1 Aula 3 - { Medidas de posição
Medidas de dispersão

Medidas de posição:

$\frac{n_i}{n}$

⇒ Tabela de Resumo (Frequências)

x_i f_i

Nº de filhos z_i	Freqüência n_i	Porcentagem $100 f_i$
0	4	20
1	5	25
2	7	35
3	3	15
5	1	5
Total	20	100

Precisamos de mais informações que nos ajude a seguir no caminho da identificação da dist.

P/ Posição {
- # Média ✓
- # Mediana ✓
- # Moda ✓

Vamos rever a tab Resumo e retirar as informações:

0, 0, 0, 0 } 4

1, 1, 1, 1, 1 } 5

2, 2, 2, 2, 2, 2, 2 } 7

3, 3, 3 } 3

5 } 1

Moda (Mo)

→ Nº q. aparece c/ maior frequência absoluta

Mo = 2

$\boxed{\text{é possível q. } \nexists \text{ Mo} \quad \text{AMODAL}}$

Média (Me)

Notação { Parâmetro Populacional (μ)
Estatística Amostral ($\bar{X}$)

$$Me = \frac{1}{N}\sum_{i=1}^{N} x_i = \mu$$

$$\boxed{Me = \frac{1}{n}\sum_{i=1} x_i = \bar{X}}$$

Fornecedor	
1	0
2	3
3	8
4	2
5	1
6	0
7	5
8	10

0 1 2 3 5 8 10

X_1 X_2 X_3 X_4 X_5 X_6 X_8

$X_{(1)}$

Prazo baixo — Md — Prazo alto

$$Md_{(x)} = \begin{cases} n \text{ par} \\ \text{Impar} \rightarrow Md = X_{\left(\frac{n+1}{2}\right)} \end{cases}$$

Par:

$$Md_{(x)} = \left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)} \right) \Big/ 2$$

② Média { importância do N ou n

P/nosso caso (n) — Amostra

$\Rightarrow n = 20$

$$\therefore \bar{X} = \frac{1}{20} \cdot \sum_{i=1}^{20} x_i$$

$$\bar{X} = \frac{\overbrace{0+0+0+0}^{4} + \overbrace{1+1+1+1+1}^{5} + \overbrace{2+2+2+2+2+2+2}^{7} + \overbrace{3+3+3}^{3} + \overbrace{5}^{1}}{20} = \frac{33}{20}$$

$n_i \quad x_i$

$$\bar{X} = \frac{4(0) + 5(1) + 7(2) + 3(3) + 1(5)}{20} = \frac{33}{20} = 1,65$$

20

Posso fazer $f_i = \frac{n_i}{n} \Rightarrow \boxed{\bar{X} = \sum_{i=1}^{n} f_i x_i = 1,65}$

Importante:

$\Rightarrow$ Só conseguimos gerar Me p/ variáveis Quantitativas {discretas/contínuas

Matemática é importante: Esperança (E)

{ Média Aritmética de x (μ ou $\bar{x}$)

Esperança matemática de x ($E(x)$ ou $E(\bar{x})$)

→ "centro de gravidade do conjunto de dados"

Veja que: $\boxed{E(x) = \mu = \frac{1}{N} \sum_{i=1}^{N} x_i}$

$E(X) = \int f(x)\,x\,dx$

Propriedades:

1) $E(x \pm k) = k + E(x)$, c/ k = constante
2) $E(x \cdot k) = k \cdot E(x)$
3) $E(x \pm y) = E(x) \pm E(y)$
4) $E(x \cdot y) = E(x) \cdot E(y)$
5) $E(x - \bar{x}) = 0$

x, y } variáveis aleatórias

(variáveis contínuas) — futuramente //

3) Se tivermos variáveis agrupadas, como fazer?

Classes de salários	Ponto médio s_i	Freqüência n_i	Porcentagem $100 f_i$
4,00 ⊢ 8,00	6,00	10	27,78
8,00 ⊢ 12,00	10,00	12	33,33
12,00 ⊢ 16,00	14,00	8	22,22
16,00 ⊢ 20,00	18,00	5	13,89
20,00 ⊢ 24,00	22,00	1	2,78
Total	—	36	100,00

SALÁRIOS NA SUA EMPRESA

Qual a $\bar{S}$?

Veja aqui que geramos $S_i = \frac{S_I + S_S}{2} \Rightarrow$ P/ C_1

classe 1 (C_1)

$$S_1 = \frac{4+8}{2} = 6$$

$$S_2 = \frac{12+8}{2} = 10$$

⋮

$$\bar{S} = \frac{10(6)+12(10)+8(14)+5(18)+(22)}{36}$$

$$\bar{S} = 11,22$$

$$\Rightarrow \bar{S} = \sum_{i=1}^{n} f_i \cdot S_i$$

c/ S_i = Pto médio

$f_i = (n_i \cdot n^{-1})$

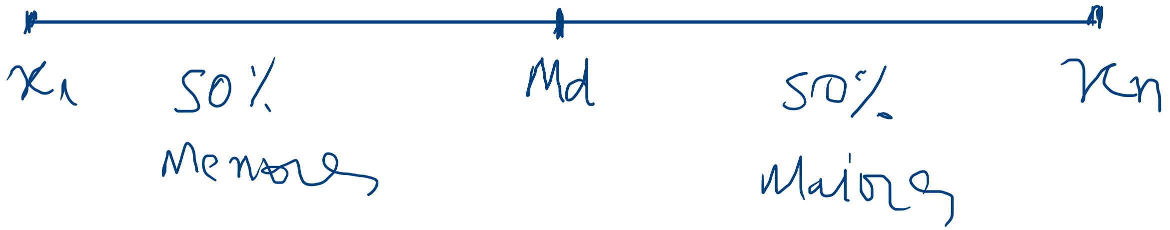
MEDIANA (Md)

↳ Divide a amostra entre dois grupos (Maiores e Menores)

* Exatamente na posição central

(50% — Maiores)

(50% — Menores)



Há um método p/ obtê-la?

1 – Ordenar os dados – (crescente)

x_1 (menor elemento)

↓

x_n (maior elemento)

$$\underbrace{x_1 < x_2 < \dots < x_n}_{\text{Estatística de ordem}}$$

4) Mediana

$$Md = \begin{cases} X_{\left(\frac{n+1}{2}\right)}, & \text{se } n \text{ é ímpar} \\ \dfrac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}, & \text{se } n \text{ é par} \end{cases}$$

Consideremos a X_i = nº de filhos, lembre que:

$$X = \{0, 0, 0, 0, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 5\}$$

$$x_1, x_2, x_3, x_4, x_5, \dots, x_{10}, \dots, x_{17}, \dots, x_{20}$$

n - par

$$\Rightarrow Md = \left(\frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}\right) = \underbrace{X_{\left(\frac{20}{2}\right)} + X_{\left(\frac{20}{2}+1\right)}}_{\frac{X_{10} + X_{11}}{2}}$$

$$\left.\begin{array}{l} X_{10} = 2 \\ X_{11} = 2 \end{array}\right\} \frac{2+2}{2} = 2$$

$$\therefore Md = 2 \;//$$

P/ o salário (variável contínua)

Relembre que:

TABELA RESUMO:

Classes de salários	Ponto médio x_i	Freqüência n_i	Porcentagem $100 f_i$
4,00 ⊢ 8,00	6,00	10	27,78
8,00 ⊢ 12,00	10,00	12	33,33
12,00 ⊢ 16,00	14,00	8	22,22
16,00 ⊢ 20,00	18,00	5	13,89
20,00 ⊢ 24,00	22,00	1	2,78
Total	—	36	100,00

Md p/ n = Par

$$Md = \left(\frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}\right)$$

$$\to X_{\left(\frac{36}{2}\right)} = X_{(18)} = 9{,}8$$

$$\to X_{(19)} = 10{,}53$$

$$Md \left(\frac{9{,}8 + 10{,}53}{2}\right) \cong 10{,}16 \;//$$

ORIGINAL

1	4,00
2	4,56
3	5,25
4	5,73
5	6,26
6	6,66
7	6,86
8	7,39
9	7,59
10	7,44
11	8,12
12	8,46
13	8,74
14	8,95
15	9,13
16	9,35
17	9,77
18	9,80
19	10,53
20	10,76
21	11,06
22	11,59
23	12,00
24	12,79
25	13,23
26	13,60
27	13,85
28	14,69
29	14,71
30	15,99
31	16,22
32	16,61
33	17,26
34	18,75
35	19,40
36	23,30

$M_o = \varnothing$

$A < B < C < D < E$

5) MEDIDAS DE DISPERSÃO

VEJA O PRÊMIO PAGO NAS AÇÕES DE EMPRESAS, AGRUPADAS EM DIFERENTES INDÚSTRIAS.

grupo A (variável X): 3, 4, 5, 6, 7
grupo B (variável Y): 1, 3, 5, 7, 9
grupo C (variável Z): 5, 5, 5, 5, 5
grupo D (variável W): 3, 5, 5, 7
grupo E (variável V): 3, 5, 5, 6, 6

Temos que:

$$\bar{X} = \sum_{i=1}^{n} \frac{x_i}{n} = 5{,}0$$

$\bar{Y} = 5{,}0$

$\bar{Z} = 5{,}0$

$\bar{W} = 5{,}0$

$\bar{V} = 5{,}0$

$\therefore \underbrace{\bar{X} = \bar{Y} = \bar{Z} = \bar{W} = \bar{V}}_{\text{onde investir?}}$

Qual grupo?

⇒ P/ responder podemos pensar na dispersão em torno da média.

USAMOS: # Variância
Desvio-Padrão
Desvio-Médio

} Sempre olhar em relação a média

GA:

Desvios: $\boxed{x_i - \bar{x}}$ como $\begin{cases} X = \{3, 4, 5, 6, 7\} \\ \bar{X} = 5 \end{cases}$

⇒ Desvio é dado por:

$(3-5) = -2$
$(4-5) = -1$
$(5-5) = 0$
$(6-5) = 1$
$(7-5) = 2$

Perceba que $\sum_{i=1}^{n} (x_i - \bar{x}) = 0$

Soluções:
- Desvios em valor absoluto
- Desvios ao quadrado

$\sum_{i=1}^{n} |x_i - \bar{x}| = 2 + 1 + 0 + 1 + 2 = 6$

$\sum_{i=1}^{n} (x_i - \bar{x})^2 = 4 + 1 + 0 + 1 + 4 = 10$

} Ajuda, mas confunde

6)

Vamos olhar p/ o desvio-médio:

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

$$dm(x) = \frac{2+1+0+1+2}{5} = \frac{6}{5} = 1,2$$

Vamos ver a variância:

$$Var(x) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} = \frac{4+1+0+1+4}{5} = 2$$

Podemos fazer isto p/ cada grupo

Grupo B:

$y = \{1, 3, 5, 7, 9\}$

$d = (1-5), (3-5), (5-5), (7-5), (9-5)$

$$dm(y) = \frac{\sum_{i=1}^{5} |y_i - \bar{y}|}{5} = \frac{4+2+0+2+4}{5} = \frac{12}{5} = 4$$

$$Var(y) = \frac{\sum_{i=1}^{5} (y_i - \bar{y})^2}{5} = \frac{16+4+4+16}{5} = 8$$

Grupo C:

$z = \{5, 5, 5, 5, 5\}$

$\{0, 0, 0, 0, 0\}$

$dm(z) = 0$

$Var(z) = 0$

Grupo D

$w = \{3, 5, 5, 7\}$

$$dm(w) = \frac{4}{4} = 1,0$$

$$Var(w) = \frac{8}{4} = 2,0$$

Grupo E

$v = \{3, 5, 5, 6, 6\}$

$d = (3-5), (0), (0), (1), (1)$

$$dm(v) = \sum_{i=1}^{5} \frac{2+1+1}{5} = \frac{4}{5} = 0,8$$

$$Var(v) = \sum_{i=1}^{5} = \frac{4+2}{5} = \frac{6}{5} = 1,2$$

7)

	GA	GB	GC	GD	GE
média	$\bar{x}=5$	$\bar{y}=5$	$\bar{z}=5$	$\bar{w}=5$	$\bar{v}=5$
dm( )	1,2	4	0	1,0	0,8
Var( )	2,0	8	0	2,0	1,2
dp( )	1,41	2,82	0	1,41	1,09

Qual grupo + heterogêneo? → B

Qual " + homogêneo? → C

E se estivermos falando de dados c/ frequências?

Reutilizando a tab:

Nº de filhos z_i	Freqüência n_i	Porcentagem $100 f_i$
0	4	20
1	5	25
2	7	35
3	3	15
5	1	5
Total	20	100

Já calculamos $\bar{z} = 1,65$

⇒ Vejamos os desvios:

$(0-1,65), (1-1,65), (2-1,65), (3-1,65), (5-1,65)$

$(-1,65), (-0,65), (0,35), (1,35), (3,35)$

Precisamos considerar suas frequências:

daí: $dm(z) = \sum_{i=1}^{n} f_i |z_i - \bar{z}|$ c/ $f_i = \frac{n_i}{n}$

$$dm(z) = \frac{4|1,65| + 5|0,65| + 7|0,35| + 3|1,35| + 1|3,35|}{20} = 0,98$$

8) $Var(z) = \sum_{i=1}^{n} f_i (z_i - \bar{z})^2$

P/o exemplo $Var(z) = 1{,}528$

$dP(z) = \sqrt{1{,}528} = 1{,}24$

Variância populacional — $Var(x)$

x

Variância Amostral — $S(x)$

$$Var(x) = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N} \text{ ou } \sum_{i=1}^{N} f_i (x_i - \bar{x})^2$$

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

$n-1$ → graus de liberdade

(o tamanho da Amostra é importante)