior	8
19	Amplitude de uma classe	9
20	Exemplo de amplitude	9
21	Classes de mesma amplitude	10

22 Amplitude total	10
23 Exemplo de amplitude total	10
24 Amplitude total e amplitude de classe	11
25 Número de classes	11
26 Exemplo de construção de classes	11
27 Ponto médio de uma classe	12
28 Exemplo de ponto médio	13
29 Interpretação do ponto médio	13
30 Tabela de distribuição de frequências	13
31 Frequência absoluta	14
32 Total de observações	14
33 Frequência relativa	15
34 Exemplo de frequência relativa	15
35 Frequência percentual	15
36 Tabela completa	16
37 Verificações fundamentais	16
38 Frequência acumulada	16
39 Exemplo de frequência acumulada	17
40 Tabela com frequência acumulada	17
41 Interpretação da frequência acumulada	17
42 Frequência relativa acumulada	18
43 Classe modal	18

44 A classe modal não fornece o valor exato da moda	19
45 Histograma	19
46 Por que as barras se encostam?	19
47 Histograma do exemplo	20
48 Leitura do histograma	20
49 Histograma versus gráfico de barras	21
50 Exemplo de erro	21
51 Construção de uma distribuição a partir de dados brutos	21
52 Determinação do intervalo dos dados	21
53 Escolha das classes	22
54 Contagem dos dados	23
55 Tabela resultante	24
56 Verificação	24
57 Perda de informação ao agrupar	25
58 Vantagem e custo do agrupamento	25
59 Classes muito largas	25
60 Classes muito estreitas	26
61 Comparação de distribuições	26
62 Exemplo de concentração	26
63 Distribuição simétrica: ideia intuitiva	27
64 Distribuição assimétrica	27
65 Classes com amplitudes diferentes	27

66 Problema das amplitudes diferentes	27
67 Densidade de frequência – aprofundamento	28
68 Por que utilizar densidade?	28
69 Exemplo de densidade	29
70 Observação de nível	30
71 Dados agrupados e valores exatos	30
72 Construção passo a passo de uma tabela	30
73 Construção passo a passo de um histograma	31
74 Leitura de porcentagens em classes	31
75 Problema inverso com porcentagem	32
76 Problema inverso com frequência acumulada	32
77 Erros comuns	33
78 Mapa conceitual	35
79 Resumo teórico	35
80 Exercícios	36
80.1 Parte A – Classificação de variáveis	36
80.2 Parte B – Intervalos	37
80.3 Parte C – Limites e amplitude	37
80.4 Parte D – Amplitude total	38
80.5 Parte E – Frequência	38
80.6 Parte F – Frequência acumulada	39
80.7 Parte G – Dados brutos	39
80.8 Parte H – Frequências relativas	40
80.9 Parte I – Construção de histograma	40
80.10 Parte J – Problema inverso	40
80.11 Parte K – Frequência acumulada inversa	41

80.12	Parte L – Histograma versus barras	41
80.13	Parte M – Classes com amplitudes diferentes	41
80.14	Parte N – Desafio	42
81	Gabarito comentado	42
81.1	Parte A	42
81.2	Parte B	43
81.3	Parte C	44
81.4	Parte D	44
81.5	Parte E	45
81.6	Parte F	46
81.7	Parte G	46
81.8	Parte H	47
81.9	Parte I	48
81.10	Parte J	49
81.11	Parte K	49
81.12	Parte L	49
81.13	Parte M	50
81.14	Parte N	50
82	Avaliação final	51
82.1	Gabarito da avaliação final	52
83	Conclusão	55

1 Introdução

Em pesquisas estatísticas, muitas variáveis podem assumir grande quantidade de valores diferentes.

Considere, por exemplo, as alturas de 1000 pessoas.

Poderíamos obter valores como:

1,621 m, 1,684 m, 1,703 m, 1,756 m,

e centenas de outros.

Representar cada valor individualmente pode tornar a análise difícil.

Uma estratégia importante é agrupar os valores em:

classes

ou:

intervalos.

Por exemplo:

$[1,50, 1,60)$,

$[1,60, 1,70)$,

$[1,70, 1,80)$,

$[1,80, 1,90)$.

Esse procedimento permite resumir grandes conjuntos de dados e construir representações como:

histogramas.

2 Variável estatística

Definição 2.1: Variável estatística

Uma **variável estatística** é uma característica observada ou medida nos elementos de uma população ou amostra e que pode assumir diferentes valores.

Exemplos:

- idade;
- altura;
- massa;
- número de irmãos;
- cor preferida;
- meio de transporte utilizado;
- temperatura.

3 Classificação das variáveis

Uma primeira classificação distingue:

variáveis qualitativas

e:

variáveis quantitativas.

4 Variáveis qualitativas

Definição 4.1: Variável qualitativa

Uma variável é **qualitativa** quando seus valores representam categorias, qualidades ou atributos, e não quantidades numéricas resultantes de contagem ou medição.

Exemplos:

cor dos olhos,

tipo sanguíneo,

cidade de nascimento,

esporte preferido.

5 Variáveis quantitativas

Definição 5.1: Variável quantitativa

Uma variável é **quantitativa** quando seus valores representam quantidades numéricas obtidas por contagem ou medição.

Exemplos:

idade,

altura,

massa,

número de livros,

temperatura.

6 Variável quantitativa discreta

Definição 6.1: Variável quantitativa discreta

Uma variável quantitativa é **discreta** quando seus possíveis valores são, em geral, separados e resultam tipicamente de uma contagem.

Exemplos:

número de irmãos,

número de carros,

número de gols.

Uma pessoa pode ter:

0, 1, 2, 3, ...

irmãos.

Não faz sentido, nesse contexto, dizer que ela possui:

2,47

irmãos.

7 Variável quantitativa contínua

Definição 7.1: Variável quantitativa contínua

Uma variável quantitativa é **contínua** quando, dentro de determinado intervalo, pode assumir uma grande variedade de valores reais, sendo normalmente obtida por medição.

Exemplos:

altura,

massa,

tempo,

temperatura,

comprimento.

8 Discreto versus contínuo

Considere:

número de filhos

e:

altura.

O número de filhos pode assumir:

$0, 1, 2, 3, \dots$

A altura pode assumir, dentro de certo intervalo, valores como:

1,70,

1,701,

1,7015,

e assim por diante, dependendo da precisão do instrumento.

9 Tabela comparativa

Variável	Tipo	Origem típica
Número de irmãos	quantitativa discreta	contagem
Número de livros	quantitativa discreta	contagem
Altura	quantitativa contínua	medição
Massa	quantitativa contínua	medição
Tempo de corrida	quantitativa contínua	medição

10 Precisão de medida

Uma variável contínua é registrada com precisão limitada pelo instrumento utilizado. Por exemplo, uma altura real pode ser registrada como:

1,72 m

mesmo que um instrumento mais preciso fornecesse:

1,7184 m.

Observação

O fato de os valores registrados terem poucas casas decimais não transforma automaticamente uma variável contínua em discreta.
O tipo da variável depende de sua natureza e do processo de obtenção.

11 Por que agrupar dados em classes?

Considere 100 medidas de altura.

Se cada altura for apresentada individualmente, a tabela poderá ficar extensa.

Podemos resumir os valores em intervalos:

$[150, 160)$,

$$[160, 170),$$

$$[170, 180),$$

$$[180, 190).$$

Cada intervalo será denominado:

classe.

12 Classe estatística

Definição 12.1: Classe

Uma **classe** é um intervalo de valores utilizado para agrupar observações de uma variável quantitativa.

Por exemplo:

$$[160, 170)$$

é uma classe que contém valores x tais que:

$$160 \leq x < 170.$$

13 Notação de intervalos

A notação:

$$[a, b)$$

significa:

$a \leq x < b.$

O colchete:

[

indica que o limite está incluído.

O parêntese:

)

indica que o limite não está incluído.

14 Exemplo

Considere:

$$[10, 20).$$

Temos:

$$10 \in [10, 20),$$

$$15 \in [10, 20),$$

$$19,999 \in [10, 20),$$

mas:

$$20 \notin [10, 20).$$

15 Por que usar intervalos semiabertos?

Considere duas classes consecutivas:

$$[10, 20)$$

e:

$$[20, 30).$$

O valor:

$$20$$

pertence apenas à segunda classe.

Assim evitamos que o mesmo valor seja contado duas vezes.

Atenção

Se utilizássemos:

$$[10, 20]$$

e:

$$[20, 30],$$

o número 20 pertenceria às duas classes.

Isso geraria ambiguidade.

16 Outra convenção possível

Também poderíamos utilizar:

$$(10, 20], \quad (20, 30],$$

desde que a convenção seja usada de forma consistente.

O essencial é:

cada observação deve pertencer a uma única classe.

17 Limite inferior

Definição 17.1: Limite inferior da classe

O **limite inferior** de uma classe é o menor valor que define o intervalo.

Na classe:

$$[160, 170),$$

o limite inferior é:

$$160.$$

18 Limite superior

Definição 18.1: Limite superior da classe

O **limite superior** de uma classe é o maior valor que delimita o intervalo.

Na classe:

$$[160, 170),$$

o limite superior é:

$$170.$$

Observe que, nessa convenção:

$$170$$

não pertence à classe, apesar de ser seu limite superior.

19 Amplitude de uma classe

Definição 19.1: Amplitude de classe

A amplitude de uma classe é a diferença entre seu limite superior e seu limite inferior.

Se:

$$L_i$$

é o limite inferior e:

$$L_s$$

o limite superior:

$$h = L_s - L_i.$$

20 Exemplo de amplitude

Para:

$$[160, 170),$$

temos:

$$h = 170 - 160.$$

Logo:

$$h = 10.$$

21 Classes de mesma amplitude

Considere:

$$[150, 160),$$

$$[160, 170),$$

$$[170, 180),$$

$$[180, 190).$$

Todas possuem amplitude:

$$10.$$

Esse tipo de agrupamento facilita a leitura e a construção de histogramas.

22 Amplitude total

Definição 22.1: Amplitude total

A **amplitude total** de um conjunto de dados é a diferença entre o maior e o menor valor observado.

Se:

$$x_{\max}$$

é o maior valor e:

$$x_{\min}$$

o menor:

$$A = x_{\max} - x_{\min}.$$

23 Exemplo de amplitude total

Considere:

$$12, \quad 15, \quad 17, \quad 21, \quad 26, \quad 29.$$

Temos:

$$x_{\min} = 12,$$

$$x_{\max} = 29.$$

Logo:

$$A = 29 - 12.$$

Portanto:

$$A = 17.$$

24 Amplitude total e amplitude de classe

Essas duas grandezas não devem ser confundidas.

Grandeza	Definição
Amplitude total	$x_{\max} - x_{\min}$
Amplitude de classe	$L_s - L_i$

25 Número de classes

Não existe uma única regra obrigatória para escolher o número de classes no nível escolar.

Devemos buscar um equilíbrio:

nem classes demais, nem classes de menos.

Classes demais podem reproduzir praticamente os dados brutos.

Classes muito largas podem esconder características importantes.

26 Exemplo de construção de classes

Suponha que as notas de um conjunto de estudantes variem de:

32

a:

94.

Podemos escolher, por exemplo, classes:

$[30, 40)$,

$[40, 50)$,

$[50, 60)$,

$[60, 70)$,

$[70, 80)$,

$[80, 90)$,

$[90, 100)$.

Cada classe possui amplitude:

10.

27 Ponto médio de uma classe

Definição 27.1: Ponto médio da classe

O ponto médio de uma classe é a média aritmética de seus limites inferior e superior.

Se os limites são:

L_i

e:

L_s ,

então:

$$m = \frac{L_i + L_s}{2}.$$

28 Exemplo de ponto médio

Para a classe:

$$[160, 170),$$

temos:

$$m = \frac{160 + 170}{2}.$$

Logo:

$$m = 165.$$

29 Interpretação do ponto médio

O ponto médio:

$$165$$

representa o centro geométrico do intervalo:

$$[160, 170).$$

Posteriormente, ele poderá ser utilizado como valor representativo da classe em determinados cálculos aproximados.

Atenção

O ponto médio não significa que todos os dados da classe sejam iguais a ele. Ele é apenas um valor representativo do intervalo.

30 Tabela de distribuição de frequências

Definição 30.1: Distribuição de frequências

Uma distribuição de frequências organiza os dados em valores ou classes e registra quantas observações pertencem a cada grupo.

Para dados contínuos, podemos ter:

Altura (cm)	Ponto médio	Frequência
[150, 160)	155	4
[160, 170)	165	8
[170, 180)	175	12
[180, 190)	185	6

31 Frequência absoluta

Definição 31.1: Frequência absoluta

A frequência absoluta de uma classe é o número de observações pertencentes a essa classe.

Indicaremos por:

$$f_i.$$

No exemplo:

$$f_1 = 4,$$

$$f_2 = 8,$$

$$f_3 = 12,$$

$$f_4 = 6.$$

32 Total de observações

O total é:

$$n = 4 + 8 + 12 + 6.$$

Logo:

$$n = 30.$$

De modo geral:

$$n = \sum_i f_i.$$

33 Frequência relativa

Definição 33.1: Frequência relativa

A frequência relativa de uma classe é a razão entre sua frequência absoluta e o total de observações.

Assim:

$$f_{r,i} = \frac{f_i}{n}.$$

34 Exemplo de frequência relativa

Para a classe:

$$[170, 180),$$

temos:

$$f_i = 12,$$

$$n = 30.$$

Logo:

$$f_r = \frac{12}{30}.$$

Simplificando:

$$f_r = \frac{2}{5}.$$

Em forma decimal:

$$f_r = 0,4.$$

35 Frequência percentual

A frequência percentual é:

$$f_{\%,i} = 100 \frac{f_i}{n}.$$

No exemplo:

$$0,4 \cdot 100\% = \boxed{40\%}.$$

36 Tabela completa

Para os dados anteriores:

Classe	m_i	f_i	$f_{r,i}$	$f_{\%,i}$
[150, 160)	155	4	4/30	13,33%
[160, 170)	165	8	8/30	26,67%
[170, 180)	175	12	12/30	40%
[180, 190)	185	6	6/30	20%
Total	—	30	1	100%

37 Verificações fundamentais

Uma tabela correta deve satisfazer:

$$\sum_i f_i = n.$$

Além disso:

$$\sum_i f_{r,i} = 1.$$

Em porcentagem:

$$\sum_i f_{\%,i} = 100\%.$$

38 Frequência acumulada

Definição 38.1: Frequência acumulada

A frequência acumulada até determinada classe é a soma das frequências dessa classe e de todas as classes anteriores.

Indicaremos por:

$$F_i.$$

39 Exemplo de frequência acumulada

Considere:

$$4, \quad 8, \quad 12, \quad 6.$$

Temos:

$$F_1 = 4,$$

$$F_2 = 4 + 8 = 12,$$

$$F_3 = 4 + 8 + 12 = 24,$$

$$F_4 = 4 + 8 + 12 + 6 = 30.$$

Logo:

$$F_i = 4, \quad 12, \quad 24, \quad 30.$$

40 Tabela com frequência acumulada

Classe	f_i	F_i
$[150, 160)$	4	4
$[160, 170)$	8	12
$[170, 180)$	12	24
$[180, 190)$	6	30

41 Interpretação da frequência acumulada

No exemplo:

$$F_2 = 12.$$

Isso significa que:

$$12$$

observações possuem valor inferior a:

170,

considerando as classes iniciadas em 150.

Analogamente:

$$F_3 = 24$$

indica que 24 observações estão abaixo de 180.

42 Frequência relativa acumulada

Também podemos acumular frequências relativas.

$$F_{r,i} = \frac{F_i}{n}.$$

Por exemplo:

$$F_3 = 24, \quad n = 30.$$

Logo:

$$F_{r,3} = \frac{24}{30} = 0,8.$$

Assim:

$$80\%$$

das observações estão abaixo de 180.

43 Classe modal

Definição 43.1: Classe modal

A **classe modal** é a classe que possui a maior frequência.

No exemplo:

$$4, \quad 8, \quad 12, \quad 6,$$

a maior frequência é:

$$12.$$

Logo, a classe modal é:

$$[170, 180).$$

44 A classe modal não fornece o valor exato da moda

Atenção

Se os dados foram agrupados, sabemos apenas que a maior concentração está na classe:

$$[170, 180).$$

Não podemos concluir, apenas pela tabela agrupada, que a moda seja:

$$175.$$

O ponto médio é um representante da classe, não necessariamente o valor mais frequente observado.

45 Histograma

Definição 45.1: Histograma

Um **histograma** é um gráfico utilizado para representar distribuições de frequências de variáveis quantitativas agrupadas em classes.

Quando as classes possuem a mesma amplitude, podemos representar a frequência pela altura das barras.

As barras são:

$$\text{adjacentes.}$$

46 Por que as barras se encostam?

Considere:

$$[150, 160),$$

$$[160, 170).$$

Esses intervalos são consecutivos.

Não existe uma categoria intermediária.

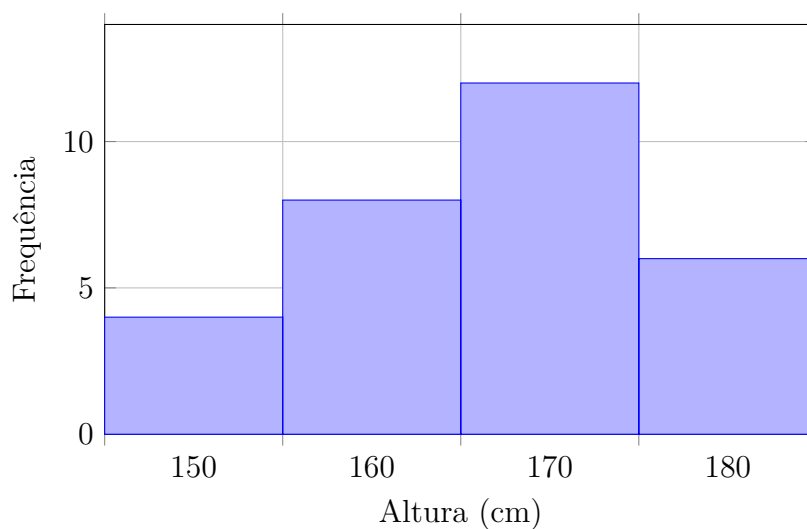
Por isso, as barras do histograma ficam encostadas, expressando a continuidade da escala.

47 Histograma do exemplo

Para:

Classe	Frequência
$[150, 160)$	4
$[160, 170)$	8
$[170, 180)$	12
$[180, 190)$	6

temos:



48 Leitura do histograma

A maior barra corresponde à classe:

$[170, 180)$.

Logo, essa é a classe com maior frequência.

A menor barra corresponde a:

$[150, 160)$,

com:

4

observações.

49 Histograma versus gráfico de barras

	Gráfico de barras	Histograma
Dados	categorias	quantitativos em classes
Barras	separadas	adjacentes
Ordem	muitas vezes flexível	determinada pela escala
Eixo horizontal	categorias	intervalos numéricos

50 Exemplo de erro

Considere as categorias:

vermelho, azul, verde.

Essas categorias não formam uma escala numérica contínua.

Logo, representar seus dados por um histograma é inadequado.

O correto seria, por exemplo:

gráfico de barras.

51 Construção de uma distribuição a partir de dados brutos

Considere os tempos, em minutos:

12, 17, 21, 15, 28, 31, 19, 23, 26, 34,

18, 22, 25, 29, 16, 20, 27, 33, 24, 14.

Queremos agrupá-los em classes de amplitude:

5.

52 Determinação do intervalo dos dados

O menor valor é:

12.

O maior é:

34.

Logo:

$$A = 34 - 12.$$

Assim:

$$A = 22.$$

53 Escolha das classes

Podemos iniciar em:

10

e utilizar:

$[10, 15),$

$[15, 20),$

$[20, 25),$

$[25, 30),$

$[30, 35).$

Todas possuem amplitude:

5.

54 Contagem dos dados

Na classe:

$$[10, 15),$$

temos:

$$12, \quad 14.$$

Logo:

$$f_1 = 2.$$

Na classe:

$$[15, 20),$$

temos:

$$15, \quad 16, \quad 17, \quad 18, \quad 19.$$

Logo:

$$f_2 = 5.$$

Na classe:

$$[20, 25),$$

temos:

$$20, \quad 21, \quad 22, \quad 23, \quad 24.$$

Logo:

$$f_3 = 5.$$

Na classe:

$$[25, 30),$$

temos:

$$25, \quad 26, \quad 27, \quad 28, \quad 29.$$

Logo:

$$f_4 = 5.$$

Na classe:

$$[30, 35),$$

temos:

$$31, \quad 33, \quad 34.$$

Logo:

$$f_5 = 3.$$

55 Tabela resultante

Classe	Ponto médio	f_i	$f_{r,i}$	F_i
$[10, 15)$	12,5	2	0,10	2
$[15, 20)$	17,5	5	0,25	7
$[20, 25)$	22,5	5	0,25	12
$[25, 30)$	27,5	5	0,25	17
$[30, 35)$	32,5	3	0,15	20
Total	—	20	1	—

56 Verificação

Somando:

$$2 + 5 + 5 + 5 + 3 = 20.$$

E:

$$0,10 + 0,25 + 0,25 + 0,25 + 0,15 = 1.$$

Logo, a distribuição está consistente.

57 Perda de informação ao agrupar

Ao transformar dados individuais em classes, perdemos parte da informação original.

Por exemplo, ao observar apenas:

$[20, 25)$

com frequência:

5,

não sabemos se os valores originais eram:

20, 21, 22, 23, 24

ou:

20,1, 20,3, 23,7, 24,2, 24,9.

58 Vantagem e custo do agrupamento

Observação

O agrupamento em classes produz uma troca:

mais síntese

em troca de:

menos detalhe.

Essa é uma ideia importante em Estatística.

59 Classes muito largas

Considere uma distribuição com valores entre:

0

e:

100.

Se utilizarmos apenas:

25

$[0, 50)$

e:

$[50, 100]$,

podemos esconder importantes diferenças internas.

60 Classes muito estreitas

Por outro lado, utilizar dezenas de classes muito estreitas pode produzir uma representação irregular e pouco resumida.

Assim, a escolha das classes deve ser:

adequada ao objetivo da análise.

61 Comparação de distribuições

Histogramas permitem comparar:

- concentração;
- dispersão;
- regiões de maior frequência;
- presença de valores em intervalos extremos.

62 Exemplo de concentração

Considere duas turmas.

Na Turma A, a maior parte das notas está entre:

$[60, 80)$.

Na Turma B, as notas estão espalhadas entre:

$[30, 100)$.

Mesmo sem calcular medidas numéricas, podemos dizer que:

a distribuição da Turma A parece mais concentrada.

63 Distribuição simétrica: ideia intuitiva

Uma distribuição é aproximadamente simétrica quando os lados do histograma apresentam comportamento semelhante em torno de uma região central.

Por exemplo:

2, 5, 8, 5, 2.

Essa sequência de frequências é simétrica.

64 Distribuição assimétrica

Considere:

10, 8, 5, 3, 1.

A distribuição apresenta maior concentração nas primeiras classes e uma extensão para as classes de valores maiores.

Esse tipo de comportamento é denominado assimetria.

Observação

A análise formal de assimetria pertence a níveis posteriores, mas a identificação visual de distribuições não simétricas já é útil.

65 Classes com amplitudes diferentes

Até agora utilizamos classes de mesma amplitude.

Por exemplo:

$[0, 10)$, $[10, 20)$, $[20, 30)$.

Nesse caso, comparar diretamente as alturas das barras é natural.

66 Problema das amplitudes diferentes

Considere:

$[0, 10)$

com frequência:

10

e:

$[10, 30)$

com frequência:

20.

A segunda classe possui o dobro da largura.

Se desenharmos a altura como frequência, a área de sua barra será:

2

vezes maior pela altura e:

2

vezes maior pela largura.

Isso produziria uma área:

4

vezes maior, apesar de a frequência ser apenas o dobro.

67 Densidade de frequência – aprofundamento

Definição 67.1: Densidade de frequência

Quando as classes possuem amplitudes diferentes, podemos definir a densidade de frequência por:

$$d_i = \frac{f_i}{h_i},$$

onde f_i é a frequência e h_i é a amplitude da classe.

Nesse caso, a altura da barra é a densidade.

68 Por que utilizar densidade?

A área da barra é:

largura $\times$ altura.

Logo:

$$h_i \cdot d_i = h_i \cdot \frac{f_i}{h_i}.$$

Assim:

$$\boxed{\text{área da barra} = f_i.}$$

Portanto, a área de cada barra permanece proporcional à frequência.

69 Exemplo de densidade

Classe 1:

$$[0, 10), \quad f_1 = 10.$$

Amplitude:

$$h_1 = 10.$$

Então:

$$d_1 = \frac{10}{10} = 1.$$

Classe 2:

$$[10, 30), \quad f_2 = 20.$$

Amplitude:

$$h_2 = 20.$$

Então:

$$d_2 = \frac{20}{20} = 1.$$

As barras possuem a mesma altura, mas a segunda possui o dobro da largura.

Logo, sua área é o dobro, exatamente como sua frequência.

70 Observação de nível

Observação

No 8º ano, histogramas são normalmente apresentados com classes de mesma amplitude. A densidade de frequência é apresentada aqui como aprofundamento para evitar uma interpretação matematicamente incorreta de histogramas com classes de larguras diferentes.

71 Dados agrupados e valores exatos

Considere a classe:

$$[50, 60)$$

com frequência:

$$12.$$

A tabela informa que existem:

$$12$$

valores no intervalo.

Ela não informa seus valores exatos.

Por isso, medidas como média e mediana calculadas apenas a partir de classes podem exigir aproximações.

Esse tema será aprofundado posteriormente.

72 Construção passo a passo de uma tabela

Notação

Para agrupar dados contínuos em classes:

1) identifique:

$$x_{\min};$$

2) identifique:

$$x_{\max};$$

3) calcule a amplitude total:

$$A = x_{\max} - x_{\min};$$

4) escolha um número razoável de classes;

- 5) escolha uma amplitude conveniente;
- 6) determine os limites das classes;
- 7) defina claramente a convenção dos intervalos;
- 8) classifique cada observação;
- 9) conte as frequências;
- 10) verifique:

$$\sum f_i = n;$$

- 11) calcule frequências relativas e percentuais;
- 12) construa o histograma.

73 Construção passo a passo de um histograma

Para construir um histograma:

- 1) coloque os limites de classe no eixo horizontal;
- 2) coloque frequências no eixo vertical, quando as classes possuem a mesma amplitude;
- 3) desenhe uma barra para cada classe;
- 4) mantenha as barras adjacentes;
- 5) utilize largura correspondente à amplitude da classe;
- 6) coloque título;
- 7) indique unidades e escalas;
- 8) verifique se cada barra corresponde corretamente à tabela.

74 Leitura de porcentagens em classes

Se uma classe possui frequência:

18

em um total de:

60,

então:

$$f_r = \frac{18}{60}.$$

Logo:

$$f_r = 0,3.$$

Em porcentagem:

$$\boxed{30\%}.$$

75 Problema inverso com porcentagem

Exemplo

Uma classe representa:

$$25\%$$

de uma amostra de:

$$120$$

pessoas.

Determine a frequência absoluta.

Resolução

$$f = 0,25 \cdot 120.$$

Logo:

$$\boxed{f = 30}.$$

76 Problema inverso com frequência acumulada

Exemplo

Uma tabela apresenta:

$$F_2 = 18$$

e:

$$F_3 = 31.$$

Determine a frequência da terceira classe.

Resolução

A terceira frequência é o acréscimo ocorrido:

$$f_3 = F_3 - F_2.$$

Logo:

$$f_3 = 31 - 18.$$

Portanto:

$$f_3 = 13.$$

77 Erros comuns

Atenção

Erro 1: confundir variável contínua com discreta.

Número de irmãos é discreto.

Altura é contínua.

Atenção

Erro 2: contar o mesmo limite em duas classes.

Evite classes como:

$$[10, 20]$$

e:

$$[20, 30].$$

Prefira, por exemplo:

$$[10, 20)$$

e:

$$[20, 30).$$

Atenção

Erro 3: confundir amplitude total e amplitude de classe.

$$A = x_{\max} - x_{\min}$$

é a amplitude total.

$$h = L_s - L_i$$

é a amplitude da classe.

Atenção

Erro 4: esquecer dados no momento da contagem.

Sempre verifique:

$$\sum f_i = n.$$

Atenção

Erro 5: obter frequências relativas cuja soma não é 1.

Deve ocorrer:

$$\sum f_{r,i} = 1.$$

Pequenas diferenças podem aparecer apenas por arredondamentos decimais.

Atenção

Erro 6: confundir histograma com gráfico de barras.

No histograma:

classes numéricas consecutivas

e:

barras adjacentes .

Atenção

Erro 7: concluir valores individuais a partir de uma classe.

Saber que uma pessoa está em:

$$[170, 180)$$

não permite determinar sua altura exata.

Atenção

Erro 8: afirmar que o ponto médio é necessariamente um dado observado.

O ponto médio:

$$m = \frac{L_i + L_s}{2}$$

pode não aparecer entre os dados originais.

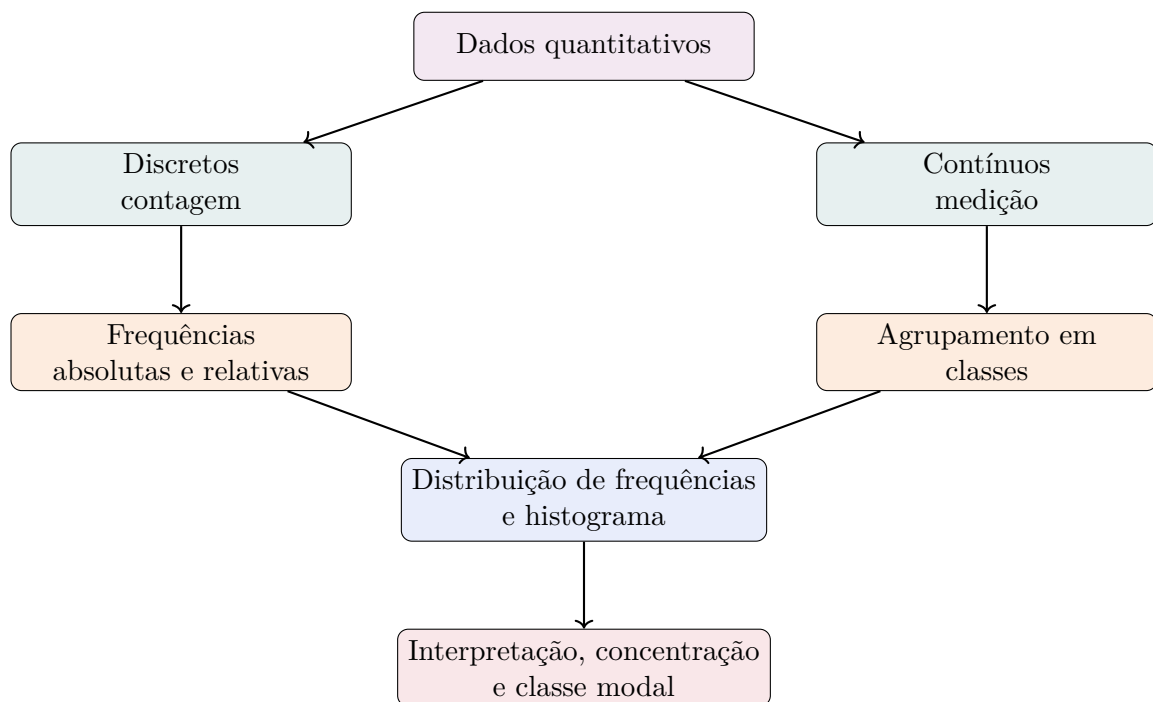
Atenção

Erro 9: comparar apenas alturas de barras em histogramas com classes de amplitudes diferentes.

Nesse caso, é necessário considerar a:

densidade de frequência.

78 Mapa conceitual



79 Resumo teórico

Conceito	Relação
Amplitude total	$A = x_{\max} - x_{\min}$
Amplitude de classe	$h = L_s - L_i$
Ponto médio	$m = \frac{L_i + L_s}{2}$
Frequência relativa	$f_r = \frac{f}{n}$
Frequência percentual	$f_{\%} = 100 \frac{f}{n}$
Frequência acumulada	$F_i = f_1 + \dots + f_i$
Densidade, se necessária	$d_i = \frac{f_i}{h_i}$

Notação

Relações fundamentais:

$$A = x_{\max} - x_{\min}$$

$$h = L_s - L_i$$

$$m = \frac{L_i + L_s}{2}$$

$$f_r = \frac{f}{n}$$

$$f_{\%} = 100 \frac{f}{n}$$

$$F_i = \sum_{j=1}^i f_j.$$

Verificações:

$$\sum_i f_i = n$$

$$\sum_i f_{r,i} = 1$$

$$\sum_i f_{\%,i} = 100\%.$$

80 Exercícios

80.1 Parte A – Classificação de variáveis

Exercício

Classifique cada variável como:

qualitativa,

quantitativa discreta

ou:

quantitativa contínua.

a) altura de uma pessoa;

- b) número de irmãos;
- c) cor preferida;
- d) massa corporal;
- e) número de gols em uma partida;
- f) temperatura do ar;
- g) tempo para concluir uma corrida;
- h) tipo sanguíneo.

80.2 Parte B – Intervalos

Exercício

Considere a classe:

$$[20, 30).$$

Determine se cada valor pertence ou não à classe:

- a) 20;
- b) 25;
- c) 29,99;
- d) 30;
- e) 19,9.

80.3 Parte C – Limites e amplitude

Exercício

Para cada classe, determine:

- limite inferior;
- limite superior;
- amplitude;
- ponto médio.

a) $[10, 20);$

b) $[30, 45);$

c)

$[100, 125);$

d)

$[1,5, 2,0).$

80.4 Parte D – Amplitude total

Exercício

Determine a amplitude total.

a)

5, 8, 12, 17, 22;

b)

31, 42, 38, 55, 47, 29;

c) dados cujo menor valor é 12,5 e maior valor é 31,8.

80.5 Parte E – Frequência

Exercício

Considere:

Classe	Frequência
$[0, 10)$	4
$[10, 20)$	9
$[20, 30)$	12
$[30, 40)$	5

Determine:

a) o total de observações;

b) a frequência relativa de cada classe;

c) a porcentagem de cada classe;

d) a classe modal.

80.6 Parte F – Frequência acumulada

Exercício

Para a tabela da questão anterior, determine:

- a) F_1 ;
- b) F_2 ;
- c) F_3 ;
- d) F_4 ;
- e) a porcentagem acumulada até o limite superior 30.

80.7 Parte G – Dados brutos

Exercício

Considere:

11, 12, 14, 15, 16, 18, 19, 20, 21, 22,

23, 24, 25, 27, 28, 30, 31, 33, 34, 37.

Agrupe os dados nas classes:

$[10, 15)$,

$[15, 20)$,

$[20, 25)$,

$[25, 30)$,

$[30, 35)$,

$[35, 40)$.

Determine a frequência de cada classe.

80.8 Parte H – Frequências relativas

Exercício

Utilizando a distribuição da Parte G:

- a) calcule todas as frequências relativas;
- b) calcule as porcentagens;
- c) determine a classe modal;
- d) verifique a soma das frequências relativas.

80.9 Parte I – Construção de histograma

Exercício

Utilize a seguinte tabela para construir um histograma:

Classe	Frequência
[40, 50)	2
[50, 60)	6
[60, 70)	10
[70, 80)	8
[80, 90)	4

Depois responda:

- a) qual é a classe modal?
- b) quantas observações existem?
- c) que porcentagem está no intervalo:
[60, 80)?

80.10 Parte J – Problema inverso

Exercício

Uma classe representa:

35%

de uma amostra de:

80

pessoas.

Determine a frequência absoluta da classe.

80.11 Parte K – Frequência acumulada inversa

Exercício

Em uma distribuição:

$$F_2 = 17,$$

$$F_3 = 29,$$

$$F_4 = 41.$$

Determine:

a) f_3 ;

b) f_4 .

80.12 Parte L – Histograma versus barras

Exercício

Explique por que:

“tipo de música preferido”

não deve ser representado por um histograma.

Indique um gráfico mais adequado.

80.13 Parte M – Classes com amplitudes diferentes

Exercício

Considere:

$$[0, 10)$$

com frequência:

$$8$$

e:

$$[10, 30)$$

com frequência:

12.

Determine:

- a) a amplitude de cada classe;
- b) a densidade de frequência de cada classe;
- c) qual deve possuir maior altura em um histograma construído por densidades.

80.14 Parte N – Desafio

Exercício

Uma distribuição possui as classes:

$[0, 5), [5, 10), [10, 15), [15, 20)$

com frequências:

$x, 8, 12, 5.$

Sabendo que há:

30

observações, determine:

- a) x ;
- b) a frequência relativa da primeira classe;
- c) a frequência acumulada da terceira classe;
- d) a porcentagem de observações abaixo de 15.

81 Gabarito comentado

81.1 Parte A

Resolução

a) altura:

quantitativa contínua;

b) número de irmãos:

quantitativa discreta;

c) cor preferida:

qualitativa;

d) massa:

quantitativa contínua ;

e) número de gols:

quantitativa discreta ;

f) temperatura:

quantitativa contínua ;

g) tempo:

quantitativa contínua ;

h) tipo sanguíneo:

qualitativa .

81.2 Parte B

Resolução

Para:

$[20, 30)$,

temos:

$20 \leq x < 30$.

Assim:

$20 \in [20, 30)$,

$25 \in [20, 30)$,

$29,99 \in [20, 30)$,

mas:

$30 \notin [20, 30)$,

e:

$19,9 \notin [20, 30)$.

81.3 Parte C

Resolução

a)

$$L_i = 10, \quad L_s = 20.$$

$$h = 20 - 10 = 10.$$

$$m = \frac{10 + 20}{2} = 15.$$

b)

$$L_i = 30, \quad L_s = 45.$$

$$h = 15.$$

$$m = \frac{75}{2} = 37,5.$$

c)

$$L_i = 100, \quad L_s = 125.$$

$$h = 25.$$

$$m = \frac{225}{2} = 112,5.$$

d)

$$L_i = 1,5, \quad L_s = 2.$$

$$h = 0,5.$$

$$m = \frac{1,5 + 2}{2} = 1,75.$$

81.4 Parte D

Resolução

a)

$$A = 22 - 5.$$

$$A = 17.$$

b)

Menor:

$$29.$$

Maior:

$$55.$$

Logo:

$$A = 55 - 29.$$

$$A = 26.$$

c)

$$A = 31,8 - 12,5.$$

$$A = 19,3.$$

81.5 Parte E

Resolução

Total:

$$n = 4 + 9 + 12 + 5.$$

$$n = 30.$$

Frequências relativas:

$$\frac{4}{30} \approx 0,1333,$$

$$\frac{9}{30} = 0,3,$$

$$\frac{12}{30} = 0,4,$$

$$\frac{5}{30} \approx 0,1667.$$

Percentuais:

$$13,33\%, \quad 30\%, \quad 40\%, \quad 16,67\%.$$

A maior frequência é:

12.

Logo, a classe modal é:

$$[20, 30).$$

81.6 Parte F

Resolução

$$F_1 = 4.$$

$$F_2 = 4 + 9 = 13.$$

$$F_3 = 4 + 9 + 12 = 25.$$

$$F_4 = 30.$$

Portanto:

$$F_1 = 4, \quad F_2 = 13, \quad F_3 = 25, \quad F_4 = 30.$$

Até 30, temos:

25

observações de um total de:

30.

Logo:

$$\frac{25}{30} \cdot 100\% \approx 83,33\%.$$

$$83,33\%.$$

81.7 Parte G

Resolução

$$[10, 15) : \quad 11, 12, 14.$$

Logo:

$$f_1 = 3.$$

$$[15, 20) : \quad 15, 16, 18, 19.$$

Logo:

$$f_2 = 4.$$

$$[20, 25) : \quad 20, 21, 22, 23, 24.$$

Logo:

$$f_3 = 5.$$

$$[25, 30) : \quad 25, 27, 28.$$

Logo:

$$f_4 = 3.$$

$$[30, 35) : \quad 30, 31, 33, 34.$$

Logo:

$$f_5 = 4.$$

$$[35, 40) : \quad 37.$$

Logo:

$$f_6 = 1.$$

Verificação:

$$3 + 4 + 5 + 3 + 4 + 1 = 20.$$

81.8 Parte H

Resolução

Como:

$$n = 20,$$

temos:

$$f_{r,1} = \frac{3}{20} = 0,15,$$

$$f_{r,2} = \frac{4}{20} = 0,20,$$

$$f_{r,3} = \frac{5}{20} = 0,25,$$

$$f_{r,4} = \frac{3}{20} = 0,15,$$

$$f_{r,5} = \frac{4}{20} = 0,20,$$

$$f_{r,6} = \frac{1}{20} = 0,05.$$

Percentuais:

15%, 20%, 25%, 15%, 20%, 5%.

Classe modal:

$$\boxed{[20, 25)}.$$

Verificação:

$$0,15 + 0,20 + 0,25 + 0,15 + 0,20 + 0,05 = 1.$$

81.9 Parte I

Resolução

Total:

$$n = 2 + 6 + 10 + 8 + 4.$$

$$\boxed{n = 30}.$$

Classe modal:

$$\boxed{[60, 70)}.$$

No intervalo:

$$[60, 80)$$

temos:

$$10 + 8 = 18$$

observações.

Logo:

$$\frac{18}{30} = 0,6.$$

Portanto:

60%.

81.10 Parte J

Resolução

$$f = 0,35 \cdot 80.$$

$$f = 28.$$

Logo:

28.

81.11 Parte K

Resolução

$$f_3 = F_3 - F_2.$$

Logo:

$$f_3 = 29 - 17.$$

$$f_3 = 12.$$

Analogamente:

$$f_4 = F_4 - F_3.$$

$$f_4 = 41 - 29.$$

$$f_4 = 12.$$

81.12 Parte L

Resolução

Tipo de música preferido é uma variável:

qualitativa.

Suas categorias não representam intervalos de uma escala numérica contínua. Portanto, um histograma é inadequado.

Uma escolha apropriada seria:

gráfico de barras.

81.13 Parte M

Resolução

Primeira classe:

$$h_1 = 10 - 0 = 10.$$

Segunda:

$$h_2 = 30 - 10 = 20.$$

Densidades:

$$d_1 = \frac{8}{10} = 0,8.$$

$$d_2 = \frac{12}{20} = 0,6.$$

Logo, em um histograma por densidade:

a primeira barra deve ser mais alta.

81.14 Parte N

Resolução

Sabemos que:

$$x + 8 + 12 + 5 = 30.$$

Logo:

$$x + 25 = 30.$$

Portanto:

$$x = 5.$$

Frequência relativa da primeira classe:

$$\frac{5}{30} = \frac{1}{6}.$$

Logo:

$$f_r = \frac{1}{6} \approx 0,1667.$$

Frequência acumulada da terceira classe:

$$5 + 8 + 12 = 25.$$

Logo:

$$F_3 = 25.$$

Abaixo de 15, temos as três primeiras classes:

$$25$$

observações.

Assim:

$$\frac{25}{30} \cdot 100\% \approx 83,33\%.$$

$$83,33\%.$$

82 Avaliação final

Avaliação final – Variáveis contínuas e classes

Resolva.

1. Explique a diferença entre uma variável quantitativa discreta e uma variável quantitativa contínua.
2. Classifique:
altura, número de filhos, temperatura.

3. Considere a classe:

$$[50, 60).$$

Determine:

- a) limite inferior;
 - b) limite superior;
 - c) amplitude;
 - d) ponto médio.
4. Explique por que:

$$60$$

não pertence à classe:

$$[50, 60).$$

5. Um conjunto possui menor valor:

18

e maior valor:

73.

Determine a amplitude total.

6. Considere:

Classe	Frequência
[0, 10)	5
[10, 20)	10
[20, 30)	15
[30, 40)	10

Determine:

- o total;
 - a frequência relativa de cada classe;
 - a classe modal.
7. Para a tabela anterior, determine as frequências acumuladas.
8. Que porcentagem das observações está abaixo de 30?
9. Explique duas diferenças entre um histograma e um gráfico de barras categórico.
10. Uma distribuição possui classes de amplitudes diferentes. Explique por que a frequência absoluta não deve necessariamente ser usada diretamente como altura das barras de um histograma.

82.1 Gabarito da avaliação final

Resolução

1. Variável discreta assume valores separados e geralmente decorre de contagem. Variável contínua decorre tipicamente de medição e pode assumir valores em intervalos.

2. :

altura $\Rightarrow$ contínua,

número de filhos $\Rightarrow$ discreta,

temperatura $\Rightarrow$ contínua.

3. :

a)

$$\boxed{50};$$

b)

$$\boxed{60};$$

c)

$$60 - 50 = 10,$$

portanto:

$$\boxed{10};$$

d)

$$\frac{50 + 60}{2} = 55,$$

portanto:

$$\boxed{55}.$$

4. Porque o parêntese à direita indica que o limite superior não está incluído:

$$[50, 60) = \{x \in \mathbb{R} \mid 50 \leq x < 60\}.$$

5. :

$$A = 73 - 18.$$

Logo:

$$\boxed{A = 55}.$$

6. Total:

$$n = 5 + 10 + 15 + 10.$$

$$\boxed{n = 40}.$$

Frequências relativas:

$$\frac{5}{40} = 0,125,$$

$$\frac{10}{40} = 0,25,$$

$$\frac{15}{40} = 0,375,$$

$$\frac{10}{40} = 0,25.$$

Portanto:

$$12,5\%, \quad 25\%, \quad 37,5\%, \quad 25\%.$$

Classe modal:

$$[20, 30).$$

7. :

$$F_1 = 5,$$

$$F_2 = 15,$$

$$F_3 = 30,$$

$$F_4 = 40.$$

Portanto:

$$5, \quad 15, \quad 30, \quad 40.$$

8. Abaixo de 30:

$$5 + 10 + 15 = 30.$$

Como:

$$n = 40,$$

temos:

$$\frac{30}{40} = 0,75.$$

Logo:

$$75\%.$$

9. No histograma, o eixo horizontal representa intervalos numéricos e as barras são adjacentes. No gráfico de barras, normalmente representamos categorias distintas e as barras ficam separadas.

10. Quando as amplitudes são diferentes, a área das barras deve representar a frequência. Por isso, utiliza-se a densidade:

$$d_i = \frac{f_i}{h_i}.$$

Assim:

$$h_i d_i = f_i.$$

83 Conclusão

O agrupamento de variáveis contínuas em classes permite resumir grandes conjuntos de dados e revelar características de sua distribuição.

As principais ideias estudadas foram:

- 1) variáveis quantitativas podem ser discretas ou contínuas;
- 2) variáveis contínuas são normalmente obtidas por medição;
- 3) dados contínuos podem ser agrupados em intervalos;
- 4) uma convenção como:

$$[a, b)$$

evita que valores de fronteira sejam contados em duas classes;

- 5) a amplitude total é:

$$A = x_{\max} - x_{\min};$$

- 6) a amplitude de uma classe é:

$$h = L_s - L_i;$$

- 7) o ponto médio é:

$$m = \frac{L_i + L_s}{2};$$

- 8) a frequência relativa é:

$$f_r = \frac{f}{n};$$

- 9) a frequência acumulada soma as frequências até determinada classe;
- 10) a classe modal é aquela de maior frequência;
- 11) histogramas representam distribuições quantitativas agrupadas em classes;
- 12) em histogramas com classes de mesma amplitude, a altura das barras pode representar diretamente a frequência;

- 13) em classes com amplitudes diferentes, é necessário considerar a densidade de frequência para que a área represente corretamente a frequência;
- 14) agrupar dados simplifica a análise, mas provoca perda de informação individual.

Notação

Resumo fundamental:

$$A = x_{\max} - x_{\min}$$

$$h = L_s - L_i$$

$$m = \frac{L_i + L_s}{2}$$

$$f_r = \frac{f}{n}$$

$$f_{\%} = 100 \frac{f}{n}$$

$$F_i = \sum_{j=1}^i f_j$$

e, quando as classes possuem amplitudes diferentes:

$$d_i = \frac{f_i}{h_i}.$$

Verificações:

$$\sum_i f_i = n$$

$$\sum_i f_{r,i} = 1$$

$$\sum_i f_{\%,i} = 100\%.$$

Finalmente:

$$\text{dados contínuos} \longrightarrow \text{classes} \longrightarrow \text{frequências} \longrightarrow \text{histograma.}$$

Fim do curso