Parte I – Estatística descritiva

4. Medidas de tendência central

Ana Maria Lopez Calvo de Feijoo

FEIJOO, AMLC. Medidas de tendência central. In: *A pesquisa e a estatística na psicologia e na educação* [online]. Rio de Janeiro: Centro Edelstein de Pesquisas Sociais, 2010, pp. 14-22. ISBN: 978-85-7982-048-9. Available from SciELO Books <http://books.scielo.org>.

4. MEDIDAS DE TENDÊNCIA CENTRAL

a) *Definição*: são as medidas típicas ou representativas de um conjunto de dados.

b) *Objetivo*: indicar o valor típico ou prevalente de uma distribuição de frequência, quando esta apresenta os valores intermediários da variável com frequência maior que os valores extremos, ou seja, uma tendência central.

c) *Tendência central*: é a tendência das notas incidirem para o centro da distribuição.

d) *Utilidade*: diante de uma distribuição de frequência, faz-se necessário dispor de um número que nos indique onde está a tendência central, ou, então, o valor mais capaz de substituir todos os outros.

e) *Medida mais usada*: média aritmética.

f) *Outras medidas*: mediana, moda.

g) *Restrições*: por si só dão informações insuficientes; precisam ser acopladas a uma medida de variabilidade.

Média aritmética

A média aritmética representa o "centro de gravidade" da distribuição, isto é, o ponto de qualquer distribuição em torno do qual se equilibram as discrepâncias positivas e negativas.

Situa-se entre o valor máximo e o valor mínimo da distribuição. Não pode, portanto, ser inferior ou superior ao valor mínimo e ao máximo da distribuição.

A média aritmética é um valor que pretende ser o resumo de todos os valores da distribuição. Dessa forma, pode vir a ser um valor não presente na distribuição.

Permite fazer interpretações quando é utilizada na comparação de dois ou mais grupos, constatando qual é o grupo com resultados mais ou menos elevados.

Cálculo da média aritmética

No cálculo do valor da média aritmética faz-se a soma dos valores e divide-se esta pelo número de observações da série.

DADOS NÃO AGRUPADOS

Acha-se o quociente do somatório dos valores da série pelo número deles.

$$X_i = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{N}$$

X_1, X_2, ..., X_n são valores particulares que as variáveis assumem naquela série de observações.

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

$\overline{X}$ = média aritmética

X_i = nota do i-ésimo elemento

N = número total de notas

Σ = somatório

Na distribuição de frequência, alguns valores podem ocorrer mais de uma vez. É possível, portanto, elaborar uma distribuição de frequência e verificar que o produto do valor da variável pela sua respectiva frequência ($F_i.X_i$) é idêntico à soma de todos os valores iguais de uma variável. Neste caso, para se obter a média aritmética, somam-se todos os produtos $F_i.X_i$ e divide-se o resultado pelo número total de casos.

$$\overline{X} = \frac{\sum_{i=1}^{N} F_i \,.\, X_i}{N}$$

$\overline{X}$ = média aritmética

$F_i.X_i$ = nota do i-ésimo elemento

N = total de casos

Σ = somatório

Exemplo: Escores obtidos pelas crianças na dimensão duração

X_i	F_i	F_iX_i
1	3	3
2	2	4
3	2	6
4	1	4
5	1	5
6	1	6
$\sum$	10	28

$$\overline{X} = \frac{\sum X_i \, . \, F_i}{N}$$

$$\overline{X} = \frac{20}{10}$$

$$\overline{X} = 2{,}8$$

Propriedades da média aritmética

1º) A soma algébrica dos afastamentos ou desvios dos valores da série em relação à média é nula.

Essa propriedade implica a seguinte afirmativa: se calcularmos os desvios em relação a um outro termo qualquer da série, diferente da sua média aritmética, a soma destes desvios será diferente de zero.

2°) A média aritmética é influenciada pelas alterações sofridas pelos valores da série.

Ficará aumentada na mesma quantidade que for adicionada a todos os valores da série.

Ficará diminuída da quantidade que for subtraída a todos os valores da série.

Ficará multiplicada pela quantidade que for multiplicada a todos os valores da série.

Ficará dividida pela quantidade pela qual foram divididos todos os valores da série.

3°) A soma dos quadrados dos afastamentos dos valores da série em relação à média aritmética é um mínimo.

Isto significa que se tornarmos a soma dos quadrados dos desvios em relação a outro termo qualquer da série, esta soma será sempre maior do que a encontrada, utilizando-se os quadrados dos desvios em relação à média aritmética.

4°) A média aritmética depende de todos os valores da série, porque todos entram no seu cálculo, sendo por isso um valor representativo da série.

5°) A média aritmética é grandemente influenciada pelos valores extremos da série.

Vantagens e desvantagens na utilização da média aritmética

– É influenciada pelos valores extremos, ou seja, um valor excepcionalmente alto ou baixo a altera. Por isso não deve ser utilizada quando a distribuição é muito assimétrica.

– Permite apenas uma descrição incompleta na distribuição, por não fornecer informações a respeito do número de casos que estão acima ou abaixo do valor da média aritmética; e também a respeito do valor mais frequente.

Na pesquisa realizada por Torres (1978), verifica-se a utilização da média aritmética para fins de comparações entre distribuições.

Médias e desvios-padrão obtidos em cada dimensão nos três grupos considerados

	Subperíodo pré-operacional (PO)	**Subperíodo de operação concreta (O)**	**Período de operações formais (OF)**
Extensão	$\overline{X}$ = 15,67 S = 5,36	$\overline{X}$ = 24,73 S = 3,99	$\overline{X}$ = 28,87 S = 3,54
Significado	$\overline{X}$ = 16,79 S = 8,34	$\overline{X}$ = 29,23 S = 5,46	$\overline{X}$ = 33,62 S = 4,07
Duração	$\overline{X}$= 3,37 S = 2,47	$\overline{X}$ = 7,85 S = 5,46	$\overline{X}$= 9,40 S = 0,57

Analisando-se as médias aritméticas em relação às três dimensões, verifica-se que houve uma evolução do conceito de morte na medida em que há uma evolução cognitiva, pois houve um aumento das médias.

Observando-se as médias aritméticas das três dimensões, conclui-se que a compreensão do conceito de morte, quanto à sua duração, é mais difícil qualquer que seja a fase cognitiva em que se encontre a criança, visto que nesta dimensão os valores das médias aritméticas são mais baixos do que os demais.

A tabela apresenta também valores dos desvios-padrão que devem ser acoplados aos valores da média aritmética como veremos mais adiante.

Mediana

A mediana é o valor médio de uma distribuição ordenada, o qual apresenta o mesmo número de valores abaixo e acima desse valor.

A mediana é um ponto da distribuição, tal que a probabilidade de um valor qualquer da distribuição, aleatoriamente escolhido, se situar acima da mediana é igual à probabilidade dele se situar abaixo da mesma. Isto é válido para qualquer distribuição, não importando a sua forma.[*]

Através da interpretação do valor mediano, pode-se afirmar, ao comparar dois ou mais grupos, qual é o que apresenta resultados mais elevados, e qual o que apresenta resultados menos elevados.

Cálculo da mediana

Para se calcular o valor da mediana, primeiramente organizam-se os dados em ordem crescente ou decrescente.

O valor da mediana é aquele abaixo e acima do qual encontra-se a mesma quantidade de valores da série.

Primeiramente, localiza-se a mediana, e depois, observa-se o seu valor.

[*] Nick, notas de aula.

Posição $\frac{N+1}{2}$

N= número de casos

Exemplo: valores obtidos pelas crianças na dimensão duração no período pré-operacional

X_i	F_i	F_{ac}
1	3	3
2	2	5
3	2	7 ⇐
4	1	8
5	1	9
6	1	10
		10

$$\frac{N+1}{2} = \frac{11}{2} = 5{,}5$$

Md (posição)
5,5 °

$$\text{Md} = L_{ir} + \frac{N/2 - Fac}{F_i} \, . \, h$$

Md = mediana
L_{ir} = limite inferior real
N / 2 = 50%
Fac = frequência acumulada da classe anterior
F_i = frequência simples da classe

$$\text{Md} = 2{,}5 + \frac{5{,}0 - 5}{2} = 2{,}5$$

Ordenar : 4, 2, 2, 1, 3, 5, 6, 3, 1, 1

1, 1, 1 ,2 ,2 ,3, 3, 4, 5, 6

$$\text{Md} = \frac{N+1}{2} = \frac{11}{2} = 5,5° \qquad \text{Md} = 2{,}5$$

Propriedades da mediana

1°) A mediana não é influenciada pelos valores extremos.

2°) A soma dos valores absolutos dos desvios a partir da mediana é um mínimo.

3°) A mediana de uma série de classes indefinidas pode ser calculada, desde que conheça o número total de observações.

Vantagens na utilização da mediana

Sua determinação é fácil e rápida, não requer cálculos complexos.

Juntando-se o mesmo número de termos nas duas extremidades, a mediana continua sendo a mesma.

Desvantagens na utilização da mediana

– A mediana flutua mais de amostra para amostra do que a média aritmética; portanto, é menos confiável. Não utiliza a totalidade dos dados.

– É um valor posicional, não vem definido por uma expressão matemática; portanto, não é susceptível de tratamento algébrico. Não é possível calcular a mediana de um grupo total a partir das medianas de dois subgrupos.

Moda

É o valor da distribuição que ocorre com a maior frequência, ou seja, o valor que mais se repete dentro de uma série de observações.

A moda só pode ser utilizada como medida de tendência central quando apenas um valor da série ocorre com maior frequência.

Exemplo: Escore obtido por 10 crianças na dimensão duração

X_i	F_i
1	3
2	2
3	2
4	1
5	1
6	1
$\sum$	10

O escore que apareceu maior número de vezes foi o 1**;** portanto, o valor da moda é igual a 1. Consequentemente, este valor seria o valor representativo da distribuição.

Vantagens e desvantagens na utilização da moda

Embora de cálculo fácil, não pode ser utilizada em distribuições bimodais ou multimodais.

Utilização das medidas de tendência central

Em um dado momento, podem surgir dúvidas sobre que medida de tendência central utilizar; no entanto, dois fatores devem ser averiguados:

1º) o aspecto ou forma da distribuição

2º) o objetivo da pesquisa

Forma da distribuição

A forma da distribuição pode influenciar o pesquisador na escolha de uma medida de tendência central. Em uma distribuição unimodal e perfeitamente simétrica, a moda, a mediana e a média serão idênticas, uma vez que o ponto de frequência máxima (Mo) é também o valor que divide a distribuição em duas partes, contendo o mesmo número de termos em cada uma das partes; é também o "centro de gravidade" distribuição.

Na distribuição assimétrica à esquerda, a média aritmética é deslocada para a esquerda da moda; na distribuição inclinada para a direita, a média incide à direita da moda.

Em cada caso, a média aritmética é deslocada na mesma direção da inclinação da distribuição.

A direção do deslocamento da mediana é o mesmo da direção do deslocamento da média, mas a extensão deste deslocamento é menos do que o da média aritmética, visto que esta é influenciada pelos valores extremos.

Em uma distribuição assimétrica, a mediana sempre se situa em algum lugar entre a média e a moda. É essa característica que a torna a medida de tendência central preferida, por alguns pesquisadores, para representar uma distribuição assimétrica.

Objetivo da pesquisa

A escolha da medida de tendência central depende das hipóteses ou objetivos do pesquisador.

Utilizará a moda se pretender uma medida descritiva, rápida e simples ainda que grosseira, e se a distribuição for unimodal. Se a pretensão for uma medida exata, ele poderá optar entre a média e a mediana. Se a distribuição for aproximadamente simétrica, a média aritmética é a mais indicada, mesmo porque esta poderá ser utilizada em estatística mais avançada e é uma medida mais estável.