1.8 MEDIDAS SEPARATRIZES

São valores que separam o rol (os dados ordenados) em quatro (quartis), dez (decis) ou em cem (percentis) partes iguais. Note que para a sua correta aplicação, exige-se que os dados estejam organizados num **rol.**

QUARTIS (Q_i)

DECIS (D_i)

PERCENTIS OU CENTIS (C_i)

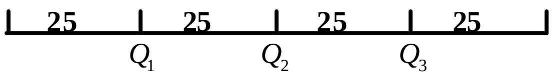
1.8.1 Quartis (Q_i)

São valores que dividem o conjunto de dados ordenados (rol) em 4(quatro) partes iguais.

Primeiro Quartil (Q_1) - valor situado de tal modo na série de dados que 25% das observações são menores que ele e 75% são maiores.

Segundo Quartil (Q_2) - valor situado de tal modo na série de dados que 50% das observações são menores que ele e 50% são maiores.

Terceiro Quartil (Q_3) - valor situado de tal modo na série de dados que 75% das observações são menores que ele e 25% são maiores.



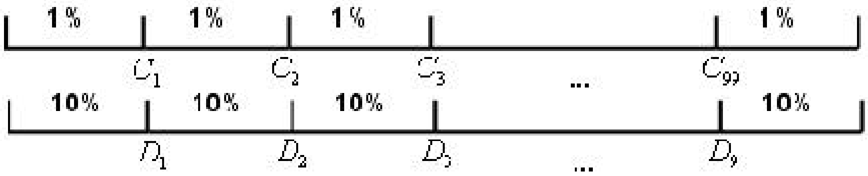
1.8.2 Decis (D_i)

São valores que dividem o conjunto de dados ordenados (rol) em 10(dez) partes iguais.

Primeiro Decil (D_1) - valor situado de tal modo na série de dados que 10% das observações são menores que ele e 90% são maiores.

Segundo Decil (D_2) - valor situado de tal modo na série de dados que 20% das observações são menores que ele e 80% são maiores.

Nono Decil (D_9) - valor situado de tal modo na série de dados que 90% das observações são menores que ele e 10% são maiores.



1.8.3 Percentis ou Centis (C_i)

São valores que dividem o conjunto de dados ordenados (rol) em 100(cem) partes iguais.

Primeiro Percentil (C_1) - valor situado de tal modo na série de dados que 1% das observações são menores que ele e 99% são maiores.

Segundo Percentil (C_2) - valor situado de tal modo na série de dados que 2% das observações são menores que ele e 98% são maiores.

Segundo Percentil (C_3) - valor situado de tal modo na série de dados que 2% das observações são menores que ele e 98% são maiores.

. . . .

Nonagésimo Nono Percentil (C_{99}) - valor situado de tal modo na série de dados que 99% das observações são menores que ele e 1% são maiores.

Cálculo dos Quartis, Decis e Percentis

Roteiro para o cálculo:

1º (Passo) Determinar as freqüências acumuladas (fac) da distribuição.

2º (Passo) Calcular a posição do Quartil, Decil ou Percentil desejado, por uma das fórmulas.

$$P_i = \left(\frac{i.\sum f_i}{4} \right) \text{(i= 1,2,3)} \qquad \rightarrow \textbf{Quartil}$$

$$P_i = \left(\frac{i.\sum f_i}{10} \right) \text{(i = 1, 2, 3,..., 9)} \qquad \rightarrow \textbf{Decil}$$

$$P_i = \left(\frac{i.\sum f_i}{100} \right) \text{(i = 1, 2, 3,..., 99)} \qquad \rightarrow \textbf{Percentil}$$

3º (Passo) Identificar a que classe que contém o Quartil, Decil ou Percentil desejado por meio da freqüência acumulada simples (f_{ac}).

4º (Passo) Calcular o Quartil, Decil ou Percentil desejado por meio de uma das fórmulas:

- **Para o Quartil:**

$$Q_i = l_i + \left(\frac{P - f_{ant}}{f_Q} \right).h \quad \text{onde: } \begin{cases} l_i \rightarrow \text{limite inferior da classe do quartil} \\ f_{ant} \rightarrow \text{frequência acumulada anterior a classe do quartil} \\ f_Q \rightarrow \text{frequência simples da classe do quartil} \\ h \rightarrow \text{amplitude de classe do quartil} \end{cases}$$

- **Para o Decil:**

$$D_i = l_i + \left(\frac{P - f_{ant}}{f_D} \right).h \quad \text{onde: } \begin{cases} l_i \rightarrow \text{limite inferior da classe do decil} \\ f_{ant} \rightarrow \text{frequência acumulada anterior a classe do decil} \\ f_D \rightarrow \text{frequência simples da classe do decil} \\ h \rightarrow \text{amplitude de classe do decil} \end{cases}$$

- **Para o Percentil:**

$$C_i = l_i + \left(\frac{P - f_{ant}}{f_C} \right).h \quad \text{onde: } \begin{cases} l_i \rightarrow \text{limite inferior da classe do percentil} \\ f_{ant} \rightarrow \text{frequência acumulada anterior a classe do percentil} \\ f_C \rightarrow \text{frequência simples da classe do percentil} \\ h \rightarrow \text{amplitude de classe do percentil} \end{cases}$$

Exemplo: os salários (em salário mínimo) de 160 professores de uma escola estão distribuídos conforme a tabela a seguir. Calcule o Q_1, D_4 e o C_{85} e interprete os resultados.

Salário	***N. º de prof. (fi)***	***fac***
01 |--- 03	20	20
03 |--- 05	40	60
05 |--- 07	60	120
07 |--- 09	30	150
09 |--- 11	10	160
Total	160	-----

Solução:

1º (Passo) Determinar as freqüências acumuladas (fac) da distribuição.

2º (Passo) Calcular a posição do Quartil, Decil ou Percentil desejado, por uma das fórmulas.

$$P_1 = \left(\frac{i.\sum f_i}{4}\right) = \frac{1.160}{4} = 40^0\ elemento.$$ → **Quartil**

$$P_4 = \left(\frac{i.\sum f_i}{10}\right) = \frac{4.160}{10} = 64^0 elemento.$$ → **Decil**

$$P_{85} = \left(\frac{i.\sum f_i}{100}\right) = \frac{85.160}{100} = 136^0 elemento$$ → **Percentil**

3º (Passo) Identificar a que classe que contém o Quartil, Decil ou Percentil desejado por meio da freqüência acumulada simples (f_{ac}). Quartil (segunda classe); Decil (terceira Classe); Percentil (Quarta classe).

4º (Passo) Calcular o Quartil, Decil ou Percentil desejado por meio de uma das fórmulas:

- **Quartil:**

$$Q_1 = 3 + \left(\frac{40-20}{40}\right).2 = 4$$ Salários mínimos

- **Decil:**

$$D_4 = 5 + \left(\frac{64-60}{60}\right).2 = 5{,}13$$ Salários mínimos

- **Percentil:**

$$C_{85} = 7 + \left(\frac{136-120}{30}\right).2 = 8{,}07$$ Salários mínimos

- **Interpretação:** 25% dos professores da escola ganham até 4 salários mínimos ou 75% dos professores ganham mais de 4 salários mínimos.

- **Interpretação:** 40% dos professores da escola ganham até 5,13 salários mínimos ou 60% dos professores ganham mais de 5,13 salários mínimos.

- **Interpretação:** 85% dos professores da escola ganham até 8,07 salários mínimos ou 15% dos professores ganham mais de 8,07 salários mínimos.

1.9 MEDIDAS DE DISPERSÃO

São medidas estatísticas utilizadas para avaliar o grau de variabilidade ou dispersão dos valores em torno de um valor central; geralmente as médias. Servem para medir a representatividade das medidas de tendência central.

Chamamos de dispersão ou variabilidade a maior ou menor diversificação dos valores de uma variável em torno de um valor de tendência central, tomado como ponto de comparação.

Consideremos os seguintes conjuntos de valores como sendo as notas de três turmas de alunos X, Y e Z.

Turma "X" = { 6,0; 6,0; 6,0; 6,0; 6,0} $\overline{X} = 6{,}0$

Turma "Y" = { 5,8; 5,9; 6,0; 6,1; 6,2}

Turma "Z" = { 1,0; 4,0; 6,0; 9,0; 10,0}

Embora as tumas X, Y e Z, apresentem a mesma média aritmética, é fácil notar que o grupo X é mais homogêneo em relação as notas, que os grupos Y e Z, já que todas as notas são iguais a média.

O grupo Y, por sua vez, é mais homogêneo que o grupo Z, pois há menor diversificação entre cada um dos seus valores.

Podemos dizer que a grupo X apresenta dispersão ou variabilidade nula e que o grupo Y apresenta uma dispersão ou variabilidade maior que o grupo Z.

Portanto, para qualificar os valores de uma dada variável, resultando a maior ou menor dispersão ou variabilidade entre esses valores e a sua medida de posição, a Estatística recorre às medidas de dispersão ou de variabilidade. A seguir vamos apresentar algumas medidas de dispersão mais usuais.

1.9.1 Variância e Desvio Padrão

Cálculo da Variância para dados não tabulados

$$S^2 = \frac{1}{n-1} \cdot \left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i \right)^2}{n} \right]$$ Variância

$$S = \sqrt{S^2}$$ Desvio padrão

Consideremos os seguintes conjuntos de valores como sendo amostras de rendimentos de três grupos de trabalhadores.

Grupo "X" = { 6,0; 6,0; 6,0; 6,0; 6,0}

Grupo "Y" = { 5,8; 5,9; 6,0; 6,1; 6,2}

Grupo "Z" = {1,0; 4,0; 6,0; 9,0; 10,0}

Exemplo: determine a variância e o desvio padrão para as amostras dos grupos relacionadas anteriormente.

- Para o **Grupo "X"** = {6,0; 6,0; 6,0; 6,0; 6,0}

x_i *(rendimentos)*	6,0	6,0	6,0	6,0	6,0	$\sum_{i=1}^{5} X_i = 30,0$
x_i^2	36,0	36,0	36,0	36,0	36,0	$\sum_{i=1}^{5} X_i^2 = 180,0$

$$S^2 = \frac{1}{n-1} \cdot \left[\sum_{i=1}^{5} x_i^2 - \frac{\left(\sum_{i=1}^{5} x_i \right)^2}{n} \right] = \frac{1}{5-1} \cdot \left[180 - \frac{(30)^2}{5} \right] = \frac{1}{4} \cdot [180 - 180] = \frac{1}{4} \cdot 0 = 0$$

$$S = \sqrt{S^2} = \sqrt{0} = 0$$

- Para a **Grupo "Y"** = { 5,8; 5,9; 6,0; 6,1; 6,2}

x_i *(rendimentos)*	5,8	5,9	6,0	6,1	6,2	$\sum_{i=1}^{5} X_i = 30,0$
x_i^2	33,64	34,81	36,00	37,21	38,44	$\sum_{i=1}^{5} X_i^2 = 180,10$

$$S^2 = \frac{1}{n-1}.\left[\sum_{i=1}^{5} x_i^2 - \frac{\left(\sum_{i=1}^{5} x_i\right)^2}{n}\right] = \frac{1}{5-1}.\left[180,10 - \frac{(30)^2}{5}\right] = \frac{1}{4}.[180,10 - 180] = \frac{1}{4}.0,10 = \frac{0,10}{4} = 0,02$$

$$S = \sqrt{S^2} = \sqrt{0,02} = 0,1$$

- **Grupo "Z"** = {1,0; 4,0; 6,0; 9,0; 10,0}

x_i *(rendimentos)*	1,0	4,0	6,0	9,0	10,0	$\sum_{i=1}^{5} X_i = 30,0$
x_i^2	1,0	16,0	36,0	81,0	100,0	$\sum_{i=1}^{5} X_i^2 = 234,0$

$$S^2 = \frac{1}{n-1}.\left[\sum_{i=1}^{5} x_i^2 - \frac{\left(\sum_{i=1}^{5} x_i\right)^2}{n}\right] = \frac{1}{5-1}.\left[234 - \frac{(30)^2}{5}\right] = \frac{1}{4}.[234 - 180] = \frac{1}{4}.54 = \frac{54}{4} = 13,5$$

$$S = \sqrt{S^2} = \sqrt{13,5} = 3,7$$

Resumo

Grupo	*Variância*	*Desvio padrão*
X	0,0	0,0
Y	0,2	0,1
Z	13,5	3,7

Analisando os resultados obtidos com o cálculo das variâncias e dos desvios padrões, observa-se que o grupo X apresenta menor dispersão dos valores em torno da média e o grupo Z foi o que apresentou maior variabilidade em torno da média

Cálculo da Variância e do Desvio padrão para dados tabulados

$$S^2 = \frac{1}{\sum_{i=1}^{k} f_i - 1} \cdot \left[\sum_{i=1}^{k} f_i x_i^2 - \frac{\left(\sum_{i=1}^{k} f_i x_i \right)^2}{\sum_{i=1}^{k} f_i} \right] \text{ Variância}$$

$$S = \sqrt{S^2} \text{ Desvio padrão}$$

Exemplo: os salários (em salário mínimo) de 160 funcionários de uma empresa estão distribuídos conforme a tabela a seguir. Calcule a variância e o desvio padrão dos salários dos funcionários.

Salários	***N.º de func. (fi)***	***X_i***	***Xi.fi***	***$Xi^2 \cdot f_i$***
01 |--- 03	20	2	40	80
03 |--- 05	40	4	160	640
05 |--- 07	60	6	360	2160
07 |--- 09	30	8	240	1920
09 |--- 11	10	10	100	1000
Total	160	----	900	5800

$$S^2 = \frac{1}{\sum_{i=1}^{k} f_i - 1} \cdot \left[\sum_{i=1}^{k} f_i x_i^2 - \frac{\left(\sum_{i=1}^{k} f_i x_i \right)^2}{\sum_{i=1}^{k} f_i} \right] = \frac{1}{160-1} \left[5800 - \frac{(900)^2}{160} \right] =$$

$$S^2 = \frac{1}{159} \cdot [5800 - 5062{,}50] = \frac{1}{159} \cdot 737{,}50 = \frac{737{,}50}{159} = 4{,}64 \text{Salarios minimos}^2$$

$$S = \sqrt{S^2} = \sqrt{4{,}64} = 2{,}15 \text{ Salários mínimos}$$

1.9.2 Coeficiente de Variação de Pearson $\left(CV_p\right)$

O coeficiente de variação de Pearson mede percentualmente a variação ocorrida da medida de dispersão absoluta (S) relativo a média aritmética (x), indica a magnitude relativa do desvio padrão quando comparado com a média da distribuição das medidas.

$$CV_p = \frac{S}{\overline{X}} x100\%$$

Exemplo: determine o coeficiente de variação para as amostras dos grupos relacionadas anteriormente.

- Para o **Grupo "X"**:

$$\overline{X} = \frac{\sum X_i}{n} = \frac{30}{5} = 6pts$$

$$CV_p = \frac{S}{\overline{X}} x100 = \frac{0}{6} x100 = \frac{0}{6} = 0{,}0\%$$

- Para o **Grupo "Y"**:

$$\overline{X} = \frac{\sum X_i}{n} = \frac{30}{5} = 6pts$$

$$CV_p = \frac{S}{\overline{X}} x100 = \frac{0{,}1}{6} x100 = \frac{10}{6} = 1.7\%$$

- Para o **Grupo "Z"**:

$$\overline{X} = \frac{\sum X_i}{n} = \frac{30}{5} = 6pts$$

$$CV_p = \frac{S}{\overline{X}} x100 = \frac{3{,}6}{6} x100 = \frac{360}{6} = 60{,}0\%$$

- Observações: alguns analistas sugerem a seguinte classificação do coeficiente de variação.

  - **Baixa variabilidade:** $CV_p < 15\%$
  - **Média variabilidade:** $15\% \leq CV_p < 30\%$
  - **Alta variabilidade:** $CV_p \geq 30\%$

Observando o coeficiente de variação das amostras, percebemos que os grupos "X" e "Y" apresentaram uma baixa variabilidade, enquanto que a grupo "Z" apresentou uma alta variabilidade, portanto, o grupo "Z" é o grupo mais heterogêneo.

1.10 MEDIDAS DE NORMALIDADE

1.10.1 Assimetria (A_s)

Para conceituar assimetria, obviamente precisamos conceituar simetria. Diremos que existe simetria quando a maioria dos valores da variável se concentra no meio da distribuição de forma simétrica.

Exemplo 1: salários (em salário mínimo) dos funcionários da empresa "Alfa" - 2004.

Salários	***N.º de funcionários***
01 |--- 03	10
03 |--- 05	30
05 |--- 07	50
07 |--- 09	30
09 |--- 11	10
Total	130

Quando os valores da variável se concentrarem em uma das extremidades da distribuição, diremos que existe assimetria.

Exemplo 2: salários (em salário mínimo) dos funcionários da empresa “Alfa” - 2004.

Salários	***N.º de funcionários***
01 |--- 03	50
03 |--- 05	40
05 |--- 07	30
07 |--- 09	20
09 |--- 11	10
Total	150

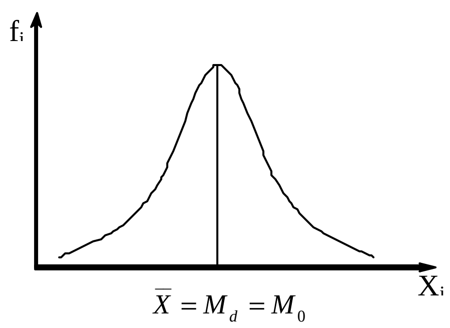
Exemplo 3: salários (em salário mínimo) dos funcionários da empresa “Alfa” - 2004.

Salários	***N.º de funcionários***
01 |--- 03	10
03 |--- 05	20
05 |--- 07	30
07 |--- 09	40
09 |--- 11	50
Total	150

1.10.2 Tipos de curva ou distribuição de freqüência

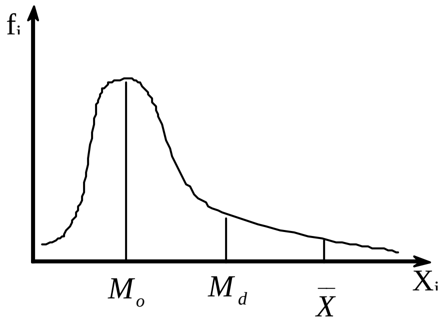
(a) Curva ou Distribuição Simétrica

Uma distribuição é considerada Simétrica quando o valor da média for igual ao da mediana e moda ($\bar{X} = M_d = M_0$), isto significa que a maioria dos valores se concentra no meio da distribuição.



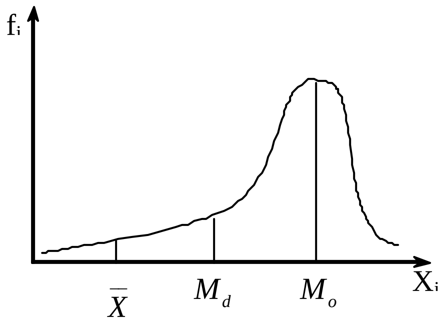
(a) Curva ou Distribuição Assimétrica Positiva:

Uma distribuição é considerada assimétrica positiva quando a o valor da moda(Mo) for menor que o da mediana(Md) e o da mediana menor que o da média(X) ($M_0 < M_d < \overline{X}$), isto significa que a maioria dos valores se concentram á esquerda.



(a) Curva ou Distribuição Assimétrica Negativa:

Uma distribuição é considerada assimétrica negativa quando a o valor da média(X) for menor que o da mediana (Md) e o da mediana menor que o da moda(Mo) ou ($\overline{X} < M_d < M_0$), isto significa que a maioria dos valores se concentram à direita.



1.10.3 Coeficiente de Assimetria

$$As = \frac{\overline{X} - Mo}{S}$$

· **Classificação da distribuição por meio do coeficiente da assimetria:**

Se **As £ -1,00** → então a distribuição é assimétrica negativa forte.

Se **–1,00 < As < -0,15** → então a distribuição é assimétrica negativa fraca.

Se **–0,15 £ As £ 0,15** → então a distribuição é simétrica.

Se **0,15 < As < 1,00** → então a distribuição é assimétrica positiva fraca.

Se **As ³ 1** → então a distribuição é assimétrica positiva forte.

Exemplo: um estudo sobre as distribuição dos pesos dos alunos da escola “FKS”, onde já calculamos os valores de $\overline{X}$ (59,3 kg), M_o (56,8 kg) e S (9,0 kg). Calcule o coeficiente de assimetria da distribuição e classifique a distribuição.

$$As = \frac{\overline{X} - Mo}{S} = \frac{59{,}3 - 56{,}8}{9} = \frac{2{,}5}{9} = 0{,}28$$

Portanto, a distribuição apresenta uma assimetria positiva moderada.