UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ

Estatística Básica

Medidas de Dispersão ou Variabilidade

Professora Ma. Tainara Volan
tainaravolan@gmail.com

2022

Medidas de Dispersão ou Variabilidade

Na outra aula vimos média aritmética, mediana e moda -> tais valores podem servir de comparação para dar a posição de qualquer elemento do conjunto.

No entanto, quando se trata de interpretar dados estatísticos, mesmo aqueles já simplificados, devemos analisar a variabilidade dos dados.

Uma variável em uma pesquisa, quando coletados os seus resultados, para cada elemento da amostra pode assumir valores diferentes entre si, ocorre que estes valores podem situar-se próximos entre si ou não, e em alguns casos chegam a ser discrepantes.

Medidas de Dispersão ou Variabilidade

Medidas de dispersão visam atribuir um valor numérico, da melhor forma possível, o qual expresse esta homogeneidade ou não entre os diversos valores obtidos.

Existem para qualificar os valores de uma dada variável, ressaltando a maior ou menor dispersão ou variabilidade entre esses valores e a sua medida de posição Dessas medidas, estudaremos a:

- amplitude total
- a variância
- o desvio padrão
- o coeficiente de variação.

UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ

Estatística Básica

Amplitude total

Professora Ma. Tainara Volan
tainaravolan@gmail.com

2022

Amplitude total

A **amplitude total** é a diferença entre o maior e o menor valor observado.

Sem intervalo de classe: $AT = x(máx) - x(mín)$

Com intervalo de classe: $AT = L(máx) - L(mín)$

Quanto maior a amplitude total, maior a dispersão ou variabilidade dos valores da variável.

Amplitude total

Exemplo: Para os valores: 40, 45, 48, 52, 54, 62 e 70

Temos:

Sem intervalo de classe: $AT = x(máx) - x(mín)$

$$AT = 70 - 40$$

$$AT = 30$$

X: 70, 70, 70, 70, 70

AT = 70 – 70 = 0 (dispersão nula)

Y: 68, 69, 70, 71, 72

AT = 72 – 68 = 4

Z: 5, 15, 50, 120, 160

AT = 160 – 5 = 155

UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ

Amplitude total

Exemplo:

Com intervalo de classe: $AT = L(máx) - L(mín)$

$$AT = 174 - 150$$

$$AT = 24\ cm$$

i	Estaturas (cm)	fi
1	150 |- 154	4
2	154 |- 158	9
3	158 |- 162	11
4	162 |- 166	8
5	166 |- 170	5
6	170 |- 174	3
		$\sum = 40$

Amplitude total

A amplitude total tem o inconveniente de só levar em conta os dois valores extremos da série, descuidando do conjunto de valores intermediários, o que quase sempre invalida a idoneidade do resultado. Ela é apenas uma indicação aproximada da dispersão ou variabilidade.

UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ

Estatística Básica

Variância e Desvio padrão

Professora Ma. Tainara Volan
tainaravolan@gmail.com

2022

Variância e Desvio padrão

Como vimos, a amplitude total é instável, por se deixar influenciar pelos valores extremos.

A variância e o desvio padrão são medidas que fogem a essa falha, pois levam em consideração a **totalidade dos valores da variável em estudo**, o que faz delas índices de variabilidade bastante **estáveis** e, por isso mesmo, os mais geralmente empregados.

Variância e Desvio padrão

A variância baseia-se nos desvios em torno da média aritmética, porém determinando a média aritmética dos quadrados dos desvios. Assim, representando a variância por s^2, temos:

$$s^2 = \frac{\sum (x_i - \overline{x})^2}{n-1}$$

ATENÇÃO: Para populações usa-se n ao invés de n-1, valor utilizado para cálculo em amostras

Variância e Desvio padrão

Exemplo: variância

Para as amostras: 3, 4, 5, 6, 7

$$s^2 = \frac{\sum(x_i - \overline{x})^2}{n-1} = \frac{[(3-5)^2 + (4-5)^2 + (5-5)^2 + (6-5)^2 + (7-5)^2]}{4} = 2{,}5$$

E para as amostras: 1, 3, 5, 7, 9

$$s^2 = \frac{\sum(x_i - \overline{x})^2}{n-1} = \frac{[(1-5)^2 + (3-5)^2 + (5-5)^2 + (7-5)^2 + (9-5)^2]}{4} = 10$$

A amostra 3, 4, 5, 6, 7 é mais homogênea.

Variância e Desvio padrão

Sendo a variância calculada a partir dos quadrado dos desvios, ela é um número em unidade quadrada em relação à variável em questão, o que, sob o ponto de vista prático, é um inconveniente.

Por isso mesmo, imaginou-se uma nova medida que tem utilidade e interpretação práticas, denominada desvio padrão, definida como a raiz quadrada da variância

e representada por s.

$$s = \sqrt{\frac{\sum(x_1 - \overline{x})^2}{n-1}}$$

Tanto o desvio padrão quanto a variância são usados como medidas de dispersão ou variabilidade. O uso de uma ou de outra dependerá da finalidade que se tenha em vista.

UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ

Variância e Desvio padrão

Exemplo: desvio padrão

Para as amostras: 3, 4, 5, 6, 7

$$s^2 = 2{,}5$$

$$s = \sqrt{2{,}5} = 1{,}58$$

E para as amostras: 1, 3, 5, 7, 9

$$s^2 = 10$$

$$s = \sqrt{10} = 3{,}16$$

Desvio padrão – dados não agrupados

Não apenas este método é usualmente mais prático, como também mais preciso.

$$s = \sqrt{\frac{\sum x_i^2 - \left[\frac{(\sum x_i)^2}{n}\right]}{n-1}}$$

Desvio padrão – dados não agrupados

Exemplo: calcular o desvio padrão da amostra 1, 3, 5, 7, 9 para a variável x:

$$s = \sqrt{\frac{\sum x_i^2 - \left[\frac{(\sum x_i)^2}{n}\right]}{n-1}}$$

$$s = \sqrt{\frac{165 - \left(\frac{25^2}{5}\right)}{4}} = \sqrt{\frac{165-125}{4}} = \sqrt{\frac{40}{4}} = \sqrt{10} = 3{,}16$$

x_i	x_i^2
1	1
3	9
5	25
7	49
9	81
$\sum = 25$	$\sum = 165$

UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ

Desvio padrão – dados agrupados

Para dados agrupados, não esquecer que se tem a presença das frequências, como fazemos no cálculo da média. Para dados agrupados com intervalos de classe, considerar o ponto médio da classe (Xi) ponderado pela fi, exatamente como vimos no cálculo de média

$$s = \sqrt{\frac{\sum f_i x_i^2 - \left[\frac{\left(\sum f_i x_i\right)^2}{n}\right]}{n-1}}$$

UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ

Desvio padrão – dados agrupados

Exemplo:

$$s = \sqrt{\frac{\sum f_i x_i^2 - \left[\frac{(\sum f_i x_i)^2}{n}\right]}{n-1}}$$

$$s = \sqrt{\frac{165 - \left[\frac{63^2}{30}\right]}{29}} = \sqrt{\frac{165 - 132{,}3}{29}} = \sqrt{\frac{32{,}7}{29}} = \sqrt{1{,}127} = 1{,}06$$

x_i	x_i^2	f_i	$f_i x_i$	$f_i x_i^2$
0	0	2	0	0
1	1	6	6	6
2	4	12	24	48
3	9	7	21	63
4	16	3	12	48
		$\sum = 30$	$\sum = 63$	$\sum = 165$

Importante

Condições para se usar o desvio-padrão ou variância para comparar a variabilidade entre grupos:

- mesmo número de observações
- mesma unidade
- mesma média

Além disso, se quisermos comparar duas ou mais amostras de valores expressos em unidades diferentes, não poderá ser possível fazer a comparação por meio do desvio-padrão, pois ele é expresso na mesma unidade dos dados.

UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ

Estatística Básica

Coeficiente de variação

Professora Ma. Tainara Volan
tainaravolan@gmail.com

2022

Coeficiente de variação

O fato de o desvio padrão ser expresso na mesma unidade dos dados limita o seu emprego quando desejamos comparar duas ou mais séries de valores, relativamente à sua dispersão ou variabilidade, quando expressas em unidades diferentes. Para isto, existe o Coeficiente de Variação.

Coeficiente de variação

Para contornar essas dificuldades e limitações, podemos caracterizar a dispersão ou variabilidade dos dados em termos relativos a seu valor médio, medida essa denominada coeficiente de variação (CV):

$$CV = \frac{s}{x} \cdot 100$$

Coeficiente de variação

	$\overline{x}$	s
Estaturas	175 cm	5,0 cm
Pesos	68 kg	2,0 kg

Exemplo:

$$CV_{estaturas} = \frac{s}{x} \cdot 100 = \frac{5}{175} . 100 = 2{,}85\%$$

$$CV_{pesos} = \frac{s}{x} \cdot 100 = \frac{2}{68} . 100 = 2{,}94\%$$

Conclusão: para este grupo os **pesos** tem maior grau de **dispersão** que as estaturas

Tipos de dispersão

O coeficiente de variação é expresso em porcentagem (%), avaliado para amostras segundo a seguinte referência:

- Baixa dispersão: CV $\leq$ 15%
- Média dispersão: 15% $<$ CV $<$ 30%
- Grande dispersão: CV $\geq$ 30%

UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS CHAPECÓ