

# Tabela de Frequências

- Por ponto (sem perda de informação)
- Por intervalo (com perda de informação)

# Tabela de Frequências

⇒ A **tabela** é uma forma muito útil de resumir a informação sobre uma variável.

⇒ Além de resumir a informação, tem por finalidade:

1. Representar a forma como **os valores da variável se distribuem** (localização da maioria dos valores, simetria, número de picos e formato das caudas)

2. Indicar qual modelo de **distribuição de probabilidade** poderia ser adequado para esses dados, pois **fornece uma ideia empírica da distribuição da população**

# Tabela de Frequências

As características dessas tabelas variam de acordo com o tipo de variável em estudo.

⇒ Se a variável é do tipo **numérica discreta** (com poucos valores), devemos obter as frequências para cada valor dessa variável (**SEM** perda de informação).

⇒ Se a variável é do tipo **numérica contínua**, devemos primeiro construir intervalos e depois obter as frequências para cada intervalo (**COM** perda de informação).

# Tabela de Frequências

⇒ Quando a variável for **numérica discreta** (com poucos valores), a tabela de distribuição de frequências apresentará a seguinte característica: **cada valor da variável constituirá uma classe.**

## Exemplo – VARIÁVEL NUMÉRICA DISCRETA

Variável em estudo: **número de filhos por casal investigado em 350 famílias de uma região.**

**Dados brutos:** 2, 5, 6, 0, 4, 4, 3, 4, 2, 2, 3, 3, 5, 3, 5, 1, 2, 4, 2, 3, 5, 4, 3, 3, 2, 3, 0, 4, 4, 3, 4, 0, 2, 0, 2, 3, 3, 1, 2, 4, 2, ...

# Construção da tabela

Para construir a tabela devemos seguir apenas dois passos:

**1º passo:** Identificar e ordenar as categorias ou valores da variável. Cada categoria ou valor constituirá uma classe.

- ♦ O número da classe é representado por  $j$ , tal que  $j=1, 2, \dots, k$ , onde  $k$  é o número total de classes.

$j$	nº de filhos
1	0
2	1
3	2
4	3
5	4
6	5
7	6

**2º passo:** Contar o número de elementos em cada classe, ou seja, contar quantas vezes o dado está repetido.

j	nº filhos	$F_j$
1	0	55
2	1	60
3	2	112
4	3	82
5	4	31
6	5	8
7	6	2
	$\Sigma$	350

Os valores provenientes desta contagem, denotados por  $F_j$ , são denominados **frequências absolutas das classes**.

# Outras frequências importantes:

**Frequência absoluta acumulada**, denotada por  $F'_j$ , expressa o número de elementos acumulados em cada classe.

j	nº filhos	$F_j$	$F'_j$
1	0	55	55
2	1	60	115
3	2	112	227
4	3	82	309
5	4	31	340
6	5	8	348
7	6	2	350
	$\Sigma$	350	-

# Outras frequências importantes:

**Frequência relativa**, denotada por  $f_j$ , expressa a proporção de elementos em cada classe.

j	nº filhos	$F_j$	$F'_j$	$f_j$
1	0	55	55	0,1571
2	1	60	115	0,1714
3	2	112	227	0,3200
4	3	82	309	0,2343
5	4	31	340	0,0886
6	5	8	348	0,0229
7	6	2	350	0,0057
	$\Sigma$	350	-	1



# Outras frequências importantes:

**Frequência relativa acumulada**, denotada por  $f'_j$ , expressa a proporção de elementos acumulada em cada classe.

j	nº filhos	$F_j$	$F'_j$	$f_j$	$f'_j$
1	0	55	55	0,1571	0,1571
2	1	60	115	0,1714	0,3286
3	2	112	227	0,3200	0,6486
4	3	82	309	0,2343	0,8829
5	4	31	340	0,0886	0,9714
6	5	8	348	0,0229	0,9943
7	6	2	350	0,0057	1
	$\Sigma$	350	-	1	-

# Interpretação

proporção de casais com até 3 filhos

número de casais com até 2 filhos

j	nº filhos	$F_j$	$F'_j$	$f_j$	$f'_j$
1	0	55	55	0,1571	0,1571
2	1	60	115	0,1714	0,3286
3	2	112	227	0,3200	0,6486
4	3	82	309	0,2343	0,8829
5	4	31	340	0,0886	0,9714
6	5	8	348	0,0229	0,9943
7	6	2	350	0,0057	1
	$\Sigma$	350	-	1	-

número de casais com 4 filhos

proporção de casais com 1 filho

## Exercício proposto:

Os dados a seguir se referem ao número diário de pães não vendidos em uma certa padaria até a hora do encerramento do expediente:

0	0	4	2	0	1	0	2	0	4
1	0	0	3	2	0	1	0	0	0
2	0	0	1	0	0	3	2	1	5
0	1	0	0	2	0	0	3	2	1

Construa a distribuição de frequências para esses dados.

# Solução:

nº pães	$F_j$	$F'_j$	$f_j$	$f'_j$
0	20	20	0,5000	0,5000
1	7	27	0,1750	0,6750
2	7	34	0,1750	0,8500
3	3	37	0,0750	0,9250
4	2	39	0,0500	0,9750
5	1	40	0,0250	1
$\Sigma$	40	-	1	-

# Tabela de frequências para variáveis contínuas

Exemplo:

Variável em estudo: valores gastos (em reais) pelas primeiras 50 pessoas que entraram num determinado Supermercado, no dia 01/03/2013.

Dados brutos:

32,03	19,54	45,40	25,13	46,69	18,36	13,78	15,23	36,37	15,62
17,00	27,65	85,76	38,64	86,37	24,58	20,16	93,34	48,65	22,22
23,04	42,97	28,06	52,75	3,11	8,88	9,26	10,81	12,69	28,38
18,43	61,22	41,02	44,67	19,50	17,39	39,16	44,08	38,98	19,27
26,24	28,08	59,07	82,70	26,26	24,47	54,80	70,32	50,39	20,59

As **variáveis contínuas**, em geral, assumem muitos valores diferentes uns dos outros.

⇒ Para contornar problemas desse tipo, as tabelas de distribuição de frequências são construídas de modo que **cada classe seja constituída por um intervalo de valores** da variável.

⇒ Quando variáveis discretas assumem muitos valores diferentes **é usual agrupar os dados discretos em intervalos de classe**.

# Construção da tabela

1º passo: Ordenar o conjunto de dados: colocar os dados brutos em ordem crescente de grandeza (rol).

2º passo: Determinar o número de classes ( $k$ ) da tabela.

De modo geral, esse valor não deverá ser inferior a 5 e nem superior a 15. Essa definição deverá ser orientada pelos objetivos do trabalho, mas existem algumas regras objetivas de determinação:

<div>Arredondar para inteiro</div>	$k = \sqrt{n}$	←	Regra empírica
	$k = 1 + 3,32 \times \log n$	←	Fórmula de Sturges

onde:  $k$ : número de classes (arredondar – inteiro)  
 $n$ : número de observações

3º passo: Determinar a amplitude do intervalo de classe.

Para isso utilizamos a expressão  $i = \frac{a_t}{k}$  

onde:  $i$ : amplitude do intervalo  
 $a_t$ : amplitude total =  $x_{(n)} - x_{(1)}$   $\left\{ \begin{array}{l} x_{(1)} = \text{Extremo Inferior} \\ x_{(n)} = \text{Extremo Superior} \end{array} \right.$

4º passo: Construir os intervalos de classe.

j	Classe
1	$x_{(1)} \text{ — } x_{(1)} + i$
2	$x_{(1)} + i \text{ — } x_{(1)} + 2i$
3	$x_{(1)} + 2i \text{ — } x_{(1)} + 3i$

5º passo: Contar o número de elementos em cada classe.



Na construção dos **intervalos de classe**, é importante observar que:

- Recomenda-se o uso de intervalos de mesma amplitude, mas eventualmente uma amplitude variável poderá ser mais adequada ao contexto;
- Deve ser garantido que todas as observações sejam classificadas;
- As classes são mutuamente exclusivas, ou seja, uma observação pertence a uma única classe;
- Com exceção da última classe, que é fechada à esquerda e à direita, os intervalos são fechados à esquerda e abertos à direita, de modo que um valor que coincida com o extremo superior será classificado na classe seguinte.

# Exemplo:

Os dados abaixo, já ordenados, se referem aos valores gastos (em reais) pelas primeiras 50 pessoas que entraram em um determinado Supermercado, no dia 01/03/2013.

3,11	8,88	9,26	10,81	12,69	13,78	15,23	15,62	17,00	17,39
18,36	18,43	19,27	19,50	19,54	20,16	20,59	22,22	23,04	24,47
24,58	25,13	26,24	26,26	27,65	28,06	28,08	28,38	32,03	36,37
38,98	38,64	39,16	41,02	42,97	44,08	44,67	45,40	46,69	48,65
50,39	52,75	54,80	59,07	61,22	70,32	82,70	85,76	86,37	93,34

Faça a distribuição de frequências desses dados.

## Resolução:

$$n = 50$$

$$k = 1 + 3,32 \times \log n = 1 + 3,32 \times 1,7 = 6,64 \cong 7$$

$$i = \frac{a_t}{k} = \frac{ES - EI}{k} = \frac{93,34 - 3,11}{7} = 12,89$$

Ponto médio ou  
centro de classe

j	Compra (R\$)	$F_j$	$F'_j$	$f_j$	$f'_j$	$c_j$
1	3,11  — 16,00	8	8	0,16	0,16	9,56
2	16,00  — 28,89	20	28	0,40	0,56	22,45
3	28,89  — 41,78	6	34	0,12	0,68	35,34
4	41,78  — 54,67	8	42	0,16	0,84	48,23
5	54,67  — 67,56	3	45	0,06	0,90	61,12
6	67,56  — 80,45	1	46	0,02	0,92	74,01
7	80,45  — 93,34	4	50	0,08	1	86,90
	$\Sigma$	50	-	1	-	-

# Exercício:

Tomemos a seguinte variável:

$X$  = peso ao nascer (em kg) de 60 bovinos machos da raça Ibagé, para a qual os valores observados (e já ordenados) foram:

16, 17, 17, 18, 18, 18, 19, 20, 20, 20, 20, 20,  
21, 21, 22, 22, 23, 23, 23, 23, 23, 23, 23, 23,  
23, 25, 25, 25, 25, 25, 25, 26, 26, 27, 27, 27,  
27, 28, 28, 28, 29, 29, 29, 30, 30, 30, 30, 30,  
30, 30, 31, 32, 33, 33, 33, 34, 34, 35, 36, 39.

Faça a distribuição de frequências desses dados.

## Resolução:

$$n = 60$$

$$k = 1 + 3,32 \times \log n = 1 + 3,32 \times 1,778 = 6,9 \cong 7$$

$$i = \frac{a_t}{k} = \frac{ES - EI}{k} = \frac{39 - 16}{7} = 3,2857 \cong 3,3$$

Ponto médio ou  
centro de classe

j	Peso	$F_j$	$F'_j$	$f_j$	$f'_j$	$c_j$
1	16  — 19,3	7	7	0,1167	0,1167	17,65
2	19,3  — 22,6	9	16	0,1500	0,2667	20,95
3	22,6  — 25,9	15	31	0,2500	0,5167	24,25
4	25,9  — 29,2	12	43	0,2000	0,7167	27,55
5	29,2  — 32,5	9	52	0,1500	0,8667	30,85
6	32,5  — 35,8	6	58	0,1000	0,9667	34,15
7	35,8  — 39,1	2	60	0,0333	1,0000	37,45
$\Sigma$		60	—	1,0000	—	—