

FORMAS DE APRESENTAÇÃO DOS DADOS

Vimos, nos itens I, V e VI das aulas sobre Medidas de Posição (Partes I e II), diversos exemplos na forma de apresentação dos dados e concluímos, então, que o modo de encontrar essas medidas (Média, Moda e Mediana) dependerá de como os dados se apresentem. Foram vistas as seguintes formas de apresentação dos dados:

- ROL (dados isolados)
- AGRUPAMENTO SIMPLES
- AGRUPAMENTO EM CLASSES

Mas há ainda uma outra forma de apresentação de dados que tem sido cobrada freqüentemente em provas recentes da ESAF: o **DIAGRAMA DE RAMOS E FOLHAS**. Podemos dar, como exemplo, uma das questões da prova para Analista do IRB em 2004:

"O diagrama de ramos e folhas apresentado abaixo corresponde à seqüência de observações amostrais (34, 38, ..., 97) de um atributo X. Assinale a opção que dá a mediana amostral de X."

| | |
|---|---------|
| 3 | 4 |
| 3 | 8 |
| 4 | 22 |
| 4 | 57 |
| 5 | 124 |
| 5 | 7889 |
| 6 | 013 |
| 6 | 5567899 |
| 7 | 0112334 |
| 7 | 556679 |
| 8 | 1123344 |
| 8 | 57 |
| 9 | 0133 |
| 9 | 7 |

- (a) 69,5
- (b) 71,0
- (c) 70,5
- (d) 72,0
- (e) 74,0

Observe então, que os números que aparecem à esquerda no diagrama são os ramos e os números à direita são as folhas. Juntando cada uma das folhas aos ramos, teremos o rol que originou o diagrama e que é citado resumidamente no início da questão.

Assim, a seqüência de observações é: 34, 38, 42, 42, 45, 47, 51, 52, 54, 57, 58, 58, 59, 60, 61, 63, 65, 65, 66, 67, 68, 69, 69, 70, 71, 71, 72, 73, 73, 74, 75, 75, 76, 76, 77, 79, 81, 81, 82, 83, 83, 84, 84, 85, 87, 90, 91, 93, 93, 97. Contando o número n de observações, temos que $n = 50$.

Como já foi visto anteriormente, quando o número de observações é par, a mediana será a média aritmética entre os dois termos centrais que, no caso, são o 25º e o 26º termos, ambos iguais a 71. Este será o valor da mediana (gabarito da letra b).

Todavia, poderíamos também transformar o diagrama de ramos e folhas em um agrupamento simples, colocando na 1ª coluna os valores observados (junção da folha com o ramo), na 2ª coluna a frequência simples e na 3ª coluna a frequência acumulada crescente. Resumidamente, teremos:

| X_i | F_i | F_{ac} |
|----------|-------|----------|
| 34 | 1 | 1 |
| 38 | 1 | 2 |
| 42 | 2 | 4 |
| 45 | 1 | 5 |
| . | . | . |
| . | . | . |
| . | . | . |
| 70 | 1 | 24 |
| 71 | 2 | 26 |
| . | . | . |
| . | . | . |
| . | . | . |
| 97 | 1 | 50 |
| Σ | 50 | - |

—————> O 25º e o 26º elemento são iguais a 71

Entretanto, há ainda uma forma mais prática e rápida de encontrar a mediana quando for dado um diagrama de ramos e folhas. Basta aproveitar o próprio diagrama, criando nele uma coluna de frequência simples (onde a frequência será dada pelo número de folhas em cada ramo) e outra de frequência acumulada crescente. Assim, teremos:

| | F_i | F_{ac} |
|---|---------|----------|
| 3 | 4 | 1 |
| 3 | 8 | 1 |
| 4 | 22 | 2 |
| 4 | 57 | 2 |
| 5 | 124 | 3 |
| 5 | 7889 | 4 |
| 6 | 013 | 3 |
| 6 | 5567899 | 7 |
| 7 | 0112334 | 7 |
| 7 | 556679 | 6 |
| 8 | 1123344 | 7 |
| 8 | 57 | 2 |
| 9 | 0133 | 4 |
| 9 | 7 | 1 |

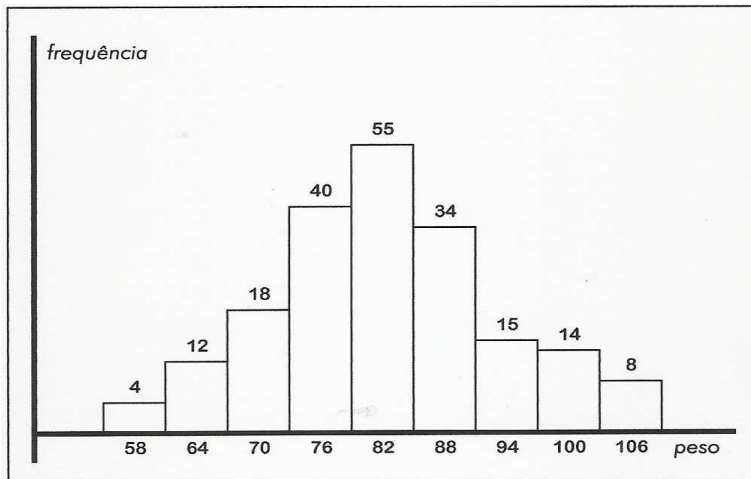
Observamos então que, até a última folha do 2º ramo com valor 6 acumulamos 23 observações (a 23ª observação será igual a 69). Logo, a 1ª folha (0) do 1º ramo com valor 7 será o 24º termo (igual a 70), a próxima folha (1) será o 25º termo (igual a 71) e a folha seguinte (1) será o 26º termo (igual a 71). Portanto, a Mediana é igual a 71.

Concluimos então que as formas de apresentação: ROL, AGRUPAMENTO SIMPLES e RAMOS E FOLHAS podem comunicar-se entre si, ou seja, podemos transformar:

- ROL em um AGRUPAMENTO SIMPLES;
- AGRUPAMENTO SIMPLES em um ROL;
- ROL em um DIAGRAMA DE RAMOS E FOLHAS;
- DIAGRAMA DE RAMOS E FOLHAS em um ROL;
- AGRUPAMENTO SIMPLES em um DIAGRAMA DE RAMOS E FOLHAS;
- DIAGRAMA DE RAMOS E FOLHAS em um AGRUPAMENTO SIMPLES;

Outra forma de apresentação dos dados em questões de provas tem sido o HISTOGRAMA. Podemos dar como exemplo uma das questões da prova de Análise de Métodos Quantitativos para o IBGE em 2002, elaborada pelo NCE/UFRJ:

“O histograma a seguir apresenta dados a respeito de uma amostra de *pesos*, em kg, de duzentos homens:



Os números indicados no eixo dos *pesos* são os pontos médios de cada intervalo. O *peso* médio desta amostra, obtido a partir destes dados agrupados, em kg, é igual a:”

- (a) 75,6
- (b) 78,9
- (c) 79,1
- (d) 82,3
- (e) 84,2

Podemos encontrar a Média e responder à questão multiplicando cada ponto médio dado pela respectiva freqüência e, a seguir, dividir o somatório desses produtos (igual a 16.460) pela freqüência total, dada no enunciado, de 200 homens, encontrando então a Média de 82,3 (gabarito da letra d). Entretanto, podemos transformar esse histograma num agrupamento em classes, o que facilitará o cálculo das outras medidas (Moda, Mediana, Quartis, Decis e Percentis) e também para resolvermos a questão pelo método simplificado para cálculo da Média, a ser visto mais adiante:

| Classes de Peso | Freqüências (F_i) | Ponto Médio da classe = X _i | X _i · F _i |
|----------------------------|--|--|---------------------------------|
| 55 – 61 | 4 | 58 | 232 |
| 61 – 67 | 12 | 64 | 768 |
| 67 – 73 | 18 | 70 | 1.260 |
| 73 – 79 | 40 | 76 | 3.040 |
| 79 – 85 | 55 | 82 | 4.510 |
| 85 – 91 | 34 | 88 | 2.992 |
| 91 – 97 | 15 | 94 | 1.410 |
| 97 – 103 | 14 | 100 | 1.400 |
| 103 – 109 | 8 | 106 | 848 |
| Σ | 200 | - | 16.460 |

Vemos então que, assim como podemos transformar um AGRUPAMENTO EM CLASSES em um HISTOGRAMA, podemos transformar um HISTOGRAMA em um AGRUPAMENTO EM CLASSES.

Sintetizando o modo de encontrar as medidas de posição de acordo com a forma de apresentação dos dados, vemos que as medidas descritas abaixo devem ser obtidas:

| Quando os dados se apresentarem em: | Média | Moda | Mediana | Quartis, Decis e Percentis |
|--|---|---------------------------|---------------------------|-----------------------------------|
| ROL | $\bar{X} = \frac{\sum X_i}{n}$ | Pela observação dos dados | Pela observação dos dados | Pela observação dos dados |
| AGRUPAMENTO SIMPLES | $\bar{X} = \frac{\sum X_i \cdot F_i}{\sum F_i}$ | Pela observação dos dados | Pela observação dos dados | Pela observação dos dados |
| RAMOS E FOLHAS | $\bar{X} = \frac{\sum X_i}{n}$ | Pela observação dos dados | Pela observação dos dados | Pela observação dos dados |
| AGRUPAMENTO EM CLASSES | $\bar{X} = \frac{\sum X_i \cdot F_i}{\sum F_i}$ | Fórmula (*) | Fórmula (**) | Fórmula (***) |

$$(*) \rightarrow Mo = l + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot h$$

$$(**) \rightarrow Md = l + \frac{\left(\frac{n}{2} - \sum f\right)}{F_{Md}} \cdot h$$

$$(***) \rightarrow Qi = l + \frac{\left(\frac{in}{4} - \sum f\right)}{F_{Qi}} \cdot h; \quad Di = l + \frac{\left(\frac{in}{10} - \sum f\right)}{F_{Di}} \cdot h; \quad Pi = l + \frac{\left(\frac{in}{100} - \sum f\right)}{F_{Pi}} \cdot h$$

OBS.: Lembrando que, na fórmula da média para agrupamento em classes, temos que considerar como X_i o ponto médio de cada classe.