



Medidas de Tendência Central





Generalidades

- **Estatística Descritiva:**
 - Resumo ou descrição das características importantes de um conjunto conhecido de dados populacionais
- **Inferência Estatística:**
 - Generalizações sobre uma população tomadas a partir da utilização de dados amostrais.



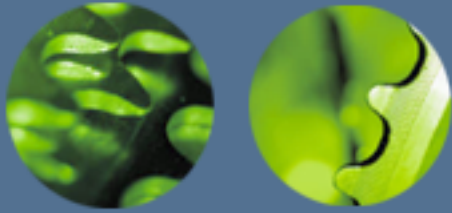
Estatística descritiva

- Através da **ESTATÍSTICA DESCRITIVA** entendemos melhor um conjunto de dados através de suas características.
- As três principais características são:
 - Um valor representativo do conjunto de dados. Ex.: uma média
 - Uma medida de dispersão ou variação.
 - A natureza ou forma da distribuição dos dados: sino, uniforme, assimétrica,...



Medidas de Tendência Central

- Determina valores típicos ou representativos de um conjunto de dados
 - Média
 - Mediana
 - Moda
 - Ponto médio



Média Aritmética ou Média

- Média aritmética ou média
 - Centro do conjunto de dados – ponto de equilíbrio
 - A mais importante medida de tendência central
 - Média = $\Sigma(x)/n$
 - \bar{x} = extraída de uma amostra
 - μ = todos os valores da população são considerados

Análise Estatística da Turma de Prob. e Estatística		
Eventos	Altura	Sexo
Aluno 1	1,72	M
Aluno 2	1,60	F
Aluno 3	1,74	M
Aluno 4	1,88	M
Aluno 5	1,82	M
Aluno 6	1,75	M
Aluno 7	1,82	M
Aluno 8	1,75	M
Aluno 9	1,73	M
Aluno 10	1,75	M
Aluno 11	1,80	M
Aluno 12	1,75	M
Aluno 13	1,73	M
Aluno 14	1,84	M
Aluno 15	1,76	M
Aluno 16	1,78	M
Aluno 17	1,75	M
Aluno 18	1,69	F
Média	1,759	-----



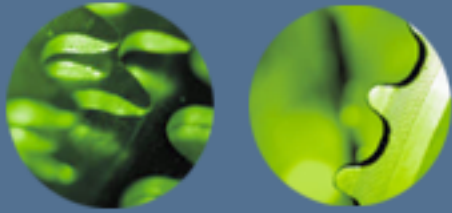
Notação

- Σ - somatório de um conjunto de valores
- x – valores individuais dos dados
- n – número de valores da amostra
- N – número de valores de uma população
- $\bar{x} = \Sigma(x)/n$ – média de um conjunto de valores amostrais
- $\mu = \Sigma(x)/N$ – média de todos os valores de uma população



Mediana (\tilde{x})

- Valor do meio do conjunto de dados, quando os valores estão dispostos em ordem crescente ou decrescente; divide um conjunto de dados em duas partes iguais.
- Para calcular:
 - Disponha os valores em ordem (crescente ou decrescente)
 - Se o número de valores é ímpar, a mediana é o número localizado no meio da lista
 - Se o número é par, a mediana é a média dos dois valores do meio



Do nosso conjunto de dados...

- Encontre a mediana:
 - Liste em ordem crescente os valores
 - Encontre a posição da mediana: $(n+1)/2$
 - Se n é ímpar, mediana é o número da posição;
 - Se n é par, mediana é a média entre os dois números em torna da posição.
- No exemplo:
 - $n=18$ (par);
 - Posição: $(n+1)/2 = 9,5$
 - Mediana \Rightarrow média entre o 9º e o 10º valor = $(1,75+1,75)/2 = 1,75$

Análise Estatística da Turma	
Altura	Sexo
1,60	M
1,69	F
1,72	M
1,73	M
1,73	M
1,74	M
1,75	M
1,75	M
1,75	M
1,75	M
1,75	M
1,76	M
1,78	M
1,80	M
1,82	M
1,82	M
1,84	M
1,88	F



Moda (M)

- É o valor que ocorre com maior frequência.
- Quando dois valores ocorrem com a mesma frequência, cada um deles é chamado de uma moda, e o conjunto se diz BIMODAL
- Se mais de dois valores ocorrem com a mesma frequência máxima, cada um deles é uma moda e o conjunto é MULTIMODAL.
- Quando nenhum valor é repetido o conjunto não tem moda



Do nosso conjunto de dados...

- Moda => 1,75

Análise Estatística da Turma	
Altura	No. ocorrências
1,60	1
1,69	1
1,72	1
1,73	2
1,74	1
1,75	5
1,76	1
1,78	1
1,80	1
1,82	2
1,84	1
1,88	1



Ponto Médio

- Valor que está a meio caminho entre o maior e o menor valor

$$\text{ponto médio} = \frac{\text{maior valor} + \text{menor valor}}{2}$$



Do nosso conjunto de dados...

$$\text{ponto médio} = \frac{1,88 + 1,60}{2} = 1,74$$

Análise Estatística da Turma	
Altura	Sexo
1,60	M
1,69	F
1,72	M
1,73	M
1,73	M
1,74	M
1,75	M
1,75	M
1,75	M
1,75	M
1,75	M
1,75	M
1,75	M
1,76	M
1,78	M
1,80	M
1,82	M
1,82	M
1,84	M
1,88	F



Medidas de Posição

- Seja o seguinte conjunto de valores:

5 7 8 10 12 15 20

média = 11

mediana = 10

ponto médio = 12,5

- Se alterarmos significativamente o último valor:

5 7 8 10 12 15 200

média = 36,7!!

mediana = 10

ponto médio = 102,5!!



Medidas de Posição

- Devemos ter cuidados ao escolhermos uma medida de posição para representar um conjunto de dados, pois:
 - “*Média*” e “*Ponto Médio*” são muito afetados por valores extremo
- *Em geral*, a melhor política é utilizar os dois parâmetros: “*média*” e “*mediana*”
 - Valores de “*Média*” e “*Mediana*” muito próximos é uma indicação que o conjunto de valores é razoavelmente simétrico em relação à posição central (*média* / *mediana*)



Média Ponderada

- Cálculo da média, atribuindo pesos diferentes para cada valor

$$\bar{x} = \frac{x_1 \cdot w_1 + x_2 \cdot w_2 + \dots + x_n \cdot w_n}{\sum w}$$

x_n = valores individuais

w_n = pesos individuais



Exercício

- A tabela de frequência a seguir resume os tempos gastos em estudo extra-classe por calouros em uma universidade americana. Faça uma estimativa do tempo médio de estudo deste grupo.

Horas de Estudo	Frequência (alunos)
0	5
1-5	96
6-10	57
11-15	25
16-20	11
+ de 20	6



Exercícios

Dados os conjuntos de dados abaixo, calcule a MÉDIA, a MEDIANA, a(s) MODA(S) e o PONTO MÉDIO.

1. 15; 26; 28; 10; 29
2. 500; 1000; 50000; 800; 500; 600
3. Altura dos alunos da sala
4. Limite de Resistência de um vergalhão kgf/mm^2
(55; 52; 55; 53; 56; 59; 58; 55; 56; 53; 52; 51; 54;
54; 55; 58; 57; 57; 56; 49; 54; 56)



Separatrizes

- Percentis
 - Decis
 - Quartis
 - Escore padronizado ou escore z
- DADOS ORDENADOS**



Percentis

- Um *percentil* indica que há $x\%$ de dados inferiores
- Ou seja, os percentis dividem o conjunto de dados em 100 partes iguais.
- Há, portanto, 99 percentis
- Ex.: o P_{92} (92º percentil) indica que há 92% de dados inferiores.



Percentis

- Dado o conjunto de valores, calcule o 25º percentil (valores devem estar *ordenados*)

72	74	75	77	78	79	82	85	86	90	93	94
----	----	----	----	----	----	----	----	----	----	----	----

- Calcule: $L = (k/100).n$
 - L – posição do percentil desejado
 - K - percentil desejado
 - n – número de valores



Percentis

L é número inteiro?

SIM

o k-ésimo percentil está a meio caminho do L-ésimo valor e o próximo valor do conjunto de dados.

- $n=12 \Rightarrow L=3$
- $(L_3 + L_4)/2 = (75+77)/2 = 76$

NÃO

Modificar L, arredondando seu valor para o inteiro maior mais próximo.

Ex.: se $n=11 \Rightarrow L=2,75$

- A posição do k-ésimo percentil será 3.
- Retirando o valor 77 do conjunto de dados...
- $P_{25} = 75$



Calcular Percentil

- Para calcular a qual percentil pertence um dado valor

$$\text{percentil do valor } x = \frac{\text{número valores inferiores a } x}{\text{número total de valores}} \cdot 100$$



Processo Alternativo

- Seja o seguinte conjunto de “ n ” valores:

10 7 15 12 20 8 5 25

- ORDENA-SE o conjunto de “ n ” valores

5 7 8 10 12 15 20 25

$$L_p = (n-1) \times p/100 + 1$$

$$L_{25} = (8-1) \times 0,25 + 1 = 2,75 \quad L_{75} = (8-1) \times 0,75 + 1 = 6,25$$

$$P_1 = 7 + (8-7) \times 0,75 = 7,75 \quad P_3 = 15 + (20-15) \times 0,25 = 16,25$$



Decis

- Dividem o conjunto de dados em 10 partes iguais de 10%.
- Portanto há 09 decis numa distribuição, que a divide em 10 grupos de 10% de dados.
- Os Decis são os percentis:
 - P_{10} , P_{20} , P_{30} , P_{40} , P_{50} , P_{60} , P_{70} , P_{80} , P_{90}



Decis

$D_1 = P_{10}$	$D_2 = P_{20}$	$D_3 = P_{30}$
$D_4 = P_{40}$	$D_5 = P_{50}$	$D_6 = P_{60}$
$D_7 = P_{70}$	$D_8 = P_{80}$	$D_9 = P_{90}$



Quartis

- Os *quartis* dividem a distribuição em quatro partes iguais de 25%
- Temos assim 3 quartis numa distribuição
 - O 1º quartil Q_1 , separa os 25% de dados inferiores
 - O 2º quartil Q_2 , separa os 50% de dados inferiores
 - O 3º quartil Q_3 , separa os 75% de dados inferiores.



Quartis

- Os Quartis são os percentis:
 - P_{25} , P_{50} , P_{75}

$Q_1 = P_{25}$	$Q_2 = P_{50}$	$Q_3 = P_{75}$
----------------	----------------	----------------

- Observa-se que $Q_2 \Rightarrow$ MEDIANA



Outras medidas descritivas

- Baseado nos conceitos de Percentis, Decis e Quartis, podemos definir outras descrições estatísticas:
 - Intervalo interquartil $\rightarrow Q_3 - Q_1$
 - Intervalo semi-interquartil $\rightarrow (Q_3 - Q_1)/2$
 - Quartil médio $\rightarrow (Q_3 + Q_1)/2$
 - Amplitude de percentis
 - Ex. 10-90 $\Rightarrow P_{90} - P_{10}$



Escore padronizado

$$z = \frac{x - \bar{x}}{s}$$

$$z = \frac{x - \mu}{\sigma}$$

- Número de desvios-padrão pelo qual um valor dista da média (para mais ou para menos)