Parte I – Estatística descritiva

5. Medidas de dispersão

Ana Maria Lopez Calvo de Feijoo

FEIJOO, AMLC. Medidas de dispersão. In: *A pesquisa e a estatística na psicologia e na educação* [online]. Rio de Janeiro: Centro Edelstein de Pesquisas Sociais, 2010, pp. 23-27. ISBN: 978-85-7982-048-9. Available from SciELO Books <http://books.scielo.org>.

---

5. MEDIDAS DE DISPERSÃO

a) *Definição: é* um índice que indica o grau de dispersão dos escores em tomo da posição central.

b) *Objetivo:* descreve a heterogeneidade do grupo.

c) *Utilidade: é* o complemento da medida de tendência central. Mediante seu uso sabe-se que se os valores estão "muito concentrados" ao redor da média aritmética, esta será muito representativa; mas se os valores estão muito dispersos ao redor da média aritmética, esta será pouco representativa.

d) *Medida mais usada*: desvio-padrão.

e) *Outras medidas:* amplitude total, desvio semi-interquartil, desvio médio.

Intervalo total ou amplitude total

A amplitude (ou intervalo total) de uma série é definida como a diferença entre o valor mais alto e o valor mais baixo da série.

A amplitude de uma série de valores é determinada rápida e facilmente, oferecendo uma ideia grosseira do grau de dispersão.

Conhecendo-se o valor da média e da amplitude, têm-se dados sobre o centro da distribuição e da dispersão em torno desse ponto.

$$It = X_{máx} - X_{min}$$

It = Intervalo total
$X_{máx}$ = valor máximo
X_{min} = valor mínimo

Desvantagens da utilização da amplitude total

– Por depender somente dos valores extremos da série, torna-se insensível à dispersão dos demais valores, compreendido entre o ponto máximo e o mínimo; principalmente quando a série é grande e existem lacunas extensas.

– Não é uma medida exata. A alteração de apenas um valor extremo da série ocasiona uma mudança brusca de amplitude. Esta também é afetada pelo tamanho da amostra.

Desvio médio

Baseia-se na primeira propriedade da média: “A soma algébrica dos afastamentos dos valores em relação à média aritmética é nula.” No entanto, para seu cálculo apenas são considerados os valores absolutos dos desvios, logo, representa o quociente do somatório dos desvios a partir da média pelo número total de casos. Trata-se, então, da média aritmética dos desvios em torno do valor central.

$$DM = \frac{\sum_{i\ 1}^{N} |X|}{N}$$

DM = Desvio médio

$|X| = X_i - \overline{X}$

N = número total de casos

Este valor é uma estimativa da amplitude dentro da qual variam as observações médias do conjunto de itens ou mensurações.

A soma dos valores absolutos das discrepâncias tende a tornar-se maior à medida que a variabilidade da distribuição aumenta.

Cálculo do desvio médio:

1°) Calcular a média aritmética da distribuição.

2°) Subtrair de cada escore bruto a média aritmética.

3°) Somar todos os valores absolutos dos desvios.

4°) Dividir o somatório dos desvios pelo número total de casos.

O desvio médio não é muito utilizado atualmente pelos pesquisadores. Na maioria dos casos ele é substituído pelo desvio-padrão.

Desvio-padrão

O desvio médio é uma medida de pouco valor, pois não considera os sinais dos desvios. Uma tentativa de superar esta dificuldade reside na possibilidade de se elevar ao quadrado os desvios, tornando-os, dessa forma, positivos.

Denomina-se *variância* a média dos quadrados dos desvios tomados a partir do conjunto.

$$V = \frac{\sum_{i=1}^{N} \cdot X_{i^2}}{N}$$

V= variância

x_i^2= desvios ao quadrado

N= número total de casos

A raiz quadrada positiva da variância é o que se chama *desvio* ou *afastamento quadrático* ou simplesmente *desvio-padrão.*

$$S = +\sqrt{\frac{\sum_{i=l}^{N} \cdot X_{i^2}}{N}}$$

S= desvio-padrão

x_i^2 = desvios ao quadrado

N= número total de casos

Cálculo do desvio-padrão

1°) Calcular a média aritmética da distribuição.

2°) Subtrair de cada escore bruto o valor da média aritmética.

3°) Elevar os desvios ao quadrado.

4°) Somatório de todos os desvios ao quadrado.

5°) Dividir o somatório pelo número total de casos.

6°) Extrair a raiz quadrada do resultado e dar-lhe sinal positivo.

Exemplo: Escores obtidos por 10 crianças na dimensão duração

X_i	F_i	F_iX_i	$x_i(X_i-\bar{X})$	x_i	x_1^2	x_iF
1	2	3	1-2,8	-1,8	3,24	9,72
2	2	4	2-2,8	-0,8	0,64	1,28
3	2	6	3-2,8	+0,2	0,04	0,08
4	1	4	4-2,8	-1,2	1,44	1,44
5	1	5	5-2,8	+2,2	4,84	4,84
6	1	6	6-2,8	+3,2	10,24	10,24
	10	28				27,6

$$S = \sqrt{\frac{27,6}{10}} = \sqrt{2,76} = +1,66 \qquad \bar{X} = \frac{28}{10} = 2,8$$

Propriedades do desvio-padrão

1°) A média aritmética, aumentada e subtraída de um desvio-padrão, indica uma faixa de normalidade na qual há uma incidência maior das observações. Estas observações representam, via de regra, 68,26% do total. O restante das observações, fora desse limite, é superior ou inferior às consideradas normais e aparece numa percentagem equivalente a 15,9% em cada extremidade da distribuição. Tais percentagens são as que aparecem nas distribuições normais.

2°) O desvio-padrão aumenta à medida que aumenta a dispersão em torno da média aritmética.

3°) Somando-se ou subtraindo-se todos os valores da série por uma constante, o desvio-padrão não será alterado.

4º) Se todos os valores da série forem multiplicados por uma constante, o desvio-padrão será aumentado na mesma proporção.

A interpretação do desvio-padrão pode ser feita observando-se os valores numéricos. À medida que estes decrescem, menor será a variabilidade, e quanto mais aumentam, maior.

Referindo-se à tabela das médias aritméticas e desvios-padrão (pág. 17), verificamos que, em relação às dimensões extensão e significado, os desvios-padrão decrescem à medida que o desenvolvimento cognitivo

evolui. Conclui-se que os grupos se tornam mais homogêneos à medida que há evolução cognitiva.

No período das operações formais, a concentração dos valores em torno da média aritmética é maior. No entanto, com relação à dimensão duração, as crianças do subperíodo das operações concretas apresentam-se mais heterogêneas.