

8. **Intervalos de Classes Desiguais** – É muito comum o uso de classes com tamanhos desiguais no agrupamento dos dados em tabelas de freqüências. Nestes casos deve-se tomar alguns cuidados especiais quanto à análise e construção do histograma.

A tabela abaixo fornece a distribuição de 250 empresas classificadas segundo o número de empregados. Uma análise superficial pode levar à conclusão de que a concentração vem aumentando até atingir um máximo na classe 40 — 60, voltando a diminuir depois, mas não tão acentuadamente. Porém, um estudo mais detalhado revela que a amplitude da classe 40 — 60 é o dobro da amplitude das classes anteriores. Assim, espera-se que mais elementos caiam nessa classe, mesmo que a concentração seja levemente inferior. Então, um primeiro cuidado é construir a coluna que indica as amplitudes Δ_i de cada classe. Estes valores estão representados na terceira coluna da tabela.

Distribuição de 250 empresas segundo o número de empregados

<i>Número de empregados</i>	<i>Freqüência n_i</i>	<i>Amplitude Δ_i</i>	<i>Densidade n_i/Δ_i</i>	<i>Proporção f_i</i>	<i>Densidade f_i/Δ_i</i>
0 — 10	5	10	0,50	0,02	0,0020
10 — 20	20	10	2,00	0,08	0,0080
20 — 30	35	10	3,50	0,14	0,0140
30 — 40	40	10	4,00	0,16	0,0160
40 — 60	50	20	2,50	0,20	0,0100
60 — 80	30	20	1,50	0,12	0,0060
80 — 100	20	20	1,00	0,08	0,0040
100 — 140	20	40	0,50	0,08	0,0020
140 — 180	15	40	0,38	0,06	0,0015
180 — 260	15	80	0,19	0,06	0,0008
TOTAL	250	—	—	1,00	—

Um segundo passo é a construção da coluna das densidades de freqüências em cada classe, que é obtida dividindo as freqüências n_i pelas amplitudes Δ_i . Ou seja, a medida que indica qual a concentração por unidade da variável. Assim, observando-se os números da quarta coluna, vê-se que a classe de maior concentração passa a ser a 30 — 40, enquanto que a última é a de menor concentração. Para compreender a distribuição, estes dados são muito mais informativos do que as freqüências absolutas simplesmente.

De modo análogo, pode-se construir a densidade da proporção (ou porcentagem) por unidade da variável (verifique a construção através da 5.^a e 6.^a colunas). A interpretação para f_i/Δ_i é muito semelhante àquela dada para n_i/Δ_i .

Para a construção do histograma, basta lembrar que a área total deve ser igual a 1 (ou 100%), o que sugere usar no eixo das ordenadas os valores de f_i/Δ_i . O histograma para estes dados está na Figura 1.10.

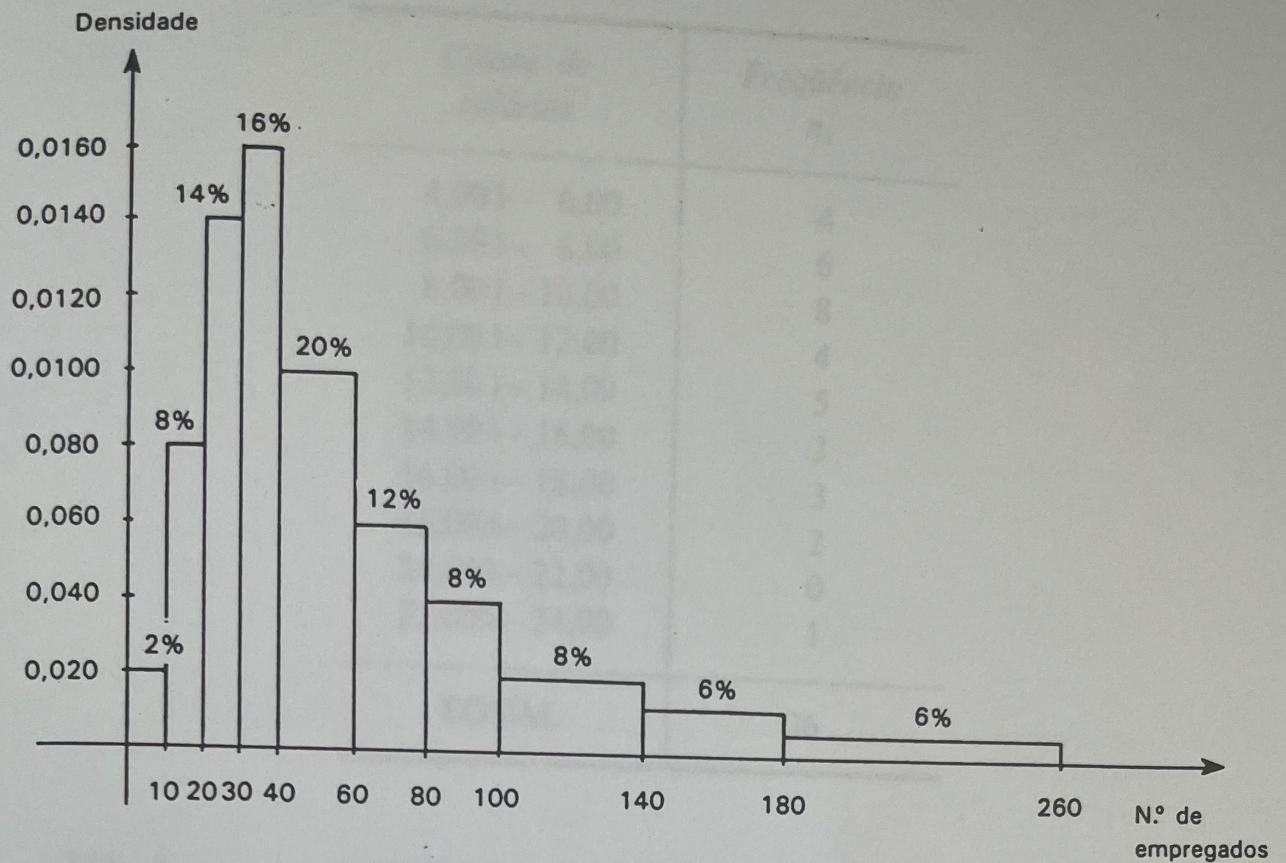


Fig. 1.10

9. Dispomos de uma relação de 200 aluguéis de imóveis urbanos e uma relação de 100 aluguéis rurais.

(a) Construa os histogramas das duas distribuições.

(b) Com base nos histogramas discuta e compare as duas distribuições.

<i>Classes de aluguéis (codificados)</i>	<i>Zona urbana</i>	<i>Zona rural</i>
2 3	10	30
3 5	40	50
5 7	80	15
7 10	50	5
10 15	20	0
TOTAL	200	100

PROBLEMAS

1. Quer se estudar o número de erros de impressão de um livro. Para isso escolheu-se uma amostra de 50 páginas, encontrando-se o seguinte número de erros por página:

(a) Qual o número médio de erros por página?

(b) E o número mediano?

(c) Qual é o desvio padrão?

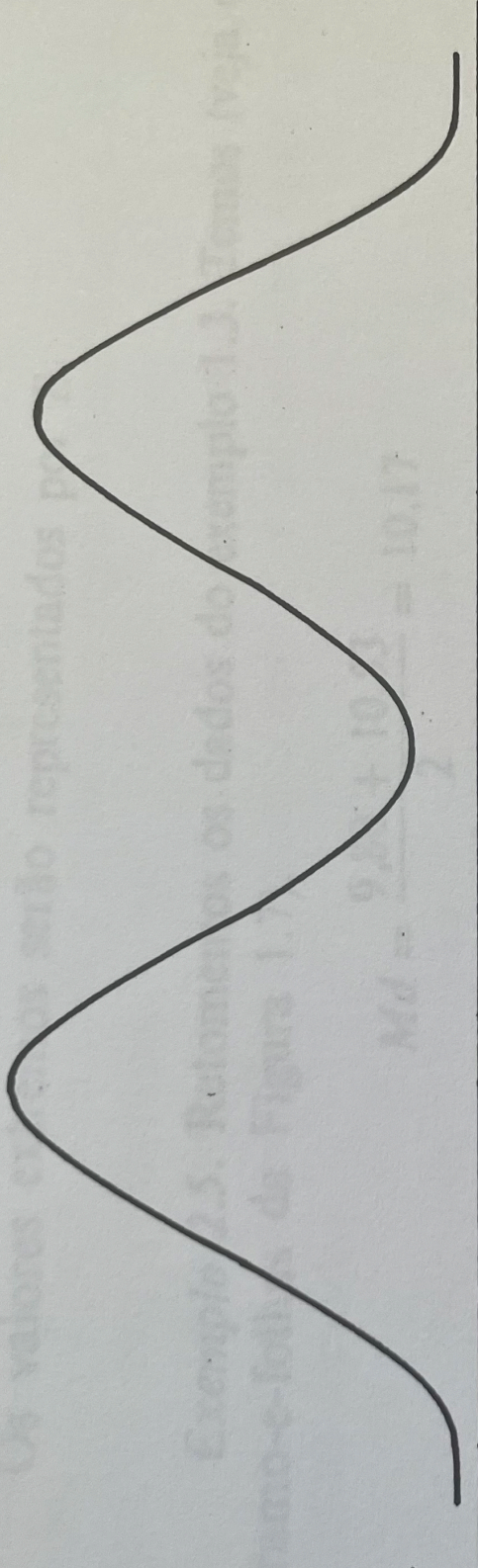
(d) Faça uma representação gráfica para a distribuição.

(e) Se o livro tem 500 páginas, qual o número total de erros esperado no livro?

<i>Erros</i>	<i>Frequência</i>
0	25
1	20
2	3
3	1
4	1

2. As taxas de juros recebidas por 10 ações durante um certo período foram (medidas em porcentagem) 2,59; 2,64; 2,60; 2,62; 2,57; 2,55; 2,61; 2,50; 2,63; 2,64. Calcule a média, a mediana e o desvio padrão.

5. Suponha que a variável de interesse tenha a distribuição como na figura abaixo.



Você acha que a média é uma boa medida de posição? E a mediana? Justifique.

6. Numa pesquisa realizada com 100 famílias levantaram-se as seguintes informações:

Número de filhos	0	1	2	3	4	5	mais que 5
Frequência de famílias	17	20	28	19	7	4	5

(a) Qual a mediana do número de filhos?

(b) E a moda?

(c) Que problemas você enfrentaria para calcular a média? Faça alguma suposição e encontre-a.