

Estatística Descritiva

Resumindo e Descrevendo Variáveis Quantitativas

Estatística Aplicada

Ana Maria Nogales Vasconcelos

Maria Teresa Leão Costa

MTLC - 2020

Medidas Resumo para Dados Agrupados

Dados Agrupados

Como determinar as medidas resumo??

DISTRIBUIÇÃO DO NÚMERO DE FALHAS EM UM SISTEMA DE COMUNICAÇÃO POR DIA -CIA. FACILCOM
SÃO PAULO – DEZEMBRO DE 2019

NÚMERO DE FALHAS	Nº de Dias
0	4
1	5
2	7
3	8
4	4
5	2
TOTAL	30

Fonte: Dados Fictícios

CONSUMO MENSAL DE ENERGIA ELÉTRICA
LOCALIDADE “CLARIDADE” - 2019

CONSUMO	Nº de Usuários
10 — 15	3
15 — 20	5
20 — 25	10
25 — 30	15
30 — 35	8
35 — 40	6
40 — 45	3
TOTAL	50

Fonte: Dados Fictícios

81

Medidas Posição para Dados Agrupados

Média

DISTRIBUIÇÃO DO NÚMERO DE FALHAS EM UM SISTEMA DE COMUNICAÇÃO POR DIA -CIA. FACILCOM
SÃO PAULO – DEZEMBRO DE 2019

NÚMERO DE FALHAS	Nº de Dias
0	4
1	5
2	7
3	8
4	4
5	2
TOTAL	30

Fonte: Dados Fictícios

00000+0+0+0 = 0 x 4
11111.....1+1+1+1+1 = 1 x 5
222222.....2+2+2+2+2+2 = 2 x 7
33333333...3+3+3+3+3+3+3+3 = 3 x 8
4444.....4+4+4+4 = 4 x 4
555+5 = 5 x 2

85

Média

X - variável quantitativa em estudo

x_i - i -ésimo valor distinto da variável X ou ponto médio da i -ésima classe

f_i - frequência do i -ésimo valor distinto da variável X ou da i -ésima classe

k - número de valores distintos ou de classes (observado da população ou da amostra)

Média populacional:

Em uma população de tamanho N , a média é dada por:

$$\mu = \frac{\sum_{i=1}^k x_i \times f_i}{N} \quad e \quad \sum_{i=1}^k f_i = N$$

Média da Amostra:

Em uma amostra de tamanho n , a média é dada por:

$$\bar{x} = \frac{\sum_{i=1}^k x_i \times f_i}{n} \quad e \quad \sum_{i=1}^k f_i = n$$

86

Média - Distribuição de Frequência por valores

DISTRIBUIÇÃO DO NÚMERO DE FALHAS EM UM SISTEMA DE COMUNICAÇÃO POR DIA - CIA. FACILCOM
SÃO PAULO - DEZEMBRO DE 2019

NÚMERO DE FALHAS	Nº. de Dias
0	4
1	5
2	7
3	8
4	4
5	2
TOTAL	30

Fonte: Dados Fictícios

x_i	f_i	$x_i \cdot f_i$
0	4	0
1	5	5
2	7	14
3	8	24
4	4	16
5	2	10
TOTAL	30	69

$$n = \sum_{i=1}^6 f_i$$

$$\sum_{i=1}^6 x_i f_i$$

$$\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{n} = \frac{69}{30} = 2,3 \text{ falhas } p / \text{ dia}$$

87

Média - Distribuição de Frequência por classes

CONSUMO MENSAL DE ENERGIA ELÉTRICA
LOCALIDADE "CLARIDADE" - 2019

CONSUMO	Nº. de Usuários
10 — 15	3
15 — 20	5
20 — 25	10
25 — 30	15
30 — 35	8
35 — 40	6
40 — 45	3
TOTAL	50

Fonte: Dados Fictícios

Classes	f_i	x_i	$x_i \cdot f_i$
10 — 15	3	12,5	37,5
15 — 20	5	17,5	87,5
20 — 25	10	22,5	225
25 — 30	15	27,5	412,5
30 — 35	8	32,5	260
35 — 40	6	37,5	225
40 — 45	3	42,5	127,5
TOTAL	50		1375

Ponto Médio

$$n = \sum_{i=1}^7 f_i$$

$$\sum_{i=1}^7 x_i f_i$$

$$\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{n} = \frac{1375}{50} = 27,5 \text{ kwh}$$

88

Variância

X - variável quantitativa em estudo

x_i - i -ésimo valor distinto da variável X ou ponto médio da i -ésima classe

f_i - frequência do i -ésimo valor distinto da variável X ou da i -ésima classe

k - número de valores distintos ou de classes (observado da população ou da amostra)

Variância populacional:

Em uma população de tamanho N :

$$\sigma^2 = \frac{\sum_{i=1}^k x_i^2 f_i - N \mu^2}{N} \quad e \quad N = \sum_{i=1}^k f_i$$

Variância da amostra:

Em uma amostra de tamanho n :

$$S^2 = \frac{\sum_{i=1}^k x_i^2 f_i - n \bar{x}^2}{n-1} \quad e \quad n = \sum_{i=1}^k f_i$$

89



Vamos trabalhar com os dados?

- Considerando os dados referentes as variáveis:
- número de falhas p/dia do sistema de comunicação
 - Consumo de energia elétrica em kwh

apresentados de forma agrupada, determine para cada caso :

- 1.as medidas de variabilidade
- 2.a medidas de assimetria.

Analise os resultados obtidos.

OBS: Baixar o arquivo com o exercício (Medidas para Dados Agrupados) no espaço da disciplina no Aprender 3.

Mediana

◆ Distribuição de frequências por valores

***n ímpar :** – Mediana é valor que ocupa a posição $\frac{n+1}{2}$*

***n par :** – Mediana é a média dos valores que ocupam as posições $\left(\frac{n}{2}\right)$ e $\left(\frac{n}{2}+1\right)$*

Mediana - Distribuição de Frequência por valores

DISTRIBUIÇÃO DO NÚMERO DE FALHAS EM UM SISTEMA DE COMUNICAÇÃO POR DIA -CIA. FACILCOM
SÃO PAULO – DEZEMBRO DE 2019

NÚMERO DE FALHAS	Nº. de Dias
0	4
1	5
2	7
3	8
4	4
5	2
TOTAL	30

Fonte: Dados Fictícios

x_i	f_i	F_i
0	4	4
1	5	9
2	7	16
3	8	24
4	4	28
5	2	30
TOTAL	30	

$$\frac{n}{2} = \frac{30}{2} = 15 \rightarrow \begin{cases} 15^a \text{ posição : } x_{(15)} = 2 \\ 16^a \text{ posição : } x_{(15)} = 2 \end{cases}$$

$$Md = \frac{2+2}{2} = 2 \text{ falhas } p / \text{ dia}$$

Mediana - Distribuição de Frequência por classes

CONSUMO MENSAL DE ENERGIA ELÉTRICA
LOCALIDADE “CLARIDADE” - 2019

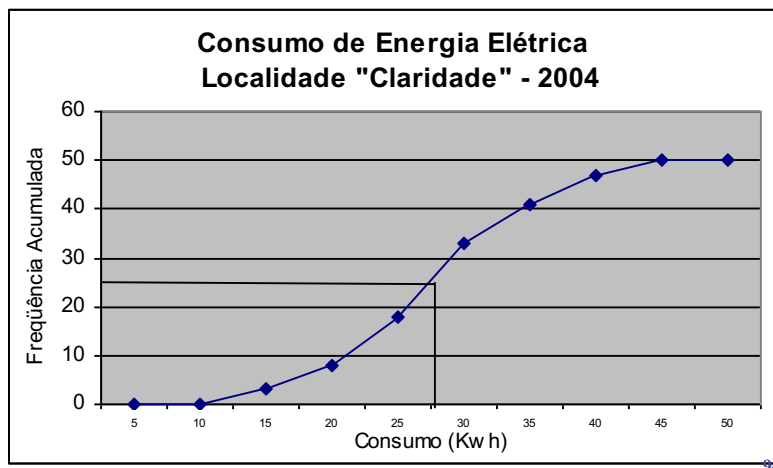
CONSUMO	Nº. de Usuários
10 — 15	3
15 — 20	5
20 — 25	10
25 — 30	15
30 — 35	8
35 — 40	6
40 — 45	3
TOTAL	50

Fonte: Dados Fictícios

Classes	f_i	F_i
10 — 15	3	3
15 — 20	5	8
20 — 25	10	18
25 — 30	15	33
30 — 35	8	41
35 — 40	6	47
40 — 45	3	50
TOTAL	50	

Mediana

◆ Distribuição de frequências por classes



Mediana

◆ Distribuição de frequências por classes

$$Md = l_i + a_i \frac{\left(\frac{n}{2} - F_{i-1}\right)}{f_i}$$

onde :

l_i – limite inferior da classe que contem a mediana

a_i – amplitude da classe que contem a mediana

F_{i-1} – frequência acumulada da classe anterior a que contem a mediana

f_i – frequência absoluta da classe que contem a mediana

Mediana - Distribuição de Frequência por classes

CONSUMO MENSAL DE ENERGIA
ELÉTRICA
LOCALIDADE "CLARIDADE" - 2019

CONSUMO	Nº de Usuários
10 — 15	3
15 — 20	5
20 — 25	10
25 — 30	15
30 — 35	8
35 — 40	6
40 — 45	3
TOTAL	50

Fonte: Dados Fictícios

Classes	f_i	F_i
10 — 15	3	3
15 — 20	5	8
20 — 25	10	18
25 — 30	15	33
30 — 35	8	41
35 — 40	6	47
40 — 45	3	50
TOTAL	50	

$$\frac{n}{2} = \frac{50}{2} = 25 \rightarrow 25^{\text{a}} \text{ posição}$$

$$Md = 25 + 5 \frac{(50/2 - 18)}{15} = 25 + 5 \times \frac{7}{15} = 25 + 2,33 = 27,33 \text{ kWh}$$

Quantis de ordem p

◆ Distribuição de frequências por classes

$$q(p) = l_i + a_i \frac{(n \times p - F_{i-1})}{f_i}$$

onde :

l_i – limite inferior da classe que contém o quantil

a_i – amplitude da classe que contém o quantil

F_{i-1} – frequência acumulada da classe anterior a que contém o quantil

f_i – frequência absoluta da classe que contém o quantil

Moda - Distribuição de Frequência por valores

DISTRIBUIÇÃO DO NÚMERO DE FALHAS EM UM
SISTEMA DE COMUNICAÇÃO POR DIA - CIA. FACILCOM
SÃO PAULO - DEZEMBRO DE 2019

NÚMERO DE FALHAS	Nº. de Dias
0	4
1	5
2	7
3	8
4	4
5	2
TOTAL	30

Fonte: Dados Fictícios

$Mo = 3 \text{ falhas p / dia}$

98

Moda - Distribuição de Frequência por classes

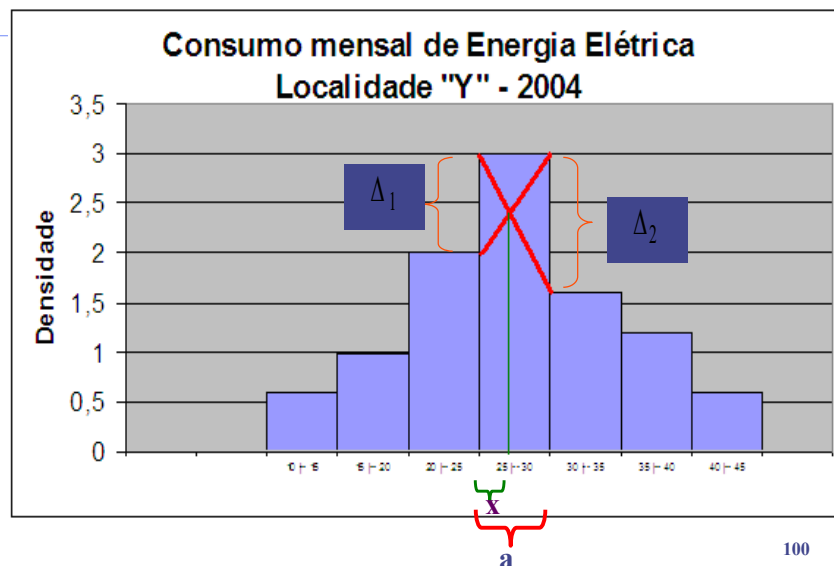
CONSUMO MENSAL DE ENERGIA
ELÉTRICA
LOCALIDADE "CLARIDADE" - 2019

CONSUMO	Nº. de Usuários
10 — 15	3
15 — 20	5
20 — 25	10
25 — 30	15
30 — 35	8
35 — 40	6
40 — 45	3
TOTAL	50

Fonte: Dados Fictícios

99

Moda de Czuber



Moda de Czuber

$$Mo = l_i + a_i \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

onde:

l_i – limite inferior da classe que contém a moda

a_i – amplitude da classe que contém a moda

Δ_1 – frequência ou densidade da classe modal-frequência da classe anterior

Δ_2 – frequência ou densidade da classe modal-frequência da classe posterior

101

Moda - Distribuição de Frequência por classes

CONSUMO MENSAL DE ENERGIA
ELÉTRICA
LOCALIDADE "CLARIDADE" - 2019

Classes	fi	ai	densidade
10 -- 15	3	5	0,6
15 -- 20	5	5	1,0
20 -- 25	10	5	2,0
25 -- 30	15	5	3,0
30 -- 35	8	5	1,6
35 -- 40	6	5	1,2
40 -- 45	3	5	0,6
TOTAL	50		

Classe
modal

$$l_i = 25 \quad a_i = 30 - 25 = 5 \quad \Delta_1 = 3 - 2 = 1$$

$$\Delta_2 = 3 - 1,6 = 1,4$$

$$Mo = l_i + a_i \frac{\Delta_1}{\Delta_1 + \Delta_2} = 25 + 5 \times \frac{1}{1 + 1,4} = 25 + \frac{5 \times 1}{2,4} = 25 + 2,08 = 27,08 kwh$$