

# Métodos Estatísticos Básicos

## Aula 2 - Distribuição de frequências

Prof. Regis Augusto Ely

Departamento de Economia  
Universidade Federal de Pelotas (UFPEL)

Março de 2014

# Distribuição de frequência

- **Tabela primitiva (dados brutos):** é uma tabela com a relação dos dados que utilizaremos, sendo que ela não foi organizada numericamente. Ex: 45, 41, 42, 41, 43, 42, 46.
- **Rol:** é uma tabela obtida após a ordenação dos dados (crescente ou decrescente). Ex: 41, 41, 42, 42, 43, 45, 46.
- **Distribuição de frequência (sem intervalos de classe):** é a condensação dos dados conforme as repetições de seus valores.

Dado	Frequência
41	2
42	2
43	1
45	1
46	1

- **Distribuição de frequência com intervalo de classe:** podemos agrupar os valores em vários intervalos e representá-los da seguinte maneira.

Classes	Frequência
41 $\vdash$ 43	4
43 $\vdash$ 45	1
45 $\vdash$ 47	2

Onde o símbolo  $\vdash$  significa que incluímos na classe o valor à esquerda e excluimos o valor à direita.

# Elementos de uma distribuição de frequência com intervalos de classe

- **Classe:** são os intervalos que utilizamos para organizar os dados. Denotamos por  $K$  o número total de classes e  $i$  o número da classe. Assim, na tabela anterior temos  $K=3$  e  $43 \vdash 45$  é a segunda classe, com  $i = 2$ .
- **Limites de classe:** são os extremos de cada classe. O menor número é o limite inferior da classe ( $l_i$ ) e o maior é o limite superior ( $L_i$ ). Ex: em  $43 \vdash 45$ , temos  $l_2 = 43$  e  $L_2 = 45$ .
- **Amplitude do intervalo de classe:** é a diferença entre o limite superior e inferior da classe. Assim,  $h_i = L_i - l_i$ . Ex: em  $43 \vdash 45$  temos  $h_2 = 2$ . Note que na distribuição de frequência com classe, o  $h_i$  será igual em todas as classes.

# Elementos de uma distribuição de frequência com intervalos de classe

- **Amplitude total da distribuição:** é a diferença entre o limite superior da última classe e o limite inferior da primeira classe.  
 $AT = L(max) - l(min)$ . Ex: na tabela anterior,  $AT = 47 - 41 = 6$ .
- **Amplitude total da amostra (Rol):** é a diferença entre o valor máximo e o valor mínimo da amostra (rol).  $AA = X_{max} - X_{min}$ . Ex: na tabela anterior,  $AA = 46 - 41 = 5$ . Note que AT será sempre maior ou igual a AA.
- **Ponto médio de classe:** é o ponto que divide o intervalo de classe em duas partes iguais  $X_i = \frac{(l_i + L_i)}{2}$ . Ex: em  $43 \vdash 45$  temos  $\frac{43+45}{2} = 44$ .

# Método para construção de uma distribuição de frequências com classe

- 1) Organize os dados brutos em Rol.
- 2) Calcule a amplitude amostral AA.
- 3) Defina o número de classes (é comum utilizarmos a Regra de Sturges):

$n$	$K$ ( $n^\circ$ de classes)
$3 \leq n < 5$	3
$6 \leq n < 11$	4
$12 \leq n < 22$	5
$23 \leq n < 46$	6
$47 \leq n < 90$	7
$91 \leq n < 181$	8
$182 \leq n < 362$	9

Essa regra é apenas uma sugestão, sendo  $n$  o número de dados distintos na amostra.

# Método para construção de uma distribuição de frequências com classe

- 4) Decidido o número de classes, calcular a amplitude de intervalo de classe, sendo  $h > \frac{AA}{K}$ . Como  $AA = 5$ , no nosso exemplo, e  $K=3$ , então  $h > \frac{5}{3} = 1,66$ ; e podemos escolher  $h=2$ .
- 5) Agora montamos a tabela, cuidando para não aparecer classes com frequência zero. A primeira classe inicia no menor número da amostra até esse mesmo número mais  $h$ . Assim, no exemplo,  $41 \vdash 41 + h \Rightarrow 41 \vdash 43$ . E assim por diante.

# Definições de frequências

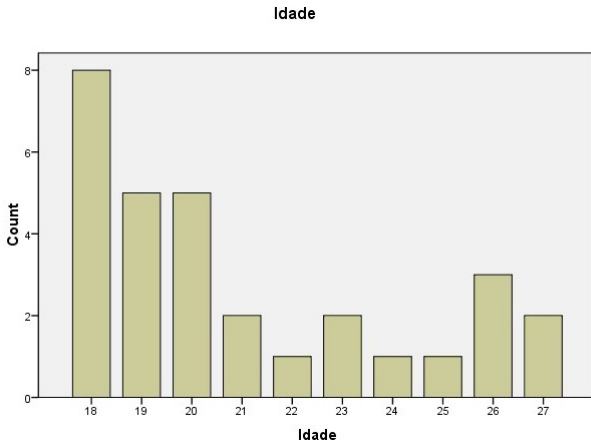
- **Frequência simples ou absoluta:** são os valores que representam o número de dados de cada classe, denotado por  $f_i$ . Temos  $\sum_i f_i = n$ .
- **Frequência relativa:** são os valores das razões entre as frequências absolutas de cada classe e a frequência total da distribuição, denotada por  $fr_i$ . Temos  $\sum_i fr_i = 1$ .
- **Frequência simples acumulada:** é o total das frequências de todos os valores inferiores ao limite superior do intervalo de uma determinada classe, denotada por  $F_i$ .
- **Frequência relativa acumulada:** é a frequência acumulada da classe, dividida pela frequência total da distribuição, denotada por  $Fr_i$ .
- Assim, se os dados são 41, 41, 42, 42, 43, 45, 46, temos:

Classes	$f_i$	$fr_i$	$F_i$	$Fr_i$	$X_i$
41 ┤ 43	4	4/7	4	4/7	42
43 ┤ 45	1	1/7	5	5/7	44
45 ┤ 47	2	2/7	7	1	46
Total	7	1			



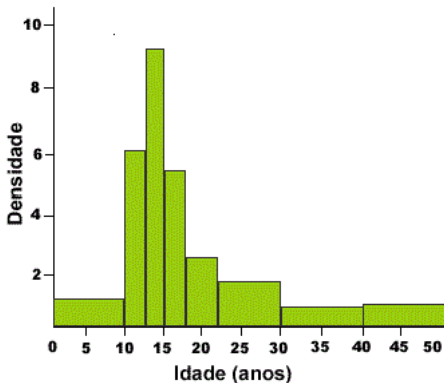
# Representação gráfica

- A representação gráfica identifica as características da distribuição dos dados, como dispersão e simetria.
- **Gráficos de barras:** utilizado para a representação das distribuições de frequência de variáveis discretas.



# Representação gráfica

- **Histograma:** utilizado para a representação das distribuições de frequência de variáveis contínuas.
- Note que para montarmos o histograma devemos primeiro agrupar os dados em intervalos de classe.



# Representação gráfica

- **Polígono de frequência:** é semelhante ao histograma, porém utilizamos linhas para conectar as frequências de cada intervalo de classe.

