REVISTA DE CIÊNCIA ELEMENTAR

Desvio padrão amostral

Maria Eugénia Graça Martins
Faculdade de Ciências da Universidade de Lisboa
memartins@fc.ul.pt

CITAÇÃO
Martins, E.G.M. (2013)
Desvio padrão amostral,
Rev. Ciência Elem., V1(01):022.
doi.org/10.24927/rce2013.022

EDITOR
José Ferreira Gomes,
Universidade do Porto

RECEBIDO EM
27 de fevereiro de 2012

ACEITE EM
20 de dezembro de 2012

PUBLICADO EM
21 de dezembro de 2012

rce.casadasciencias.org

Desvio padrão de uma amostra (ou coleção) de dados, de tipo quantitativo, é uma medida de dispersão dos dados relativamente à média, que se obtém tomando a raiz quadrada da variância amostral.

Uma vez que a variância amostral se exprime nas unidades dos dados elevados ao quadrado, considera-se como medida de dispersão, não a variância, mas a sua raiz quadrada. Se representarmos os dados por x_1, x_2, ..., xn, e por $\bar{x}$ a sua média, o desvio padrão obtém-se a partir da expressão

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}.$$

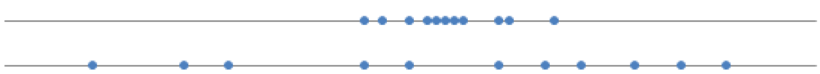
O desvio padrão é uma medida que só pode assumir valores não negativos e quanto maior for o seu valor, maior será a dispersão dos dados.

Por exemplo, os dois conjuntos de dados, que têm a mesma média (igual a 4,9),

$$4 \quad 4,2 \quad 4,5 \quad 4,7 \quad 4,8 \quad 4,9 \quad 5 \quad 5,1 \quad 5,5 \quad 5,6 \quad 6,1$$

$$1 \quad 2 \quad 2,5 \quad 4 \quad 4,5 \quad 5,5 \quad 6 \quad 6,4 \quad 7 \quad 7,5 \quad 8$$

têm desvio padrão, respetivamente 0,6 e 2,3.



Como se verifica, tanto visualmente como a partir dos valores obtidos para o desvio padrão, a dispersão do segundo conjunto de dados é muito superior à do primeiro conjunto.

Além da expressão anterior, por vezes também se utiliza a expressão

$$s' = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

quando a dimensão da amostra n é suficientemente grande (é usual considerar um valor de n superior a 30). Repare-se que nestas condições os valores de s' são muito próximos de s, pois $s'/s = \sqrt{(n-1)/n} \approx 1$.

Costuma-se utilizar o desvio padrão amostral, s, para estimar o desvio padrão populacional, σ.