Introdução,
Estatística Descritiva

Medidas de dispersão

- Possibilitam obter informações sobre a variabilidade de um conjunto de observações. Indicar o grau de afastamento de um conjunto de números em relação a média

AMPLITUDE

- Diferença entre o maior e o menor valor em um conjunto de observações

Exemplo: No caso da venda de pizzas.

38, 38, 40, 46, 49, 52, 56, 59, 63

- ✓ Amplitude = 63 – 38
- ✓ Amplitude = 25

Medidas de dispersão

DESVIO MÉDIO ABSOLUTO

➢ Determina quão distante da média cada valor observado está, para tanto, calcula-se o módulo das diferenças de cada valor pela média.

Vendas de pizza por dia	Distância da média
40	9 = |49 – 40|
56	7 = |49 – 56|
38	11 = |49 – 38|
38	11 = |49 – 38|
63	14= |49 – 63|
59	10 = |49 – 59|
52	3 = |49 – 52|
49	0 = |49 – 49|
46	3 = |49 – 46|

$$dm(x) = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

$$dm = \frac{9+7+11+11+14+10+3+0+3}{9} = \frac{68}{9} = 7{,}556$$

Medidas de dispersão

VARIÂNCIA

➢ Calcula-se o quadrado de cada desvio e toma-se a média de todos esses quadrados.

$$\mathrm{var}(x) = \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

Exemplo: No caso da venda de pizzas.

$$\begin{aligned}
&(40-49)^2 + (56-49)^2 + (38-49)^2 + (38-49)^2 + (63-49)^2 + \\
&+ (59-49)^2 + (52-49)^2 + (49-49)^2 + (46-49)^2 = \\
&= (-9)^2 + (7)^2 + (-11)^2 + (-11)^2 + (14)^2 + (10)^2 + (3)^2 + (0)^2 + (-3)^2 = \\
&= 81 + 49 + 121 + 121 + 196 + 100 + 9 + 0 + 9 = \\
&= 686
\end{aligned}$$

$$\mathrm{var}(x) = \sigma^2 = \frac{686}{9} = 76{,}222$$

Medidas de dispersão

DESVIO PADRÃO

- É a raiz quadrada da variância

$$\text{desvio} - \text{padrão}(x) = \sigma = \sqrt{\text{var}(x)}$$

Exemplo: No caso da venda de pizzas.

$$\sigma = \sqrt{\text{var}(x)} = \sqrt{76{,}222} = 8{,}731$$

Medidas de dispersão

COEFICIENTE DE VARIAÇÃO

- Mede a dispersão dos dados em relação à média;
- Indica se a distribuição é homogênea ou não;
- Se o Coeficiente de Variação for menor do que 20%, a distribuição é considerada homogênea;
- Usado para comparação entre 2 ou mais grupos.

Exemplo: No caso da venda de pizzas.

$$\sigma = \sqrt{\text{var}(x)} = \sqrt{76{,}222} = 8{,}731$$

$$\text{coeficiente de variação} = \frac{\text{desvio - padrão}}{\text{média}} = \frac{\sigma}{\bar{x}}$$

$$\text{coeficiente de variação} = \frac{8{,}731}{49} = 0{,}178 = 17{,}8\%$$

Distribuição é considerada homogênea

Medidas de dispersão

Funcionários	Quantidade de peças produzidas por dia				
	Segunda	Terça	Quarta	Quinta	Sexta
A	10	9	11	12	8
B	15	12	16	10	11
C	11	10	8	11	12
D	8	12	15	9	11

Deseja saber em média, quantos produtos são produzidos por cada funcionário em um dia.

Funcionários	Média Aritmética ($\bar{x}$)	
A	$\bar{X}_A = \frac{10 + 9 + 11 + 12 + 8}{5} = \frac{50}{5}$	$\bar{X}_A = 10{,}0$
B	$\bar{X}_B = \frac{15 + 12 + 16 + 10 + 11}{5} = \frac{64}{5}$	$\bar{X}_B = 12{,}8$
C	$\bar{X}_C = \frac{11 + 10 + 8 + 11 + 12}{5} = \frac{52}{5}$	$\bar{X}_C = 10{,}4$
D	$\bar{X}_D = \frac{8 + 12 + 15 + 9 + 11}{5} = \frac{55}{5}$	$\bar{X}_D = 11{,}0$

Medidas de dispersão

Variância → Funcionário A:

$$\text{var (A)} = \frac{(10-10)^2 + (9-10)^2 + (11-10)^2 + (12-10)^2 + (8-10)^2}{5}$$

$$\text{var (A)} = \frac{10}{5} = 2{,}0$$

Variância → Funcionário B:

$$\text{var (B)} = \frac{(15-12{,}8)^2 + (12-12{,}8)^2 + (16-12{,}8)^2 + (10-12{,}8)^2 + (11-12{,}8)^2}{5}$$

$$\text{var (B)} = \frac{26{,}8}{} = 5{,}36$$

Variância → Funcionário C:

$$\text{var (C)} = \frac{(11-10{,}4)^2 + (10-10{,}4)^2 + (8-10{,}4)^2 + (11-10{,}4)^2 + (12-10{,}4)^2}{5}$$

$$\text{var (C)} = \frac{9{,}2}{5} = 1{,}84$$

Variância → Funcionário D:

$$\text{var (D)} = \frac{(8-11)^2 + (12-11)^2 + (15-11)^2 + (9-11)^2 + (11-11)^2}{5}$$

$$\text{var (D)} = \frac{30}{5} = 6{,}0$$

variância, *uma medida de dispersão que mostra quão distantes os valores estão da média.*

Quanto maior for a **variância**, mais distantes da média estarão os valores, e quanto menor for a **variância**, mais próximos os valores estarão da média.

Mediada muito influenciada por valores que estão muito distantes da média

Medidas de dispersão

Desvio Padrão → Funcionário A:

$$dp(A) = \sqrt{var\,(A)}$$
$$dp(A) = \sqrt{2,0}$$
$$dp(A) \approx 1,41$$

Desvio Padrão → Funcionário B:

$$dp(B) = \sqrt{var\,(B)}$$
$$dp(B) = \sqrt{5,36}$$
$$dp(B) \approx 2,32$$

Desvio Padrão → Funcionário C:

$$dp(C) = \sqrt{var\,(C)}$$
$$dp(C) = \sqrt{1,84}$$
$$dp(C) \approx 1,36$$

Desvio Padrão → Funcionário D:

$$dp(D) = \sqrt{var\,(D)}$$
$$dp(D) = \sqrt{6,0}$$
$$dp(D) \approx 2,45$$

O **desvio padrão (dp)** é simplesmente o resultado positivo da raiz quadrada da variância. Na prática, o desvio padrão indica qual é o “erro” se quiséssemos substituir um dos valores coletados pelo valor da média.

$$dp = \sqrt{\sigma^2}$$

Medidas de dispersão

Podemos ver a utilização do desvio padrão na apresentação da média aritmética, informando o quão "confiável" é esse valor. Isso é feito da seguinte forma:

média aritmética (x) ± desvio padrão (dp)

Se o dono da empresa de nosso exemplo pretende concluir seu relatório com a produção média diária de seus funcionários, ele fará da seguinte forma:

Funcionário A: 10,0 ± 1,41 peças por dia
Funcionário B: 12,8 ± 2,32 peças por dia
Funcionário C: 10,4 ± 1,36 peças por dia
Funcionário D: 11,0 ± 2,45 peças por dia