ESTATÍSTICA

A Estatística, ramo da Matemática Aplicada, teve origem na história do homem. Desde a Antiguidade, vários povos registravam o número de habitantes, de nascimentos, de óbitos, distribuíam equitativamente terras ao povo. Na Idade Média colhiam-se informações, geralmente com finalidades tributárias ou bélicas. No início do século XVIII o estudo de tais fatos foi adquirindo feição verdadeiramente científica. Schmetizel foi o primeiro a utilizar a palavra em latim no séculoXVII.

Adiante, a palavra Estatística, que etimologicamente deriva da palavra Estado, foi incorporada pelo acadêmico alemão Godofredo Achenwall.

ESTATÍSTICA

As tabelas tornaram-se mais completas, surgiram as representações gráficas e o cálculo das probabilidades, e a Estatística deixou de ser uma simples catalogação de dados numéricos para se tornar o estudo de “como chegar a conclusões sobre o todo (população), partindo da observação de partes do todo (amostras)”.

Podemos dizer, então, que a Estatística é uma parte da Matemática Aplicada que fornece métodos para a coleta, organização, descrição, análise e interpretação de dados para a utilização dos mesmos na tomada de decisões.

ESTATÍSTICA

As tabelas tornaram-se mais completas, surgiram as representações gráficas e o cálculo das probabilidades, e a Estatística deixou de ser uma simples catalogação de dados numéricos para se tornar o estudo de “como chegar a conclusões sobre o todo (população), partindo da observação de partes do todo (amostras)”.

Podemos dizer, então, que a Estatística é uma parte da Matemática Aplicada que fornece métodos para a coleta, organização, descrição, análise e interpretação de dados para a utilização dos mesmos na tomada de decisões.

ESTATÍSTICA

- A coleta, organização e a descrição dos dados estão a cargo da **Estatística Descritiva**.

- A análise e a interpretação dos dados ficam a cargo da **Estatística Inferencial**.

$$\boxed{\begin{array}{c} \overline{X} \rightarrow \textit{média amostral (estimador)} \\ \mu \rightarrow \textit{média populacional (parâmetro)} \end{array}}$$

ESTATÍSTICA

$$\overline{X} \rightarrow média\ amostral\ (estimador)$$

$$\mu \rightarrow média\ populacional\ (parâmetro)$$

O parâmetro populacional é, em geral, um valor fixo. A estimativa depende da amostra. Se pegarmos uma amostra diferente, obteremos estimativas diferentes.

EXERCÍCIO 01

O diretor de um sistema penitenciário, com o propósito de estimar o percentual de detentos que possuem filhos, entregou a um analista um cadastro com os nomes de 500 detentos da instituição para que esse profissional realizasse entrevistas com os indivíduos selecionados.

A partir dessa situação hipotética e dos múltiplos aspectos a ela relacionados, julgue o item, referente a técnicas de amostragem.

A diferença entre um censo e uma amostra consiste no fato de esta última exigir a realização de um número maior de entrevistas.

EXERCÍCIO 01 - Resolução

O diretor de um sistema penitenciário, com o propósito de estimar o percentual de detentos que possuem filhos, entregou a um analista um cadastro com os nomes de 500 detentos da instituição para que esse profissional realizasse entrevistas com os indivíduos selecionados.

A partir dessa situação hipotética e dos múltiplos aspectos a ela relacionados, julgue o item, referente a técnicas de amostragem.

A diferença entre um censo e uma amostra consiste no fato de esta última exigir a realização de um número maior de entrevistas.

Errado. Realizamos um censo quando obtemos as informações de todos os elementos da população. Trabalhamos com uma amostra quando obtemos informações de um subconjunto da população. Portanto, precisamos realizar mais entrevistas no censo.

DISTRIBUIÇÕES DE FREQUÊNCIAS

À primeira fase de um processo estatístico -> organização!

Imagine que você foi o encarregado para fazer uma pesquisa sobre a altura dos alunos de Estatística.

Como são muitos alunos, você decidiu realizar uma pesquisa com apenas 40 alunos. Suponhamos termos feito uma coleta de dados relativos às estaturas de quarenta alunos, em centímetros, que compõem uma amostra dos alunos, resultando a seguinte tabela ao lado

166	**160**	**161**	**150**	**162**	**160**	**165**	**167**
164	160	162	161	168	163	156	**173**
160	155	164	168	155	152	163	**160**
155	155	169	151	170	164	154	**161**
156	**172**	**153**	**157**	**156**	**158**	**158**	**161**

DISTRIBUIÇÕES DE FREQUÊNCIAS

Obviamente, quando você começa a sua pesquisa, os seus dados não estão organizados. A esses dados desorganizados denominamos dados brutos.

O próximo passo, após realizar a coleta dos dados, é **organizar** esses dados em ordem crescente ou decrescente. Denominamos os dados dispostos em ordem crescente ou decrescente de **rol.**

166	**160**	**161**	**150**	**162**	**160**	**165**	**167**
164	160	162	161	168	163	156	**173**
160	155	164	168	155	152	163	**160**
155	155	169	151	170	164	154	**161**
156	**172**	**153**	**157**	**156**	**158**	**158**	**161**

DISTRIBUIÇÕES DE FREQUÊNCIAS

À diferença entre o maior e o menor número do rol chama-se **amplitude total** dos dados.

173 – 150 = 23 cm.

166	160	161	150	162	160	165	167
164	160	162	161	168	163	156	173
160	155	164	168	155	152	163	160
155	155	169	151	170	164	154	161
156	172	153	157	156	158	158	161

150	155	156	160	161	162	164	168
151	155	156	160	161	163	165	169
152	155	157	160	161	163	166	170
153	155	158	160	161	164	167	172
154	156	158	160	162	164	168	173

EXERCÍCIO 02

Verifique os conjuntos A, B, C e D abaixo, no formato de rol e assinale a alternativa correta.

A	3	2	5	1	0,5	0,05	5,1
B	0,25	1	3	7	10,35	10	0,5
C	0,01	0,02	0,03	0,04	0,05	0,06	0,07
D	5	4	3	3	2	2	1

a)A amplitude total do conjunto C é igual a 0,28.

b)Não é possível calcular a amplitude total do conjunto D, pois estamos diante de um rol decrescente.

c)A amplitude de todos os conjuntos é igual a 7.

d)A amplitude total do conjunto A é 2,1.

e)A amplitude total do conjunto B é o dobro da amplitude total do conjunto A.

EXERCÍCIO 02 - Resolução

O primeiro passo é organizar os conjuntos A, B, C e D em formato de rol

Rol							
A	0,05	0,5	1	2	3	5	**5,1**
B	0,25	0,5	1	3	7	10	**10,35**
C	0,01	0,02	0,03	0,04	0,05	0,06	**0,07**
D	**1**	**2**	**2**	**3**	**3**	**4**	**5**

A amplitude total do conjunto B é igual a 10,35 – 0,25 = 10,1.

Portanto, a amplitude total do conjunto B é o dobro da amplitude total do conjunto A e a alternativa E é **verdadeira.**

ELEMENTOS DE UMA DISTRIBUIÇÃO DE FREQUÊNCIAS

Estaturas (cm)	Frequência
150⊢154	4
154⊢158	9
158⊢162	11
162⊢166	8
166⊢170	5
170⊢174	3
Total	40

Classe é cada um dos grupos ou intervalos obtidos a partir do agrupamento ou conjunto de dados.

Por exemplo, a terceira classe é 158 ⊢ 162.

Denominamos limites de classe os extremos de cada classe. O menor número é o limite inferior da classe e o maior é o limite superior.

ELEMENTOS DE UMA DISTRIBUIÇÃO DE FREQUÊNCIAS

Estaturas (cm)	Frequência
150⊢154	4
154⊢158	9
158⊢162	11
162⊢166	8
166⊢170	5
170⊢174	3
Total	40

É importante notar que nos preocupamos com os extremos da classe (se o intervalo é aberto ou fechado) apenas para montar a tabela. No cálculo de medidas de posição e dispersão (média, mediana, desvio padrão, etc), não nos importamos com os extremos das classes. Como as tabelas já virão prontas na prova, você não precisa se preocupar em saber se o intervalo é aberto ou fechado.

ELEMENTOS DE UMA DISTRIBUIÇÃO DE FREQUÊNCIAS

Estaturas (cm)	Frequência
150⊢154	4
154⊢158	9
158⊢162	11
162⊢166	8
166⊢170	5
170⊢174	3
Total	40

AMPLITUDE DE UM INTERVALO DE CLASSE

Amplitude de um intervalo de classe ou, simplesmente, intervalo de classe é a medida do intervalo que define a classe. É obtida pela diferença entre os limites superior e inferior dessa classe e designamos por h

$$\boxed{h = l_{sup} - l_{inf}}$$

ELEMENTOS DE UMA DISTRIBUIÇÃO DE FREQUÊNCIAS

Estaturas (cm)	Frequência
150⊢154	4
154⊢158	9
158⊢162	11
162⊢166	8
166⊢170	5
170⊢174	3
Total	40

AMPLITUDE TOTAL DA DISTRIBUIÇÃO

Amplitude total da distribuição (AT) é a diferença entre o limite superior da última classe (limite superior máximo) e o limite inferior da primeira classe (limite inferior mínimo).

$$A_T = 174 - 150 = 24\ cm$$

TIPOS DE FREQUÊNCIA

Estaturas (cm)	Frequência
150⊢154	4
154⊢158	9
158⊢162	11
162⊢166	8
166⊢170	5
170⊢174	3
Total	40

FREQUÊNCIA ABSOLUTA SIMPLES

São os valores que realmente representam o número de dados de cada classe. A soma das frequências simples é igual ao número total dos dados.

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

$$\sum_{i=1}^{6} f_i = f_1 + f_2 + f_3 + f_4 + f_5 + f_6 = 4 + 9 + 11 + 8 + 5 + 3 = 40$$

TIPOS DE FREQUÊNCIA

Estaturas (cm)	Frequência
150⊢154	4
154⊢158	9
158⊢162	11
162⊢166	8
166⊢170	5
170⊢174	3
Total	40

FREQUÊNCIA RELATIVA SIMPLES

São os valores das razões entre as frequências simples e a frequência total, normalmente expressas em porcentagem.

$$fr_i = \frac{f_i}{n}$$

$$fr_3 = \frac{f_3}{n} = \frac{11}{40} = 0{,}275$$

TIPOS DE FREQUÊNCIA

FREQUÊNCIAS ACUMULADAS

Estaturas (cm)	Frequência	fac
150⊢154	4	4
154⊢158	9	13
158⊢162	11	24
162⊢166	8	32
166⊢170	5	37
170⊢174	3	40
Total	**40**	

Para obter as frequências acumuladas, podemos utilizar frequências absolutas ou relativas. O procedimento é o mesmo.

$$fr_i = \frac{f_i}{n}$$

$$fr_3 = \frac{f_3}{n} = \frac{11}{40} = 0,275$$

TIPOS DE FREQUÊNCIA

Estaturas (cm)	Frequência	$d_i = \frac{f_i}{h}$
150⊢154	4	$\frac{4}{4} = 1$
154⊢158	9	$\frac{9}{4} = 2{,}25$
158⊢162	11	$\frac{11}{4} = 2{,}75$
162⊢166	8	$\frac{8}{4} = 2$
166⊢170	5	$\frac{5}{4} = 1{,}25$
170⊢174	3	$\frac{3}{4} = 0{,}75$
Total	**40**	

DENSIDADE DE FREQUÊNCIA

Densidade de frequência de uma classe é a razão (quociente) entre a frequência da classe (absoluta ou relativa) e a amplitude da classe.

$$Densidade\ de\ frequência = \frac{frequência}{amplitude}$$

MÉDIAS

MÉDIA ARITMÉTICA SIMPLES

$$\bar{x} = \frac{\text{SOMA DOS TERMOS}}{\text{NÚMERO DE TERMOS}}$$

EX.: MÉDIA ARITMÉTICA SIMPLES DOS NÚMEROS $\underbrace{3, 5, 9, 2, 11}_{\text{5 TERMOS } (m)}$:

$$\bar{x} = \frac{3 + 5 + 9 + 2 + 11}{5}$$

$$= \frac{30}{5} \rightarrow \boxed{\bar{x} = 6}$$

SE TODOS OS NÚMEROS FOREM SUBSTITUÍDOS POR $\bar{x}$, A SOMA DOS TERMOS SERÁ PRESERVADA.

$\rightarrow$ SOMA $= \bar{x} . m$

PROPRIEDADES DA MÉDIA ARITMÉTICA

- SEMPRE EXISTE E É ÚNICA
- MENOR NÚMERO DO CONJUNTO $\leq \bar{x} \leq$ MAIOR NÚMERO DO CONJUNTO
- A SOMA DOS DESVIOS EM RELAÇÃO À MÉDIA É NULA.
- A SOMA DO QUADRADO DOS DESVIOS EM RELAÇÃO À MÉDIA É MÍNIMA.
- SOMANDO-SE UMA CONSTANTE C A TODOS OS NÚMEROS, A NOVA MÉDIA ($\bar{x}'$) SERÁ $\bar{x}' = \bar{x} + c$ (O EQUIVALENTE P/ SUBTRAÇÃO)
- MULTIPLICANDO-SE TODOS OS NÚMEROS POR UMA CONSTANTE C, A NOVA MÉDIA ($\bar{x}'$) SERÁ $\bar{x}' = \bar{x} . c$ (O EQUIVALENTE P/ DIVISÃO)

USE P/ SIMPLIFICAR OS TERMOS ANTES DE CALCULAR AS MÉDIAS

MÉDIA ARITMÉTICA PONDERADA

- COMO A SIMPLES, MAS OS ELEMENTOS (x_i) PODEM TER PESOS DIFERENTES (p)

(COMO EM UMA PROVA, EM QUE AS QUESTÕES DE UMA MATÉRIA VALE MAIS QUE DE OUTRA)

$$\bar{x}_p = \frac{\text{SOMA DOS TERMOS MULTIPLICADOS PELOS RESPECTIVOS PESOS}}{\text{SOMA DOS PESOS}}$$

EX.: MÉDIA ARITMÉTICA PONDERADA DOS SEGUINTES NÚMEROS E SEUS PESOS:

3, PESO 2
4, PESO 1
2, PESO 5

$$\bar{x}_p = \frac{3.2 + 4.1 + 2.5}{2 + 1 + 5}$$

$$= \frac{6 + 4 + 10}{8}$$

$$= \frac{20}{8}$$

$$= \boxed{\bar{x}_p = 2.5}$$

MÉDIAS

MÉDIA P/ DADOS AGRUPADOS POR VALOR

- = MESMA IDÉIA DA MÉDIA PONDERADA
- NORMALMENTE P/ DADOS DISCRETOS.
- EXEMPLO: IDADE DOS ALUNOS DE UMA ESCOLA.

VALOR · NÚMERO DE OCORRÊNCIAS · VOCÊ CALCULA! (COLUNA AUXILIAR)

IDADE (x_i)	FREQUÊNCIA (f_i)	$x_i \cdot f_i$
25	21	525 = 25 . 21
30	47	1.410
34	54	1.836
38	41	1.558
41	37	1.517
TOTAL:	200	6.846

(= TOTAL DE ALUNOS)

$$\overline{X} = \frac{\sum x_i \cdot f_i}{\sum f_i}$$

$$= \frac{6.846}{200} \rightarrow \boxed{\overline{X} = 34{,}23 \text{ ANOS}}$$

MÉDIA P/ DADOS AGRUPADOS POR CLASSE

- NORMALMENTE P/ DADOS CONTÍNUOS.
- EXEMPLO: ALTURA DOS ALUNOS DE UMA ESCOLA.

CLASSE · NÚMERO DE OCORRÊNCIAS · VOCÊ CALCULA!

ALTURA	FREQUÊNCIA (f_i)	PONTO MÉDIO (x_i)	$x_i \cdot f_i$
150 - 154	4	152	608 = 152 . 4
154 - 158	9	156	1.404
158 - 162	11	160	1.760
162 - 166	8	164	1.312
166 - 170	5	168	840
170 - 174	3	172	516
TOTAL:	40		6.440

(= TOTAL DE ALUNOS)

$$\overline{X} = \frac{\sum x_i \cdot f_i}{\sum f_i}$$

$$= \frac{6.440}{40} \rightarrow \boxed{\overline{X} = 161 \text{ cm}}$$

CÁLCULO DO PONTO MÉDIO (x_i):

$$x_i = \frac{\text{LIMITE INFERIOR DA CLASSE} + \text{LIMITE SUPERIOR DA CLASSE}}{2}$$

EX.: $x_1 = \frac{150 + 154}{2} = 152$

MÉDIAS

MÉDIA GEOMÉTRICA

- RAIZ n-ÉSIMA DO PRODUTO DOS TERMOS (n = NÚMERO DE TERMOS)

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$$

EX.: MÉDIA GEOMÉTRICA DOS TERMOS $\underbrace{3, 8, 9}_{\text{3 TERMOS}}$: 3 TERMOS ($n$)

$$G = \sqrt[3]{3 \cdot 8 \cdot 9} = \sqrt[3]{216} =$$

$$\boxed{G = 6}$$

SE TODOS OS NÚMEROS FOREM SUBSTITUÍDOS POR G, O PRODUTO DOS TERMOS SERÁ PRESERVADO.

$$\text{PRODUTO DOS TERMOS} = \underbrace{G \cdot G \cdot \ldots \cdot G}_{n}$$

MÉDIA HARMÔNICA

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

- = INVERSO DA MÉDIA ARITMÉTICA DOS INVERSOS: (FÓRMULA ALTERNATIVA)

$$H = \left(\frac{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}{n} \right)^{-1}$$

SE TODOS OS NÚMEROS FOREM SUBSTITUÍDOS POR $\overline{H}$, A SOMA DOS INVERSOS DOS TERMOS SERÁ PRESERVADA.

$$\text{SOMA DOS INVERSOS} = \frac{n}{\overline{H}}$$

EX.: MÉDIA HARMÔNICA DOS TERMOS $\underbrace{3, 4, 9}_{\text{3 TERMOS}}$: 3 TERMOS ($n$)

$$H = \frac{3}{\frac{1}{3} + \frac{1}{4} + \frac{1}{9}}$$

$$= \frac{3}{\frac{12 + 9 + 4}{36}} = \frac{3}{\frac{25}{36}} = \frac{108}{25}$$

$$\boxed{H = 4,32}$$

DESIGUALDADE DAS MÉDIAS

CAI MUITO EM PROVA!

- P/ UMA SEQUÊNCIA DE NÚMEROS POSITIVOS:

$$\overline{X} \geq G \geq H$$

(SÓ É IGUAL QUANDO TODOS OS NÚMEROS FOREM IGUAIS)