ESTATÍSTICA

na Contabilidade – Revisão - Parte 3

Luiz A. Bertolo

Medidas Estatísticas

A distribuição de frequências permite-nos descrever, de modo geral, os grupos de valores (*classes*) assumidos por uma variável. Com ela, por exemplo, podemos localizar se a maior concentração de valores de uma dada distribuição se encontra no início, no meio, ou no final dos valores.

Quando confrontamos distribuições e queremos destacar as tendências de cada uma, isoladamente, necessitamos de conceitos que expressem através de números estas tendências. Esses conceitos são denominados **elementos típicos** da distribuição (ou *estatísticas*) e são:

- Medidas de Posição (locação ou tendência central)
- Medidas de Dispersão (variabilidade)
- Medidas de Assimetria
- Medidas de Curtose

Cap.3 – Medidas de Posição (ou tendência central)

Elas mostram o valor representativo em torno do qual os dados tendem a agrupar-se com maior ou menor freqüência.

A medida de tendência central é um número que está representando todo o conjunto de dados, isto é, ela resume o conjunto de dados; nas pesquisas tal número pode ser encontrado a partir das medidas:

a) **média aritmética,**

b) **moda,**

c) **mediana.**

O uso de cada uma delas é mais conveniente de acordo com o nível de mensuração, o aspecto ou forma da distribuição de dados e o objetivo da pesquisa.

Outras medidas de posição são as **separatrizes**, que englobam:

- a própria mediana;
- os quartis;
- os percentis.

Médiana ($\tilde{x}$) p. 52 e 53

É o valor "do meio" de um conjunto de dados, quando os dados estão dispostos em ordem crescente ou decrescente, cortando, assim, a distribuição em duas partes com o mesmo número de elementos.

É também uma medida separatriz definida e exata, de fácil compreensão. Ela serve para análise comparativa e é representada por $\tilde{x}$ ou Md.

Para dados **não agrupados** em classes:

Se n é **ímpar** → é o termo $\tilde{x} = \left(\frac{n+1}{2}\right)^{o} termo$ POSIÇÃO da Mediana

Se *n* é **par** → é o termo $\tilde{x} = \dfrac{\frac{n}{2} termo + \left(\frac{n}{2}+1\right) termo}{2}$ POSIÇÃO da Mediana

EX1: Em um colégio, estão matriculados numa determinada classe 21 alunos. Durante o 1º bimestre foi feito um levantamento da freqüência destes alunos e foram observadas as seguintes faltas: 0, 0, 3, 5, 7, 9, 0, 1, 2, 3, 11, 2, 3, 5, 6, 4, 10, 12, 0, 1, 2. Qual a mediana das faltas? Dica: Primeiro construa o ROL.

Resposta: 0, 0, 0, 0, 1, 1, 2, 2, 2, 3, 3, 3, 4, 5, 5, 6, 7, 9, 10, 11, 12....temos 21 elementos (21+1)/2 = 11ª posição = 3

EX2: As idades dos atletas amadores de uma determinada modalidade esportiva são 14, 12, 16, 13, 17, 16 anos. Encontre a mediana da série. Dica: Primeiro construa o ROL

Resposta: 12, 13, 14, 16, 16, 17 ....temos 6 elementos {(6/2)+ [(6/2)+1]}/2 = 15 anos

Médiana ($\tilde{x}$) p. 54

Dados agrupados

Se os dados se agrupam em uma *distribuição de frequência*, o cálculo da mediana se processa de modo muito semelhante àquele dos dados não-agrupados, implicando, porém, a determinação prévia das **frequências acumuladas**. Ainda aqui, temos que determinar um valor tal que divida a distribuição em dois grupos que contenham o mesmo *número de elementos*.

Para o caso de uma distribuição, porém, a **ordem** ou **posição**, a partir de qualquer um dos extremos, é dada por:

$$\boxed{\frac{\Sigma f_i}{2}}$$

Dados Agrupados Sem intervalos de classe ou Variável Discreta

Neste caso, é o bastante identificar a frequência acumulada **imediatamente superior** à metade da soma das frequências (ordem). A mediana será aquele valor da variável que corresponde a tal frequência acumulada.

Por exemplo,

Nº de filhos	f_i	FA_i
0	2	2
1	6	8
2	10	18
3	12	30
4	4	34
	$\Sigma = 34$	

Sendo

$$\frac{\Sigma f_i}{2} = \frac{34}{2} = 17$$

A menor freqüência acumulada que supera esse valor é 18, que corresponde ao valor 2 da variável nº de filhos, sendo este o valor mediano. Logo, Md = 2 filhos.

Médiana ($\tilde{x}$) cont...

Dados Agrupados Com Intervalos de classe

Neste caso, o problema consiste em determinar o ponto do intervalo em que está compreendida a mediana. Para tanto, temos inicialmente que determinar a classe na qual se acha a mediana – classe mediana. Tal classe será, evidentemente, aquela correspondente à frequência acumulada imediatamente superior a $\frac{\Sigma f_i}{2}$.

Seja a distribuição:

i	X	f_i	F_i
1	150 |— 154	4	4
2	154 |— 158	9	13
3	158 |— 162	11	24 → Classe mediana
4	162 |— 166	8	32
5	166 |— 170	5	37
6	170 |—174	3	40

Temos:

$$\frac{\Sigma f_i}{2} = \frac{40}{2} = 20$$

POSIÇÃO

Encontramos a Classe Mediana. E o valor da **MEDIANA** ?

Médiana ($\tilde{x}$) cont...

Existem três (03) maneiras de encontrarmos o valor da mediana quando os dados estão agrupados com intervalo de classe:

#01. Como há 24 valores incluídos nas três primeiras classes da distribuição e como pretendemos determinar o valor que ocupa o 20º lugar, a partir do início da série, vemos que este deve estar localizado na terceira classe (i = 3), supondo que as frequências dessas classes estejam uniformemente distribuídas.

Como há 11 elementos nessa classe e o intervalo da classe é igual a 4, devemos tomar, a partir do limite inferior, a distância:

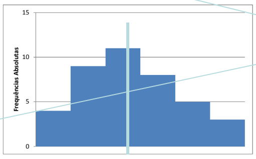
11 ....4 $\Rightarrow x = \frac{7}{11} . 4 = 2{,}54$

7 ... x

E a mediana será dada por:

$$\tilde{x} = Md = 158 + 2{,}54 = 160{,}54$$

#02. Poderíamos num histograma determinar graficamente a mediana como sendo aquele ponto do eixo das abcissas por onde passa a vertical que divide o histograma em duas áreas iguais:



4 . 4 + 4 . 9 + x . 11 = (4-x) . 11 + 4 . 8 + 4 . 5 + 4 . 3 $\Rightarrow$ 52 + 11x = 44 – 11x + 64 ou 22x = 56 $\Rightarrow$ x = 2,5454

Md = 158 + 2,5454 = **160,54.**

Médiana ($\tilde{x}$) cont...

#03. Existe, também, uma fórmula para calcularmos a mediana diretamente da tabela de distribuição de frequências:

$$Md = l_i^* + \frac{\left[\frac{\sum f_i}{2} - FA_{anterior}\right] h}{f^*}$$

Onde: l_i^* é o **limite inferior** da classe mediana;
$FA_{anterior}$ é a frequência acumulada da classe anterior à classe mediana;
f^* é a *frequência absoluta* da classe mediana;
h é a amplitude do intervalo da classe mediana.

No exemplo anterior:

$$\frac{\sum f_i}{2} = \frac{40}{2} = 20$$

$$Md = 158 + \frac{[20 - 13]4}{11} = 160{,}54$$

i	X	f_i	F_i
1	150 |— 154	4	4
2	154 |— 158	9	13
3	158 |— 162	11	24 → Classe mediana
4	162 |— 166	8	32
5	166 |— 170	5	37
6	170 |—174	3	40

Situação-Problema 02 Dados Agrupados Sem Intervalos de Classe

O rótulo em uma embalagem de detergentes líquido afirma que a embalagem contém 500 ml. O controle de qualidade do fabricante, ao inspecionar a linha de produção, seleciona uma amostra de 12 recipientes desse produto, para verificar se os recipientes estão dentro das normas especificadas no rótulo da embalagem. Os dados obtidos estão registrados no quadro 8. Sabendo-se que o controle de qualidade aceita uma produção que esteja dentro de 500,02 ml do valor de mililitros especificados no rótulo do recipiente, determine o valor médio e mediano dos dados amostrais. Interprete os resultados obtidos.

Quadro 8 Dados de mililitros de embalagens de detergente líquido

500,03	500,00	500,04	500,02
500,02	500,05	500,03	500.00
500,02	500,03	500,00	500,02

Fonte: dados simulados

Que tal fazermos neste momento uma reflexão a respeito de dois pontos de vista distintos nesta situação:

Olhar do Consumidor: quais benefícios você observa nessa realidade apresentada?

Olhar do Gestor: quais os impactos positivos e negativos você nota quando analisa essa situação pelo ponto de vista da empresa?

Tabela 8 Cálculos da média em mililitros.

x_i	f_i	$x_i f_i$	F_i
500,00	3	1500,00	3
500,02	**4**	**2000,08**	**7**
500,03	3	1500,09	10
500,04	1	500,04	11
500,05	1	500,05	12
$\sum$	12	6000,26	-

Fonte: dados simulados

$$\overline{X} = \frac{\sum x_i f_i}{n} = \frac{6000,26}{12} \cong 500,02$$

Portanto, em média, o valor de mililitros por embalagem de detergente é de aproximadamente 500,02 ml, embora o rótulo da embalagem de detergente líquido afirme que o conteúdo é de 500 ml por embalagem. E, sabendo-se que o controle de qualidade aceita uma produção que esteja dentro de 500,02 ml do valor de mililitros especificados no rótulo do recipiente, conclui-se que através da média o resultado satisfaz o volume aceito pelo controle de qualidade.

Situação-Problema 02 – Cont...

Sempre ordene os valores. Assim, em *ordem crescente* temos: 500; 500; 500; 500;02; 500,02; 500,02; 500,02; 500,03; 500,03; 500,03; 500,04; 500,05.
Como temos

$\frac{n}{2} = \frac{12}{2} = 6$ e $\frac{n}{2} + 1 = 6 + 1 = 7$ Logo, a mediana apresenta-se entre o 6º e 7º elementos, ou seja, entre 500,02 e 500,02. Assim, o valor mediano é:

$$Md = \frac{500,02 + 500,02}{2} = 500,02$$

Portanto, concluiu-se que 50% do valor de mililitros por embalagem de detergente apresentam-se *menores*, ou *iguais* a 500,02 ml e, 50% dos preços do valor de mililitros por embalagem de detergente são *maiores* ou *iguais* a 500,02.

Situação-Problema 03 Dados Agrupados Com Intervalos de Classe

Uma comissão criada por donas de casa, "*Estamos de Olho no Preço*", faz a cotação e divulga entre os moradores do bairro onde residem os preços de produtos em geral. A tabela 7 apresenta os dados referentes à variação de preço do macarrão em 19 supermercados ou lojas do gênero, pesquisados durante 6 meses. A partir dos dados registrados na tabela, a que conclusão a comissão de donas de casa chegou a respeito do preço mediano do pacote de macarrão?

Tabela 7 Intervalo de classes em função do preço do pacote de macarrão.

Classes	Preço	f_i	F_i
1	1,32 I--- 1,34	2	2
2	1,34 I--- 1,36	5	7
3	1,36 I--- 1,38	8	15
4	1,38 I--- 1,40	3	18
5	1,42 I--- 1,44	1	19
Total (n)	-	19	-

Fonte: dados simulados

Usaremos a expressão:

$$Md = l_{inf_{Md}} + \frac{\frac{n}{2} - F_{anterior}}{f_{i_{Md}}} \times h$$

Limite inferior da classe mediana

Frequência da classe mediana

Como temos n = 19, $\frac{n}{2} = \frac{19}{2} = 9,5$

A seguir, identificamos a **classe da mediana** através da frequência acumulada F e, como podemos observar o 10º elemento pertence a **3ª classe**. Substituindo-se os valores conhecidos na expressão da mediana, encontramos:

$$Md = l_{inf_{Md}} + \frac{\frac{n}{2} - F_{anterior}}{f_{i_{Md}}} xh = 1,36 + \frac{9,5-7}{8} x0,02 = 1,36 + 0,00625$$

$$Md = 1,36625$$

Portanto, a comissão de donas-de-casa concluiu que 50% dos preços do pacote de macarrão pesquisados apresentam-se *menores*, ou *iguais* a \$1,37 e, 50% dos preços de macarrão pesquisados são *maiores* ou *iguais* a \$1,37.

Média versus Mediana

A média é muito sensível a valores extremos de um conjunto de observações, enquanto a mediana não sofre muito com a presença de alguns valores muito altos ou muito baixos. A mediana é mais "robusta" do que a média.

Devemos preferir a **mediana** como <u>medida sintetizadora</u> quando o histograma do conjunto de valores é **assimétrico**

Ex.: { 200, 250, 250, 300, 450, 460, 510 }
$\bar{x}$=345,7 $\tilde{x}$=300
Tanto $\bar{x}$ como $\tilde{x}$, são boas medidas de posição.

Ex.: { 200, 250, 250, 300, 450, 460, 2300 }
$\bar{x}$ = 601 $\tilde{x}$ = 300
Devido ao valor 2300, $\tilde{x}$ é preferível a $\bar{x}$.

Moda e Classe Modal

É o valor que ocorre com **maior freqüência** em um conjunto de observações individuais. Para dados agrupados temos a **classe modal**. Em alguns casos pode haver mais de uma moda. Assim temos uma distribuição bimodal, trimodal, etc...

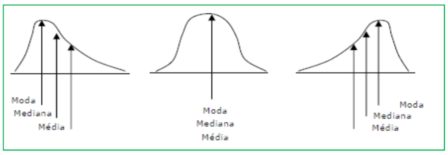
A moda é o valor em torno do qual os dados estatísticos tendem a estar mais pesadamente concentrados e é representada por Mo, também conhecida pelo nome de norma ou modo.

Situalçao 1 - Exemplos para dados NÃO agrupados

01 - Em um grupo de pessoas cujas idades são: 3, 2, 5, 2, 6, 2, 4, 4, 2, 7, 2 anos, a moda é **2** anos (Mo = 2). Portanto, denomina-se **unimodal**.

02 - Algumas pessoas freqüentaram a escola por estes números de anos: 5, 3, 7, 5, 5, 8, 5, 3, 1, 1, 3, 3, 10, 3, 5. Nesta série de números, podem-se ter duas modas: Portanto **bimodal**.

03 - Seja a amostra 4, 6, 9, 11,13, e 14. Observe que *todos os elementos apresentam a mesma freqüência*, logo a amostra é **amodal**.



Comparação das medidas!!

MODA – Cont...

Situação 2: Dados agrupados **sem intervalo**

Nesta situação, basta identificar o elemento de maior frequência. Observe a exemplificação da tabela 9, com o correspondente valor da moda.

Tabela 9 Distribuição de frequência

	1	5	**6**	8	13
f_i	1	3	**5**	2	4

Fonte: dados simulados

Portanto, a moda, Mo = 6.

Situação 3: Dados agrupados **com intervalo ou em classes**

Nesta situação, primeiro *identificamos a classe modal*, ou seja, aquela que possui maior frequência. Em seguida, *aplicamos a fórmula*. Acompanhe o exemplo da tabela 10:

Tabela 10 Intervalo de classes

Classe	Amostra	f_i
1	0 I--- 10	1
2	10 I--- 20	3
3	20 I--- 30	6
4	30 I--- 40	2
Total	-	12

Fonte: dados hipotéticos

$$Mo = l_{\inf Mo} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$$

Fórmula de Czuber

Limite inferior da classe modal

Diferença entre a f_i da classe modal e a f_i imediatamente anterior

Diferença entre f_i da classe modal e a f_i imediatamente posterior

$$Mo = l_{\inf Mo} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 20 + \frac{3}{3+4} \times 10 = 20 + 0,43 \times 10 = 24,3 \Rightarrow Mo = 24,3$$

Exercícios de Aplicação de Moda

2. Considere a tabela de frequência com os dados agrupados em intervalos de classe como mostrado na tabela abaixo e calcule a moda:

i	X	f_i	F_i
1	150 ⊢ 154	4	4
2	154 ⊢ 158	9	13
3	158 ⊢ 162	11	24
4	162 ⊢ 166	8	32
5	166 ⊢ 170	5	37
6	170 ⊢ 174	3	40

classe Modal

Aplicando a fórmula da Czuber:

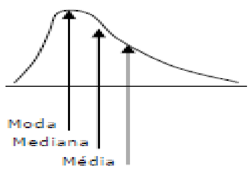
$$Mo = l_i^* + \frac{D_1}{D_1 + D_2} h$$

$D_1 = f^* - f_{anterior}$ e $D_2 = f^* - f_{posterior}$

Temos:

$$M_0 = 158 + \frac{(11-9)}{(11-9)+(11-8)} x4 = 158 + \frac{2}{2+3} x4 = 158 + 1{,}6 = 159{,}6$$

Para esta Tabela de Frequências com dados agrupados com intervalos de classe a média = 161 e mediana = 160,54 (encontrados anteriormente) e agora a moda = 158,60, mostra, claramente que os dados estão distribuídos assimetricamente com distorção (assimetria ou *skewness*) à esquerda.

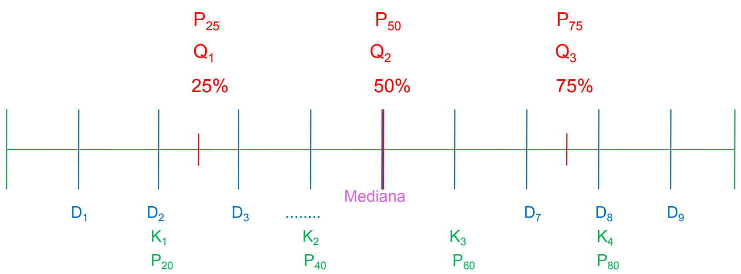


Cap. 04 - MEDIDAS SEPARATRIZES

p. 71

As **medidas separatrizes** representam números reais que separam o rol ou a distribuição de freqüências em partes iguais. Assim, a mediana, que divide a sequência ordenada em duas partes, cada uma delas contendo 50% dos valores da sequência, é também uma *medida separatriz*.

Além da mediana, estudaremos outras medidas separatrizes denominadas *quartis* (Q_i, sendo i = 1,2,3), *quintis* (K_i, sendo i=1,2,3,4) , *decis* (D_i, sendo i = 1,2, ...,9) e ***percentis*** (P_j, sendo j = 10, 20,...,90) . Observe, com atenção, a representação esquemática de uma sequência ordenada e, a seguir, relacione-a com o enunciado denominado para cada divisão desta sequência.



Atenção

Q_2 é a mediana da série

Q_3 (terceiro quartil) – separa a sequência ordenada, deixando 75% de seus elementos à esquerda e 25% de seus elementos à direita.

K_1 (primeiro quintil) – separa a sequência ordenada, deixando 20% de seus elementos à esquerda e 80% de seus elementos à direita.

D_1 (primeiro decil) – separa a sequência ordenada, deixando 10% de seus elementos à esquerda e 90% de seus elementos à direita.

P_1 (primeiro percentil) – separa a sequência, deixando 1% de seus elementos à esquerda e 99% de seus elementos à direita.

$Q_1 = P_{25}$	$K_1 = P_{20}$	$D_1 = P_{10}$
$Q_2 = P_{50}$	$K_2 = P_{40}$	$D_2 = P_{20}$
$Q_3 = P_{75}$	$K_3 = P_{60}$	$D_3 = P_{30}$
	$K_4 = P_{80}$	$D_4 = P_{40}$
		$D_5 = P_{50}$
		$D_6 = P_{60}$
		$D_7 = P_{70}$
		$D_8 = P_{80}$
		$D_9 = P_{90}$

podemos estabelecer a fórmula de cálculo de percentis onde todas as outras medidas podem ser identificadas como percentis, basta observarmos que os quartis, quintis e decis são múltiplos dos percentis. Desta forma, observe as correspondentes colunas que apresentamos a seguir:

Fórmulas

Para **dados agrupados em classes**, encontraremos os **Quartis** de maneira semelhante à usada para o cálculo da mediana:

$$Q_i = l_{\inf_{Q_i}} + \frac{E_{Q_i} - F_{anterior}}{f_{i_{Q_i}}} \times h$$

Q_i : denota o **Quartil** i , sendo i = 1, 2, 3.

$l_{\inf_{Q_i}}$: denota o **limite inferior** da classe que contém o quartil desejado.

E_{Q_i}: denota o elemento quartílico.

$F_{anterior}$: denota a **frequência acumulada até a classe anterior** à classe mediana.

h : denota a **amplitude do intervalo de classe**.

$f_{i_{Q_i}}$: denota a **frequência absoluta simples** da classe quartílica.

Decis - De maneira geral, para encontrar a ordem ou posição do ***Decil*** a ser calculado usaremos a expressão:

$$E_{D_i} = \frac{i \times n}{10}$$

E_{D_i} : denota o **Decil** i , sendo i = 1, 2, 3.

i : denota o **nº de elementos**.

n : denota o elemento quartílico.

Exemplo 1.12

Os dados do quadro 9 informam a *produção de barris de petróleo bruto por mês* de 16 poços de petróleo. Calcule o primeiro quartil.

Quadro 9 Dados da produção de barris de petróleo bruto por mês.

390	420	390	510	420	450	410	510
430	420	400	520	450	420	450	510

Fonte: dados simulados

Ordenando a sequência, em rol crescente, temos: 390, 390, 400, **410**, 420, 420, 420, 420, 430, 450, 450, 450, 510, 510, 510, 520.
Como sabemos, $Q_1 = P_{25}$, assim calculamos 25% de 16 (n = 16) , obtendo:

$$E_{P_i} = \frac{i \times n}{100} = \frac{25 \times 16}{100} = 4 \Rightarrow$$

P_{25} é o 4º elemento do **rol**, ou seja, o valor 4 indica a posição do **vigésimo quinto percentil (P_{25})** que corresponde *ao* ***primeiro quartil*** ($Q_1 = P_{25}$), pedido na situação-problema.

25% dos poços de petróleo produzem valores menores ou iguais a 410 barris de petróleo bruto por mês, e 75% desses poços produzem valores maiores ou iguais a 410.

Atividade 1.5

Fornecida a tabela com a distribuição de frequência do consumo médio de eletricidade entre usuários de uma cidade do interior de São Paulo, encontre o primeiro quartil, o septuagéssimo quinto centil e o nono decil. Interprete os resultados obtidos.

Tabela 11 Distribuição de frequência

Consumo kwh	Número de usuários (f_i)	F
5 ⊢ 25	4	4
25 ⊢ 45	6	10
45 ⊢ 65	14	24
65 ⊢ 85	26	50
85 ⊢ 105	14	64
105 ⊢ 125	8	72
125 ⊢ 145	6	78
145 ⊢ 165	2	80
Total	80	-

Fonte: dados hipotéticos

Antes de calcularmos o Q_1 , o C_{75} e o D_9 vamos observar as informações que a tabela 11 nos fornece. Além da frequência de usuários distribuídos nas 8 classes de valores do consumo médio de eletricidade (kwh) temos também uma outra coluna com a freqüência acumulada ***F***.

Atenção!
Lembre–se sempre das dicas quanto a recordar os conceitos estudados! Então, recordar o conceito de **F** , **a frequência acumulada** - corresponde à soma das frequências absolutas (f_i) de sua classe, mais as anteriores caso existir. E o **F** acumulado representado na última classe, deverá ser igual ao valor de n. Assim, $F = f_i^{\,atual} + F_{anterior}$.

Para calcular o primeiro quartil, primeiro temos que encontrar a sua posição .

$$E_{Q_1} = \frac{i*n}{4} \Rightarrow E_{Q_1} = \frac{1*80}{4} = 20$$

Como podemos ver o Q_1 encontra-se na 20ª posição logo, com o auxílio dos valores dispostos na coluna de frequência acumulada, verificamos que esta posição localiza-se na 3ª classe. Com base nessas informações calculamos o valor de Q_1 .

$$Q_i = l_{\inf_{Q_i}} + \frac{E_{Q_i} - F_{anterior}}{f_{i_{Q_i}}} * h \Rightarrow Q_1 = 45 + \frac{20-10}{14} * 20 = 45 + 14,29 = 59,29 km/h$$

Concluímos que 25% dos usuários consomem até 59,29 kwh . De forma análoga, 75% consomem acima disso.
Da mesma forma, para o cálculo do septuagésimo quinto centil C_{75} , primeiro temos que encontrar a sua posição (E_{C75})

Atividade 1.5 – página 43 continuação

$$E_{C_i} = \frac{i*n}{100} \Rightarrow E_{C_{75}} = \frac{75*80}{100} = \frac{6.000}{100} = 60$$

Como podemos ver o C_{75} encontra-se na 60ª posição logo, com o auxílio dos valores dispostos na coluna de frequência acumulada, verificamos que esta posição localiza-se na 4ª classe. Com base nessas informações calculamos o valor de C_{75} .

$$C_{75} = l_{\inf_{C_i}} + \frac{E_{C_i} - F_{anterior}}{f_{i_{C_i}}} * h \Rightarrow C_{75} = 65 + \frac{60-24}{26} * 20 = 65 + 27,70 = 92,70 km/h$$

Concluímos que 75% dos usuários consomem até 92,70 kwh. De forma análoga, 25% consomem acima disso .

No cálculo do nono decil D_9, primeiro também temos que encontrar a sua posição.

$$E_{D_i} = \frac{i*n}{10} \Rightarrow E_{D_9} = \frac{9*80}{10} = \frac{720}{10} = 72$$

Verificamos que o D_9 encontra-se na 72ª posição logo, com o auxílio dos valores dispostos na coluna de frequência acumulada, verificamos que esta posição localiza-se na 6ª classe. Com base nessas informações calculamos o valor de D_9 .

Concluímos que 90% dos usuários consomem até 125 kwh . De forma análoga, 10% consomem acima disso .

Exercícios Propostos p. 77

1. Em uma série ordenada, qual éo percentual de elementos que ficam à esquerda de cada uma das medidas separatrizes:

a. D_1 b. Q_1 c. K_1 d. D_2 e. K_3 f. Q_3 g. K_4 h. Q_2 i. D_8 j. P_{70}

Solução

D_1 – representa o 1º decil, portanto, temos que 10% dos elementos ficam à sua esquerda.

Q_1 – representa o 1º quartil, portanto, temos que 25% dos elementos ficam à sua esquerda.

K_1 – representa o 1º quintil, portanto, temos que 20% dos elementos ficam à sua esquerda.

D_2 – representa o 2º decil, portanto, temos que 20% dos elementos ficam à sua esquerda.

K_3 – representa o 3º quintil, portanto, temos que 60% dos elementos ficam à sua esquerda.

Q_3 – representa o 3º quartil, portanto, temos que 75% dos elementos ficam à sua esquerda.

K_4 – representa o 4º quintil, portanto, temos que 80% dos elementos ficam à sua esquerda.

Q_2 – representa o 2º quartil, portanto, temos que 50% dos elementos ficam à sua esquerda.

D_8 – representa o 8º decil, portanto, temos que 80% dos elementos ficam à sua esquerda.

P_{70} – representa o 70º percentil, portanto, temos que 70% dos elementos ficam à sua esquerda.

Exercícios Propostos p. 77

2. Em uma série ordenada, qual éo percentual de elementos que ficam à direita de cada uma das medidas separatrizes:
a. D_4 b. P_{80} c. Q_3 d. K_2 e. P_{20} f. D_5 g. Q_1 h. P_2

Solução

D_4 – representa o 4º decil, portanto, temos que 60% dos elementos ficam à sua direita.

P_{80} – representa o 80º percentil, portanto, temos que 20% dos elementos ficam à sua direita.

Q_3 – representa o 3º quartil, portanto, temos que 25% dos elementos ficam à sua direita.

K_2 – representa o 2º quintil, portanto, temos que 60% dos elementos ficam à sua direita.

P_{20} – representa o 20º percentil, portanto, temos que 80% dos elementos ficam à sua direita.

D_5 – representa o 5º decil, portanto, temos que 50% dos elementos ficam à sua direita.

Q_1 – representa o 1º quartil, portanto, temos que 75% dos elementos ficam à sua direita.

P_2 – representa o 2º percentil, portanto, temos que 98% dos elementos ficam à sua direita.

Exercícios Propostos p. 77

3. Qual é o percentual de elementos de uma série ordenada que se situam entre:
a. Q_1 e Q_3 b. P_{10} e P_{90} c. D_2 e D_6 d. Q_1 e K_3
e. D_3 e K_4 f. K_2 e D_8 g. K_3 e Q_3

Solução

Q_1 e Q_3 – representa o intervalo interquartílico de 50% (=75% - 25%).

P_{10} e P_{90} – representa o percentual de 80% dos elementos da série.

D_2 e D_6 – representa o percentual de 40% dos elementos da série.

Q_1 e K_3 – representa o percentual de 35% (=60% - 25%) dos elementos da série.

D_3 e K_4 – representa o percentual de 50% (80% - 30%) dos elementos da série.

K_2 e D_8 – representa o percentual de 40 (80% - 40%) dos elementos da série.

K_3 e Q_3 – representa o percentual de 15% (75% - 60%) dos elementos da série.

Exercícios Propostos p. 77

4. Se uma série ordenada possui 180 elementos, dê o número aproximado de elementos que se situam:

a. acima do P_{20} b. abaixo do K_3 c. acima do Q_3 d. abaixo do P_{90}
e. entre o P_{10} e P_{80} f. entre o Q_1 e Q_3 g. entre Q_3 e P_{80} h. entre P_{90} e P_{92}

Solução

A posição do percentil é dada por E_{Pi} = (i . n)/100

E_{P20} = (20x180)/100 = 36ª posição. Logo, temos 144 elementos acima desta posição.

E_{K3} = (3x180)/5 = 108ª posição. Logo, temos 108 elementos abaixo desta posição.

E_{Q3} = (3x180)/4 = 135ª posição. Logo, temos 45 elementos acima desta posição.

E_{P90} = (90x180)/100 = 162ª posição. Logo, temos 162 elementos acima desta posição.

E_{P10} = (10x180)/100 = 18 e E_{P80} = (80x180)/100 = 144. Logo entre as posições 18ª e 144ª, temos 126 elementos.

E_{Q1} = (1x180)/4 = 45 e E_{Q3} = (3x180)/4 = 135. Logo entre as posições 45ª e 135ª, temos 90 elementos.

E_{Q3} = (3x180)/4 = 135 e E_{P80} = (80x180)/100 = 144. Logo entre as posições 135ª e 144ª, temos 9 elementos.

E_{P90} = (90x180)/100 = 162 e E_{P92} = (92x180)/100 = 165,6. Logo entre as posições 162ª e 166ª (arredondando), temos 4 elementos.

Exercícios Propostos p. 77

5. Dada a série X: 3, 15, 6, 9, 10, 4, 12, 15, 17, 20, 29, calcule:

a. Q_1 b. K_2 c. D_4 d. Q_3 e. P_{90}

Solução

Primeiramente formemos o ROL: 3, 4, 6, 9, 10, 12, 15, 15, 17, 20, 29

E_{Q1} = (1x11)/4 = 2,75. Esta posição não é inteira, significa que o Q_1 é um valor compreendido entre o 2º e o 3º elemento da série. O 2º é o 4 e o 3º é o 6. Tirando a média, temos 5. Assim, temos que 25% dos valores da série são menores ou iguais a 5 e 75% são maiores ou iguais a 5.

E_{K2} = (2x11)/5 = 4,40. Esta posição não é inteira, significa que o K_2 é um valor compreendido entre o 4º e o 5º elemento da série. O 4º é o 9 e o 5º é o 10. Tirando a média, temos 9,5. Assim, temos que 40% dos valores da série são menores ou iguais a 9,5 e 60% são maiores ou iguais a 5.

E_{D4} = (4x11)/10 = 4,40. Esta posição não é inteira, significa que o D_4 é um valor compreendido entre o 4º e o 5º elemento da série. O 4º é o 9 e o 5º é o 10. Tirando a média, temos 9,5. Assim, temos que 40% dos valores da série são menores ou iguais a 9,5 e 60% são maiores ou iguais a 5.

E_{Q3} = (3x11)/4 = 8,25. Esta posição não é inteira, significa que o Q_3 é um valor compreendido entre o 8º e o 9º elemento da série. O 8º é o 15 e o 9º é o 17. Tirando a média, temos 16. Assim, temos que 75% dos valores da série são menores ou iguais a 16 e 25% são maiores ou iguais a 16.

E_{P90} = (90x11)/100 = 9,90. Esta posição não é inteira, significa que o P_{90} é um valor compreendido entre o 9º e o 10º elemento da série. O 9º é o 17 e o 10º é o 20. Tirando a média, temos 18,5. Assim, temos que 90% dos valores da série são menores ou iguais a 18,5 e 10% são maiores ou iguais a 18,5.

Exercícios Propostos p. 77

6. A distribuição de frequência abaixo representa idade de 50 alunos de uma classe de primeiro ano de uma Faculdade:

Idade (anos)	Nº de alunos
17	3
18	18
19	17
20	8
21	4

Calcule:

a. Q_1 b. K_2 c. D_4 d. Q_3 e. P_{90}

Solução

Primeiramente montemos a coluna de frequências acumuladas

Idade (anos)	Nº de alunos	$F_{acumulada}$
17	3	3
18	18	21
19	17	38
20	8	46
21	4	50

E_{Q1} = (1x50)/4 = 12,50. Como não é inteira, significa que o Q_1 é um elemento compreendido entre a 12ª e a 13ª posição. Pela $F_{acumulada}$, o 12ª é o 18 e o 13º é o 18, também. Logo, Q_1 = 18 anos.

E_{K2} = (2x50)/5 = 20. Significa que o K_2 é o elemento na 20ª posição. Pela $F_{acumulada}$, o elemento da 20ª posição é o 18. Logo, K_2 = 18 anos.

E_{D4} = (4x50)/10 = 20. Significa que o K_2 é o elemento na 20ª posição. Pela $F_{acumulada}$, o elemento da 20ª posição é o 18. Logo, K_2 = 18 anos.

E_{Q3} = (3x50)/4 = 37,5. Como não é inteira, significa que o Q_3 é um elemento compreendido entre a 37ª e a 38ª posição. Pela $F_{acumulada}$, o 37ª é o 19 e o 38º é o 19, também. Logo, Q_3 = 19 anos.

E_{P90} = (90x50)/100 = 45. Significa que o P_{90} é o elemento na 45ª posição. Pela $F_{acumulada}$, o elemento da 45ª posição é o 20. Logo, P_{90} = 20 anos.

Exercícios Propostos p. 77

7. A distribuição de frequência abaixo representa o consumo por nota de 54 notas fiscais emitidas durante um dia em uma loja de departamentos.

Classe	Consumo por nota US$	Nº de notas
1	0|-- 50	10
2	50|--100	28
3	100|--150	12
4	150|--200	2
5	200|--250	1
6	250|--300	1

Calcule:

a. Q_1 b. K_2 c. D_3 d. Q_3 e. D_7 f. P_{98}

Solução

Primeiramente montemos a coluna de frequências acumuladas

Classe	Consumo por nota US$	Nº de notas	F_{acum}
1	0|-- 50	10	10
2	50|--100	28	38
3	100|--150	12	50
4	150|--200	2	52
5	200|--250	1	53
6	250|--300	1	54

E_{Q1} = (1x54)/4 = 13,5. Isto nos dá a posição de Q_1 na série. A classe que contém o elemento que ocupa a posição 13,5 é a 2ª classe. Substituindo os valores indicados na fórmula, obtém-se:

$$Q_1 = l_{inf} + \frac{E_{Q1} - F_{ant}}{f_i} h = 50 + \frac{13{,}5 - 10}{28} .50 = 56{,}5$$

E_{K2} = (2x54)/5 = 21,60. Isto nos dá a posição de K_2 na série. A classe que contém o elemento que ocupa a posição 21,60 é a 2ª classe. Substituindo os valores indicados na fórmula, obtém-se:

$$K_2 = l_{inf} + \frac{E_{K2} - F_{ant}}{f_i} h = 50 + \frac{21{,}60 - 10}{28} .50 = 70{,}71$$

E_{D3} = (3x54)/10 = 16,2. Isto nos dá a posição de D_3 na série. A classe que contém o elemento que ocupa a posição 16,2 é a 2ª classe. Substituindo os valores indicados na fórmula, obtém-se:

$$D_3 = l_{inf} + \frac{E_{D3} - F_{ant}}{f_i} h = 50 + \frac{16{,}2 - 10}{28} .50 = 61{,}07$$

Respostas:
d. 110,42 e. 99,64 f. 246

Exercícios Propostos p. 77

8. Interprete os valores obtidos no problema anterior.

Solução

a. 25% das notas indicavam consumo menor ou igual a US$ 56,25 e 75% indicavam consumo maior ou igual a US$ 56,25

b. 40% das notas indicavam consumo menor ou igual a US$ 70,71 e 60% indicavam consumo maior ou igual a US$ 70,71

c. 30% das notas indicavam consumo menor ou igual a US$ 61,07 e 70% indicavam consumo maior ou igual a US$ 61,07

d. 75% das notas indicavam consumo menor ou igual a US$ 110,42 e 25% indicavam consumo maior ou igual a US$ 110,42

e. 70% das notas indicavam consumo menor ou igual a US$ 99,64 e 75% indicavam consumo maior ou igual a US$ 99,64

f. 98% das notas indicavam consumo menor ou igual a US$ 246,00 e 2% indicavam consumo maior ou igual a US$ 246,00

Exercícios Propostos p. 77

9. Tomando como amostra a tabela do problema 7 apresentada novamente abaixo, o gerente desta loja de departamentos decidiu premiar a nível promocional com um brinde, 10% dos fregueses que mais consumirem, nos próximos 30 dias. A partir de qual valor de consumo da nota fiscal os clientes seriam premiados?

Classe	Consumo por nota US$	Nº de notas
1	0|-- 50	10
2	50|--100	28
3	100|--150	12
4	150|--200	2
5	200|--250	1
6	250|--300	1

Solução

Os 10% dos clientes que mais consumirem serão aqueles cujo consumo seja superior a D_9.

E_{D9} = (9x54)/10 = 48,60. Isto nos dá a posição de D_9 na série. A classe que contém o elemento que ocupa a posição 48,60 é a 3ª classe. Substituindo os valores indicados na fórmula, obtém-se:

Classe	Consumo por nota US$	Nº de notas	F_{acum}
1	0|-- 50	10	10
2	50|--100	28	38
3	100|--150	12	50
4	150|--200	2	52
5	200|--250	1	53
6	250|--300	1	54

$$D_9 = l_{inf} + \frac{E_{D9} - F_{ant}}{f_i} h = 100 + \frac{48,60 - 38}{12} . 50 = 144,17$$

A partir do valor de consumo de R$ 144,17, os clientes seriam premiados!!!

Exercícios Propostos p. 77

10. Uma empresa estabelece o salário de seus vendedores com base na produtividade. Desta forma, 10% é fixo e 90% são comissões sobre a venda. Uma amostra de salários mensais nesta empresa revelou o quadro abaixo. Se a empresa decidir, a nível de incentivo, fornecer uma cesta básica para 5% dos vendedores que pior desempenho tiveram durante o próximo mês com base nesta amostra, qual será o maior salário que receberá esta cesta básica

Classe	Salários US$	Nº de vendedores
1	70|-- 120	8
2	120|--170	28
3	170|--220	54
4	220|--270	32
5	270|--320	12
6	320|--370	6

Solução

Primeiramente montemos a coluna de frequências acumuladas

Classe	Salários US$	Nº de vend.	F_{acum}
1	70|-- 120	8	8
2	120|--170	28	36
3	170|--220	54	90
4	220|--270	32	122
5	270|--320	12	134
6	320|--370	6	140

Estamos querendo premiar os 5% piores vendedores, isto é, queremos o P_5 para premiarmos os salários abaixo daquele valor

E_{P5} = (5x140)/100 = 7,00. Isto nos dá a posição de P_5 na série. A classe que contém o elemento que ocupa a posição 7,00 é a 1ª classe. Substituindo os valores indicados na fórmula, obtém-se:

$$P_5 = l_{inf} + \frac{E_{P5} - F_{ant}}{f_i} h = 70 + \frac{7{,}00 - 0}{8} .50 = 113{,}75$$

A empresa premiará com incentivo da cesta básica todos aqueles funcionários que conseguirem um salário abaixo a R$ 113,75.

Exercícios Propostos p. 77

11. A tabela abaixo representa a venda de livros didáticos em uma editora na primeira semana de março.

Determine: a. Q_1 b. Q_3 c. P_{90} d. P_{10}

Classe	Preço Unitário US$	Nº de livros vendidos
1	0|-- 10	4.000
2	10|--20	13.500
3	20|--30	25.600
4	30|--40	43.240
5	40|--50	26.800
6	50|--60	1.750

Solução

Primeiramente montemos a coluna de frequências acumuladas

Classe	Preço Unitário US$	Nº de livros vendidos	F_{acum}
1	0|-- 10	4.000	4.000
2	10|--20	13.500	17.500
3	20|--30	25.600	43.100
4	30|--40	43.240	86.340
5	40|--50	26.800	113.140
6	50|--60	1.750	114.890

E_{Q1} = (1x114890)/4 = 28.722,5. Isto nos dá a posição de Q_1 na 3ª classe da série. Substituindo os valores indicados na fórmula, obtém-se:

$$Q_1 = l_{inf} + \frac{E_{Q1} - F_{ant}}{f_i} h = 20 + \frac{28722{,}5 - 17500}{25600} .10 = 24{,}38$$

E_{Q3} = (3x114890)/4 = 86.167,5. Isto nos dá a posição de Q_3 na 4ª classe da série. Substituindo os valores indicados na fórmula, obtém-se:

$$Q_3 = l_{inf} + \frac{E_{Q3} - F_{ant}}{f_i} h = 30 + \frac{86167{,}5 - 43100}{43240} .10 = 39{,}96$$

E_{P90} = (90x114890)/100 = 103.401. Isto nos dá a posição de P_{90} na 5ª classe da série. Substituindo os valores indicados na fórmula, obtém-se:

$$P_{90} = l_{inf} + \frac{E_{P90} - F_{ant}}{f_i} h = 40 + \frac{103401 - 86340}{26800} .10 = 46{,}37$$

E_{P10} = (10x114890)/100 = 11.489. Isto nos dá a posição de P_{10} na 2ª classe da série. Substituindo os valores indicados na fórmula, obtém-se:

$$P_{10} = l_{inf} + \frac{E_{P10} - F_{ant}}{f_i} h = 10 + \frac{11489 - 4000}{13500} .10 = 15{,}55$$

Exercícios Propostos p. 77

12. Interprete os valores obtidos no problema anterior.

Solução

a. 25% dos livros vendidos tinham preço unitário menor ou igual a US$ 24,38 e 75% dos livros vendidos tinham preço unitário maior ou igual a US$ 24,38

b. 75% dos livros vendidos tinham preço unitário menor ou igual a US$ 39,96 e 25% dos livros vendidos tinham preço unitário maior ou igual a US$ 39,96

c. 90% dos livros vendidos tinham preço unitário menor ou igual a US$ 46,37 e 10% dos livros vendidos tinham preço unitário maior ou igual a US$ 46,37

d. 10% dos livros vendidos tinham preço unitário menor ou igual a US$ 15,55 e 90% dos livros vendidos tinham preço unitário maior ou igual a US$ 15,55

Exercícios Propostos p. 77

13. A tabela a seguir representa o número de faltas anuais dos funcionários de uma empresa.
Determine: a. D_3 b. K_3 c. P_{90} d. Q_3 e. Q_1 f. P_{10}

Nº de faltas	Nº de empregados
0	20
1	42
2	53
3	125
4	84
5	40
6	14
7	3
8	2

Solução

Primeiramente montemos a coluna de frequências acumuladas

Nº de faltas	Nº de empregados	F_{acum}
0	20	20
1	42	62
2	53	115
3	125	240
4	84	324
5	40	364
6	14	378
7	3	381
8	2	383

a. E_{D3} = (3x383)/10 = 114,9. Como é uma posição NÃO inteira, o valor D_3 está compreendido entre as posições 114ª (53) e 115ª (53). A média é 53 (2 faltas). Logo 30% dos empregados tiveram 2 faltas ou menos e 70% dos empregados 2 faltas ou mais.

b. E_{K3} = (3x383)/5 = 229,8. Como é uma posição NÃO inteira, o valor K_3 está compreendido entre as posições 229ª (125) e 230ª (125). A média é 125 (3 faltas). Logo 60% dos empregados tiveram 3 faltas ou menos e 40% dos empregados tiveram 3 faltas ou mais.

c. E_{P90} = (90x383)/100 = 344,7. Como é uma posição NÃO inteira, o valor P_{90} está compreendido entre as posições 344ª (40) e 345ª (40). A média é 40 (5 faltas). Logo 90% dos empregados tiveram 5 faltas ou menos e 10% dos empregados tiveram 5 faltas ou mais.

d. E_{Q3} = (3x383)/4 = 287,25. Como é uma posição NÃO inteira, o valor Q_3 está compreendido entre as posições 287ª (84) e 288ª (84). A média é 84 (4 faltas). Logo 75% dos empregados tiveram 4 faltas ou menos e 25% dos empregados tiveram 4 faltas ou mais.

Exercícios Propostos p. 77

13. A tabela a seguir representa o número de faltas anuais dos funcionários de uma empresa.
Determine: a. D_3 b. K_3 c. P_{90} d. Q_3 e. Q_1 f. P_{10}

Nº de faltas	Nº de empregados
0	20
1	42
2	53
3	125
4	84
5	40
6	14
7	3
8	2

Solução

Nº de faltas	Nº de empregados	F_{acum}
0	20	20
1	42	62
2	53	115
3	125	240
4	84	324
5	40	364
6	14	378
7	3	381
8	2	383

e. E_{Q1} = (1x383)/4 = 95,75. Como é uma posição NÃO inteira, o valor Q_1 está compreendido entre as posições 95ª (53) e 96ª (53). A média é 53 (2 faltas). Logo 25% dos empregados tiveram 2 faltas ou menos e 75% dos empregados tiveram 2 faltas ou mais.

f. E_{P10} = (10x383)/100 = 38,3. Como é uma posição NÃO inteira, o valor P_{10} está compreendido entre as posições 38ª (42) e 39ª (42). A média é 42 (1 faltas). Logo 10% dos empregados tiveram 1 falta ou menos e 90% dos empregados tiveram 1 falta ou mais.

Exercícios Propostos p. 77

14. Interprete os valores obtidos no problema anterior.

Solução

a. 30% dos empregados tiveram um número de faltas anuais menor ou igual a 2 e 70% dos empregados tiveram um número de faltas anuais maior ou igual a 2

b. 60% dos empregados tiveram um número de faltas anuais menor ou igual a 3 e 40% dos empregados tiveram um número de faltas anuais maior ou igual a 3

c. 90% dos empregados tiveram um número de faltas anuais menor ou igual a 5 e 10% dos empregados tiveram um número de faltas anuais maior ou igual a 2

d. 75% dos empregados tiveram um número de faltas anuais menor ou igual a 4 e 25% dos empregados tiveram um número de faltas anuais maior ou igual a 4

e. 25% dos empregados tiveram um número de faltas anuais menor ou igual a 2 e 75% dos empregados tiveram um número de faltas anuais maior ou igual a 2

f. 10% dos empregados tiveram um número de faltas anuais menor ou igual a 1 e 90% dos empregados tiveram um número de faltas anuais maior ou igual a 1

Exercícios Propostos p. 77

15. Uma amostra do tempo de vida útil de uma peça forneceu a distribuição abaixo:

Classe	Nº de horas (vida útil)	Nº de peças
1	0|-- 100	6
2	100|--200	42
3	200|--300	86
4	300|--400	127
5	400|--500	64
6	500|--600	8

Se o produtor deseja estaberlecer uma garantia mínima para o número de horas de vida útil de uma peça, trocando a peça que não apresentar este número mínimo de horas, qual é a garantia, se ele está disposto a trocar 8% das peças.

Solução

Primeiramente montemos a coluna de frequências acumuladas. Queremos a vida útil inferior a 8%, ou seja, inferior ao P_8

Classe	Nº de horas (vida útil)	Nº de peças	F_{acum}
1	0|-- 100	6	6
2	100|--200	42	48
3	200|--300	86	134
4	300|--400	127	261
5	400|--500	64	325
6	500|--600	8	333

E_{P8} = (8x333)/100 = 26,64. Isto nos dá a posição de P_8 na 2ª classe da série. Substituindo os valores indicados na fórmula, obtém-se:

$$P_8 = l_{inf} + \frac{E_{P8} - F_{ant}}{f_i} h = 100 + \frac{26{,}64 - 6}{42} \cdot 100 = 149{,}14 \text{ horas}$$

CONCLUSÃO: Para peças com vida útil inferior ou igual a 149,14 h, o produtor TROCARÁ a peça. Para peças com vida útil superior ou igual a 149,14 h, o produtor NÃO TROCARÁ a peça.

Colocar os Exercícios Propostos da p. 10 da Apostila