CONCEITOS BÁSICOS

1 – Estatística

É uma **metodologia** ou conjunto de técnicas que utiliza a coleta de dados, sua classificação, sua apresentação ou representação, sua análise e sua interpretação visando a sua utilização dentro de um processo decisório.

2 – Estatísticas

É um **conjunto ou coleção de dados numéricos** que fornecem informações acerca de um fenômeno observado: estatísticas econômicas (que fornecem dados sobre preços, vendas, produção, etc.), estatísticas demográficas (que fornecem dados sobre nascimentos, mortes, etc.).

3 – Estatística Descritiva

É a parte da Estatística que procura **descrever** e **analisar** um certo grupo de observações, normalmente denominado de amostra, procurando expressar estas observações através de medidas e formas de representação (tabelas, gráficos, curvas, etc.). É também denominada **Estatística Dedutiva**.

4 – Estatística Inferencial

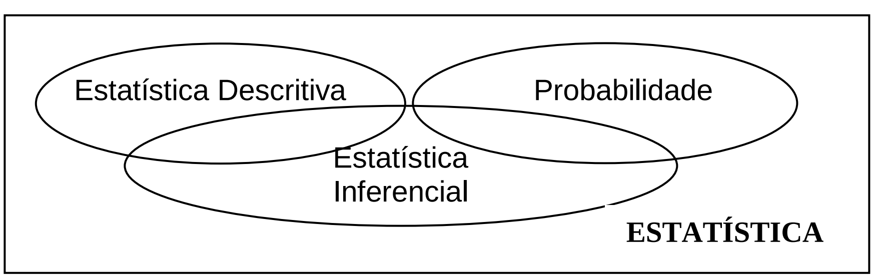
É a parte da Estatística que compreende um processo de generalização, a partir da análise e interpretação de dados amostrais. É também denominada **Estatística Indutiva.**

5 – Teoria da Probabilidade

Como todo processo de generalização envolve uma margem de risco ou incerteza, temos então a parte da Estatística que se ocupa em estudá-la utilizando métodos e técnicas apropriadas, o que constitui os fundamentos da **Teoria da probabilidade.**

6 – Ramos da Estatística

Podemos assim caracterizar três áreas de interesse (ramos) da Estatística:

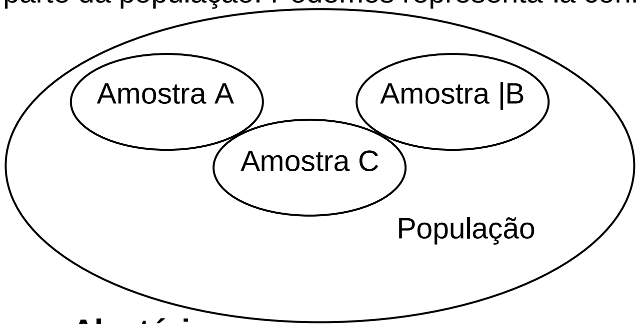


7 - População ou Universo Estatístico

É o maior conjunto tomado como referência na observação de um fenômeno. Pode ser **finita** ou **infinita.** Por exemplo, a produção diária de uma fábrica de determinados componentes (população finita) ou o conjunto de todos os resultados (cara ou coroa) em seguidos lançamentos de uma moeda não viciada (população infinita).

8 – Amostra

É qualquer **subconjunto** não vazio de uma população, excetuando-se a própria população. É, portanto, uma parte da população. Podemos representá-la conforme a figura a seguir:



9 – Amostragem Aleatória

É um plano ou técnica de amostragem, onde cada elemento de uma população tem a mesma probabilidade de ser incluído na amostra.

Uma técnica de obtenção de uma amostragem aleatória é o **sorteio** dos elementos. A fim de garantir uma proteção maior contra a tendência ou vício, utiliza-se atualmente tabelas de **números aleatórios**, que são geradas a partir de algoritmos em computadores.

10 – Experimento Aleatório

É uma experiência ou observação estatística de um fenômeno qualquer em estudo. Esta experiência pode ser, por exemplo:

- lançar um dado e observar o número obtido;
- tirar uma carta de um baralho e observar o resultado;
- retirar uma bola de uma sacola e observar sua cor;
- anotar o número de chamadas telefônicas em uma central telefônica, em um determinado horário;
- registrar o número de acidentes de trânsito em uma determinada rodovia, em um determinado período de tempo.

Uma característica importante de um experimento aleatório é a sua **possibilidade de repetição** contínua, observando-se as mesmas condições de experimentação.

A outra característica importante é **aleatoriedade** do experimento, não podendo existir vícios ou tendências na obtenção de resultados.

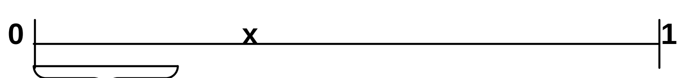
11 – Dados Contínuos

Neste tipo de dados existem **variáveis** que podem assumir qualquer valor num intervalo de valores, inclusive valores quebrados. Normalmente, associamos a estes dados a característica de **medidas.**

Exemplo: altura, peso, comprimento, temperatura, etc.

Desta maneira, podemos dizer que as **variáveis contínuas** podem assumir qualquer valor num intervalo contínuo.

INTERVALO CONTÍNUO



A variável contínua **x** pode assumir qualquer valor no intervalo considerado.

12 – Dados Discretos

Neste tipo de dados existem **variáveis** que só podem assumir valores inteiros num intervalo de valores. Normalmente, associamos a estes dados a característica de **contagem.**

Exemplos: número de livros em uma biblioteca, número de alunos numa sala de aula, etc.

13 – Dados Nominais (ou Categóricos)

Estes dados ocorrem quando são definidas "categorias" tais como sexo (masculino ou feminino), cor dos olhos (pretos, azuis, etc.), campo de estudo (Engenharia, Direito, Economia, etc.). É comum associar-se a estes dados variáveis não-quantitativas e sim (qualitativas) que são denominadas **variáveis nominais**. Antes de serem processadas estatisticamente devem ser atribuídos valores numéricos a tais variáveis.

14 – Dados por Postos

São os dados que, de um modo geral, são sujeitos a avaliações subjetivas quanto à preferência ou desempenho em um conjunto de observações.

Exemplos: competições de beleza, concursos de modas, concursos de canto, etc.

As **variáveis** utilizadas podem receber valores, normalmente, inteiros como, 1º, 2º, 3º, etc... ou então os atributos o mais, o menos, o melhor, o pior, etc.

15 – Método Estatístico

É uma técnica utilizada, normalmente, na área da Estatística Descritiva que visa estruturar e organizar as fases ou etapas que devem ser estabelecidas na abordagem de uma observação estatística. Suas fases ou etapas principais são:

- definição do problema;
- planejamento;
- coleta de dados;
- apuração dos dados;
- apresentação dos dados;
- análise e interpretação dos dados.

16 – Séries Estatísticas

Podemos dizer que uma série estatística é um conjunto de dados estatísticos referenciados aos seguintes fatores: tempo, local e fenômeno.

As séries estatísticas mais comuns são:

- A SÉRIE TEMPORAL, CRONOLÓGICA, HISTÓRICA OU EVOLUTIVA
--→ VARIA O TEMPO
- A SÉRIE GEOGRÁFICA, TERRITORIAL OU ESPACIAL
--→ VARIA O LOCAL
- A SÉRIE ESPECÍFICA OU ESPECIFICATIVA
--→ VARIA O FENÔMENO

Observação:

Existem ainda outros tipos de séries estatísticas, tais como:

a) séries mistas → onde variam mais de um dos fatores citados (tempo, local e fenômeno);
b) distribuições de freqüência → que é uma representação tabular de grande utilização na Estatística, a qual será estudada detalhadamente mais adiante;
c) seqüências de dados numéricos ou quadros com valores numéricos → os quais serão abordados em exercício no decorrer do curso.

TESTES E EXERCÍCIOS

Assinale a alternativa correta.

1. População ou universo é:

a) conjunto de pessoas;
b) conjunto de indivíduos apresentando uma característica especial;
c) conjunto de todos os indivíduos apresentando uma característica comum objeto de estudo.

2. A série Estatística é chamada cronológica quando:

a) o elemento variável é o tempo;
b) o elemento variável é o local;
c) o elemento variável é o fenômeno
d) não tem elemento variável.

3. Estabelecer quais dos dados seguintes são discretos e quais são contínuos:

a) número de ações vendidas diariamente na Bolsa de Valores; **(Discreto)**
b) temperaturas registradas em um posto de meteorologia; **(Contínuos)**
c) vida média das válvulas de televisão produzidas por uma determinada companhia; **(Contínuos)**
d) salários anuais de professores do colégio; **(Contínuos)**
e) comprimentos de 1000 parafusos produzidos pro uma fábrica. **(Contínuos)**

4. Dê o domínio de cada uma das seguintes variáveis e diga se são contínuas ou discretas:

a) quantidade G de litros de água numa máquina de lavar roupa;
R: Domínio: 0 a G → Contínua
b) número B de livros em uma estante de biblioteca;
R: Domínio: 0 a B → Discreta
c) Soma S de pontos obtidos ao lançar um par de dados.
R: Domínio: 2 a 12 → Discreta.

5. Considerando-se a tabela a seguir indicada, pode-se concluir que seus dados refletem uma série:

1972

PRODUTOS	QUANTIDADES (ton.)
CAFÉ	400.000
AÇÚCAR	200.000
MILHO	100.000
FEIJÃO	20.000

Fonte: Dados Fictícios.

a) temporal; → Varia o Tempo
b) geográfica; → Varia o Local
c) especificativa ou específica; → Varia o Fenômeno
d) evolutiva. → Varia o Tempo

10. (AFC – 1994) A tabela abaixo apresenta a distribuição de um grupo de 200 estudantes segundo o curso que fazem (Estatística ou Economia) e o sexo (homem ou mulher).

	HOMEM	MULHER
ESTATÍSTICA	40	20
ECONOMIA	80	60

A única afirmação **errada** é:

a) 60% dos estudantes são homens.
b) 75% das mulheres fazem o curso de Economia.
c) Dois em cada três estudantes de Estatística são homens.
d) Um em cada três homens faz o curso de Estatística.
e) 40% dos homens estudam Economia.

12. (TCDF – 1995) Assinale a opção correta.

a) Em Estatística, entende-se por população um conjunto de pessoas.
b) A variável é discreta quando pode assumir qualquer valor dentro de determinado intervalo.
c) Freqüência relativa de uma variável aleatória é o número de repetições dessa variável.
d) A série estatística é cronológica quando o elemento variável é o tempo.
e) Amplitude Total ou Range é a diferença entre dois valores quaisquer do atributo.

DISTRIBUIÇÃO DE FREQÜÊNCIAS

A representação tabular é uma das modalidades mais utilizadas para a apresentação dos dados estatísticos coletados na amostragem. Para isto, nós, freqüentemente, listamos os valores encontrados na amostra. Uma classificação metodológica usual é verificar a natureza dos dados estatísticos. Vejamos, o tratamento tabular dado aos dados discretos e contínuos.

1 – Representação de Dados Discretos

Exemplo: Seja a seguinte amostra: 3, 4, 4, 5, 7, 6, 6, 7, 7,4, 5, 5, 6, 6, 7, 5, 8, 5, 6, 6 onde encontramos valores discretos.

1.1 – Freqüência simples absoluta (f_i)

É o número de observações correspondente a um determinado valor. Por exemplo: a freqüência do valor 3 é 1, do valor 4 é 3, conforme podemos ver no quadro abaixo.

VALOR	FREQÜÊNCIA
3	1
4	3
5	5
6	6
7	4
8	1
TOTAL	20

1.2– Amplitude amostral (A)

É a diferença entre o maior e o menor valor da amostra.

A = 8 – 3 → **A = 5**

1.3. Tipos de Freqüências

Simples { **Absolutas** / **Relativas**

OBS.: pode-se empregar nos cálculos tanto a absoluta como a relativa

Acumuladas
- "Abaixo de" (crescente) { Absolutas / Relativas
- "Acima de" (Decrescente) { Absolutas / Relativas

Utilizando o exemplo citado anteriormente, podemos fazer o seguinte quadro:

VALOR x_i	FREQÜÊNCIA SIMPLES ABSOLUTA f_i	FREQÜÊNCIA SIMPLES RELATIVA f_{ri}	FREQÜÊNCIA ACUMULADA CRESCENTE (ABSOLUTAS)	FREQÜÊNCIA ACUMULADA DECRESCENTE (ABSOLUTAS)
3	1	1/20 = 0,05 ou 5%	1	20
4	3	3/20 = 0,15 ou 15%	4	19
5	5	5/20 = 0,25 ou 25%	9	16
6	6	6/20 = 0,30 ou 30%	15	11
7	4	4/20 = 0,20 ou 20%	19	5
8	1	1/20 = 0,05 ou 5%	20	1
TOTAL	20	20/20 = 1 ou 100%		

2- Representação de Dados Contínuos

Na representação de dados contínuos utiliza-se a forma de intervalos (distribuição intervalar) em virtude desses dados serem obtidos, normalmente, através de medidas.Isto melhora a distribuição dos erros cometidos (erro do observador, do método utilizado, do equipamento de medida, etc). Usa-se também a representação intervalar para dados discretos, levando-se em conta, principalmente, o tratamento de grandes amostras (número de elementos ≥ 30).

Desta maneira poderemos encontrar os seguintes elementos nesta representação:

2.1. – Classe

É cada um dos grupos ou intervalos obtidos a partir do agrupamento ou conjunto de dados (normalmente de um rol).

Para a determinação do número de classes existem diversos métodos, dentre os quais destacam-se:

A Regra de Sturges

$$\boxed{K = 1 + 3{,}3 \log_{10} N}$$

Onde:

K = número de classes
N = número total de observações

Por exemplo, se o número de observações for 500:
N = 500
$K = 1 + 3{,}3 \log_{10} 500$
K = 9,906 ou **K = 10 aproximadamente**

A Regra do Quadrado

K = √N

Utilizando-se o quadrado perfeito mais próximo.

Por exemplo, se N = 50:

K = **√50 = 7 aproximadamente → K = 7**

A prática recomenda para K a distribuição intervalar de $5 \leq K \leq 16$.

Representação moderna

Utiliza a simbologia de intervalos abertos ou fechados.

├ ou ┤ ou ├┤

Assim, teríamos:

10 ├ 20, inclui o 10 e exclui o 20

20 ├ 30, inclui o 20 e exclui o 30

E assim por diante

2.2 – Amplitude do Intervalo de Classe (C)

É a diferença entre os limites superiores e inferiores.

Exemplo:

Classes	Freqüências
140 – 149	1
150 – 159	3
160 – 169	5
170 – 179	7
180 – 189	2

$h = Lim_{superior} - Lim_{inferior}$ ou $C = Lim_{superior} - Lim_{inferior}$

h = 159 – 150 → **9**

2.3 – Ponto Médio da Classe (Pm)

É a média aritmética simples dos limites (superior e inferior) de cada classe.

CLASSES	FREQÜÊNCIAS	PONTOS MÉDIOS
2 ├ 4	6	$\frac{2+4}{2}=3$
4 ├ 6	9	$\frac{4+6}{2}=5$
6 ├ 8	2	$\frac{6+8}{2}=7$

2.4 – Passos ou roteiro a ser seguido na construção de uma tabela de freqüência que utiliza dados agrupados em classes

1º PASSO – Pegar os dados brutos (amostra) e definir o rol.

2º PASSO – Calcular a amplitude amostral.

A = MAIOR VALOR – MENOR VALOR

3º PASSO – Calcular o número de classes **K** (Regra de Sturges ou Regra do Quadrado, etc.).

4º PASSO – Determinar **C** (amplitude do intervalo de classe):

C = $\frac{A}{K}$ **A**

5º PASSO – Escolher os limites de classe, preferindo, sempre que possível, números inteiros.

6º PASSO – Construir a tabela de freqüências.

TESTES E EXERCÍCIOS

1. Dada amostra: 3, 4, 4, 5, 7, 6, 6, 7, 7, 4, 5, 5, 6, 6, 7, 5, 8, 5, 6, 6, pede-se:

a) construir a distribuição de freqüência;
b) determinar as freqüências relativas;
c) determinar as freqüências acumuladas;
d) qual é a amplitude amostral?;
e) qual é a porcentagem de elementos maiores que 5?

2. Dadas as notas de 50 alunos, determine os pedidos.
Notas:

60	85	33	52	65	77	84	65	74	57
71	35	81	50	35	64	74	47	54	68
80	61	41	91	55	73	59	53	77	45
41	55	78	48	69	85	67	39	60	76
94	98	66	66	73	42	65	94	88	89

Pede-se:

Construa uma tabela que possua freqüências : absoluta, relativa % e acumulada crescente, e ainda uma coluna com os pontos médios das classes.

3. A tabela abaixo representa os salários pagos a 100 operários de uma empresa.

Nº DE SALÁRIOS MÍNIMOS	**Nº DE OPERÁRIOS**
0 ├ 2	40
2 ├ 4	30
4 ├ 6	10
6 ├ 8	15
8 ├ 10	5
TOTAL	100

Pede-se:
a) nº de operários que ganham até três salários mínimos;
b) nº de operários que ganham até seis salários mínimos;
c) porcentagem de operários com salário entre 6 e 8 salários mínimos;
d) porcentagem de operários com salário igual ou inferior a 4 salários mínimos.
Respostas:
a) 55
b) 80
c) 15%
d) 70%

5. Freqüência absoluta é:

a) o número de repetições de uma variável;
b) a soma das freqüências simples;
c) o número de valores que se repetem dividido pelo total de valores;
d) n.r.a.
Resposta: Alternativa A

6. Freqüência total é:

a) o número de repetições de uma variável;
b) a soma das freqüências simples absolutas;
c) a soma das freqüências relativas menos as freqüências absolutas;
d) n.r.a.
Resposta: Alternativa B

7. Considere a seguinte distribuição de freqüências, que corresponde aos diferentes preços de um determinado produto em uma pesquisa realizada em vinte lojas:

PREÇOS (R$)	**Nº DE LOJAS**
50,00	2
51,00	5
52,00	6
53,00	6
54,00	1
TOTAL	20

Pede-se:
a) quantas lojas apresentaram um preço de R$ 52,00;
b) construir uma coluna de freqüências simples relativas;
c) construir uma coluna de freqüências absolutas acumuladas;
d) qual a porcentagem de lojas que apresentou um preço de até R$ 53,00;
e) que porcentagem de lojas apresentou um preço maior do que R$ 51,00 e menor do que R$ 54,00.

Respostas: a) 6 d) 95% e) 60%

14. Assinale, entre as alternativas, aquela que contiver uma afirmação verdadeira.

a) A amplitude do intervalo de classe é calculada pela soma entre os limites reais inferior e superior de uma classe.
b) Obtém-se o ponto médio de uma classe pela média aritmética dos limites inferior e superior.
c) Um intervalo de classe aberto em seus dois limites inclui ambos os números extremos.
d) Intervalos de classe fechados têm seus limites superior e inferior reais excluídos dos números que os compõem.

Resposta: Alternativa B

17. (Fiscal de Tributos de Minas Gerais / 96) Ouvindo-se 300 pessoas sobre o tema "Reforma da Previdência, contra ou a favor?", foram obtidas 123 respostas a favor, 72 contra, 51 pessoas não quiseram opinar, e o restante não tinha opinião formada sobre o assunto. Distribuindo-se esses dados numa tabela, obtém-se:

OPINIÃO	FREQÜÊNCIA	FREQÜÊNCIA RELATIVA
FAVORÁVEL	123	X
CONTRA	72	Y
OMISSOS	51	0,17
SEM OPINIÃO	54	0,18
TOTAL	300	1,00

Na coluna freqüência relativa, os valores de x e y são, respectivamente:

a) 0,41 e 0,24
b) 0,38 e 0,27
c) 0,37 e 0,28
d) 0,35 e 0,30
e) 0,30 e 0,35

Resposta: Alternativa A

(Fiscal de Tributos de Minas Gerais / 96) Responda às questões 18 e 19 com base na seguinte situação: a distribuição a seguir indica o número de acidentes ocorridos com 40 motoristas de uma empresa de ônibus.

NÚMERO DE ACIDENTES	0	1	2	3	4	5	6
NÚMERO DE MOTORISTAS	13	7	10	4	3	2	1

18. O número de motoristas que sofreram pelo menos 4 acidentes é:

a) 3 **b) 6** c) 10 d) 27 e) 30

Resposta: Alternativa B

19. A porcentagem de motoristas que sofreram no máximo 2 acidentes é:

a) 25% b) 32,5% c) 42,5% d) 57,5% **e) 75%**

Resposta: Alternativa E

Exercício sobre Interpolação

INTERPOLAÇÃO

14. (AFRF-2002) Em um ensaio para o estudo da distribuição de um atributo financeiro (X),foram examinados 200 itens de natureza contábil do balanço de uma empresa. Esse exercício produziu a tabela de freqüências abaixo. A coluna *Classes* representa intervalos de valores de X em reais e a coluna *P* representa a freqüência relativa acumulada. Não existem observações coincidentes com os extremos das classes.

Classes	P (%)
70-90	5
90-110	15
110-130	40
130-150	70
150-170	85
170-190	95
190-210	100

Assinale a opção que corresponde à estimativa da freqüência relativa de observações de X menores ou iguais a 145.

a) 62,5% b) 70,0% c) 50,0% d) 45,0% e) 53,4%

3. (AFRF-2002.2) O atributo do tipo contínuo X, observado como um inteiro, numa amostra de tamanho 100, obtida de uma população de 1000 indivíduos, produziu a tabela de freqüências seguinte:

Classes	Freqüência (f)
29,5-39,5	4
39,5-49,5	8
49,5-59,5	14
59,5-69,5	20
69,5-79,5	26
79,5-89,5	18
89,5-99,5	10

Assinale a opção que corresponde à estimativa do número de indivíduos na população com valores do atributo X menores ou iguais a 95,5 e maiores do que 50,5.

a) 700 b) 638 **c) 826** d) 995 e) 900

1. (AFRF-2000) Freqüências Acumuladas de Salários Anuais, em Milhares de Reais, da Cia. Alfa

Classes de Salário	Freqüências Acumuladas
(3 ; 6]	12
(6 ; 9]	30
(9 ; 12]	50
(12 ; 15]	60
(15 ; 18]	65
(18 ; 21]	68

Suponha que a tabela de freqüências acumuladas tenha sido construída a partir de uma amostra de 10% dos empregados da Cia. Alfa. Deseja-se estimar, utilizando interpolação linear da ogiva, a freqüência populacional de salários anuais iguais ou inferiores a R$ 7.000,00 na Cia. Alfa. Assinale a opção que corresponde a este número.

a) 150 b) 120 c) 130 d) 160 **e) 180**

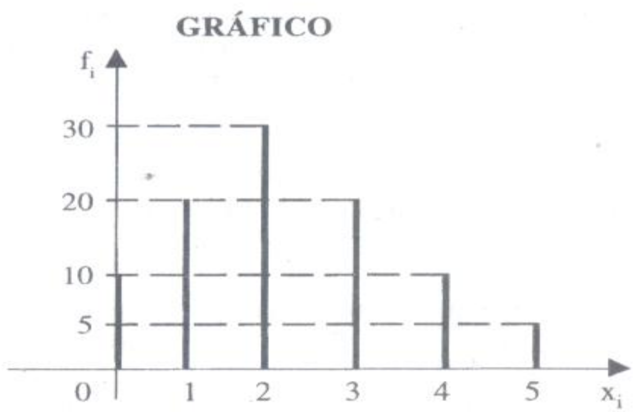
REPRESENTAÇÃO GRÁFICA

1 – Gráficos de Hastes ou Bastões

É muito utilizado na representação gráfica de dados **não agrupados em classes**, o que ocorre normalmente com dados discretos. Dizemos que neste caso **não há perda de informação**, pois os valores da variável aparecem individualmente, como constam da amostra.

TABELA

X_i	F_i
0	10
1	20
2	30
3	20
4	10
5	5



2- Histograma

É muito utilizado na representação gráfica de dados **agrupados em classes**, o que ocorre normalmente com dados contínuos. Dizemos que neste caso **há perda de informação**. O seu uso é recomendado quando:

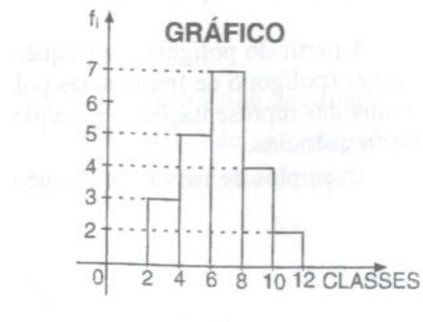
- os valores da variável são inteiros e não-inteiros, ou somente não-inteiros;
- a quantidade de valores da variável é grande, no caso de valores inteiros (discretos);
- não for importante a perda de informação ocasionada pelos dados apresentados.

OBS.:

- **O histograma pode ser construído tanto com freqüência absoluta como relativa**
- **É destinado a representação de freqüências, não se aplicando a séries temporais e geográfica.**

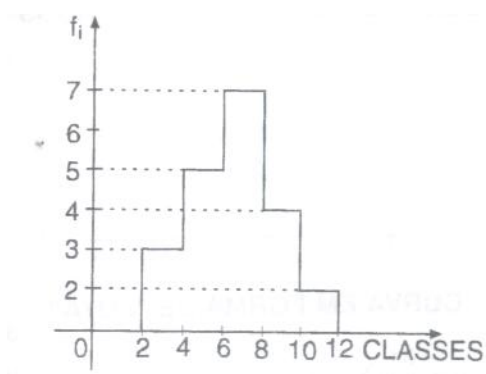
TABELA

CLASSES	**F_i**
2 ├ 4	3
4 ├ 6	5
6 ├ 8	7
8 ├ 10	4
10 ├ 12	2



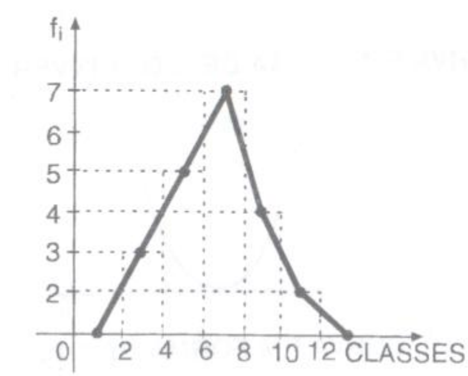
3- Poligonal Característica

É o gráfico que utiliza em sua representação o **contorno** do histograma.



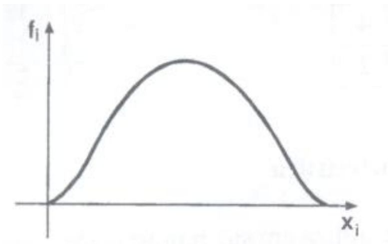
4 – Polígono de Freqüências

É o gráfico que se obtém unindo-se por linhas retas os pontos médios das bases superiores dos retângulos de um histograma.



5 – Curva de Freqüências

A partir do polígono de freqüências podemos representar contornos mais suaves (polígono de freqüências polido), utilizando-se curvas para chegarmos a uma das representações de grande utilidade para a Estatística que é a curva de freqüências.

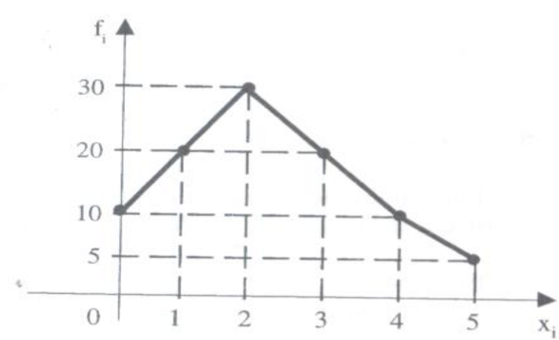


6 – Gráficos de Linhas

É muito utilizado na representação gráfica de dados **não agrupados em classes**, de valores que aumentam gradativamente como temperatura, consumo e etc...,e também para a representação de **séries temporais** (cotação de ações, vendas, etc.).

TABELA

X_i	F_i
0	10
1	20
2	30
3	20
4	10
5	5



Observação:

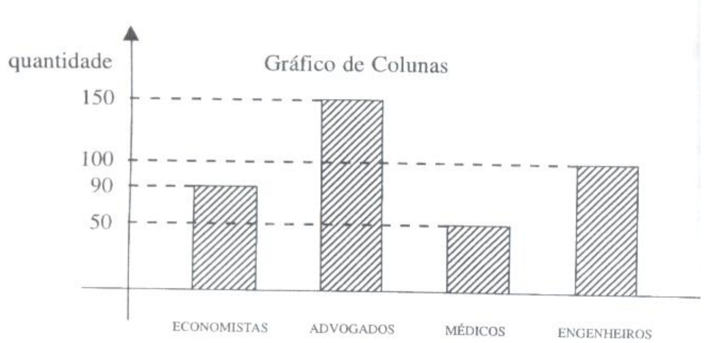
Neste gráfico são usadas linhas retas. Quando o número de valores aumenta, podemos utilizar curvas, que irão constituir "curvas de freqüência", que desempenharão a mesma importância das já citadas para dados agrupados em classes.

7 – Gráfico de Colunas

É o gráfico que corresponde ao histograma, porém utilizado na representação de **dados nominais** (ou categóricos) ou em **séries temporais**.

Exemplo:

Representação de algumas classes profissionais formadas por uma Universidade em determinado ano.



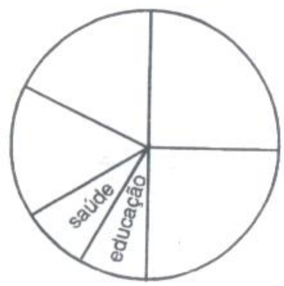
Obs.: o gráfico de barras nada mais é do que o gráficos de colunas na horizontal.

8 – Gráficos em Setores

Muito utilizado quando se deseja mostrar "partes de um todo", conforme ocorre em orçamentos, produções, vendas, etc., e na comparação de proporções ou quantidades percentuais.

Exemplo:

Orçamento da União



9 – Ogivas

É obtida pelo polígono das freqüências acumuladas. Como se fosse um histograma das freqüências acumuladas, só que no polígono da freqüência acumulada crescente liga-se o canto superior direito de cada coluna.

TESTES E EXERCÍCIOS

1. Veja a tabela abaixo:

CLASSES	FREQÜÊNCIA
100 ├ 110	2
110 ├ 120	4
120 ├ 130	6
130 ├ 140	6
140 ├ 150	2
TOTAL	$\Sigma f_i = 20$

Pede-se:
a) construir o histograma da distribuição de freqüências dadas;
b) construir o polígono de freqüências;
c) construir uma curva de freqüências acumuladas (ogiva).

2. Uma empresa apresentou a seguinte evolução de vendas de um de seus artigos, no período de dois anos, computados trimestralmente.

Período	1991				1992			
Trimestres	1º	2º	3º	4º	1º	2º	3º	4º
Vendas	100	120	95	90	120	140	113	108

Representar a evolução das vendas, através de um gráfico em hastes ou bastões.

3. Um produto é vendido por apenas três empresas, em um determinado mercado. Em 1993, para um total de 18.000 unidades vendidas, tivemos a distribuição das vendas, conforme o quadro abaixo:

EMPRESA	A	B	C
VENDAS	7.200	4.800	6.000

Represente graficamente (gráfico de setores) a distribuição das vendas de cada empresa em termos absolutos e percentuais.

4. (TCU / 93) Gráficos são instrumentos úteis na análise estatística. Assinale a definição / afirmação incorreta.

a) Um histograma representa uma distribuição de freqüências para variáveis do tipo contínuo.
b) O gráfico de barras representa, por meio de uma série de barras, quantidades ou freqüências para variáveis categóricas.
c) O gráfico de setores é apropriado, quando se quer representar as divisões de um montante total.
d) Um histograma pode ser construído utilizando-se, indistintamente, as freqüências absolutas ou relativas de um intervalo de classe.
e) Uma ogiva pode ser obtida ligando-se os pontos médios dos topos dos retângulos de um histograma.

Resposta: Alternativa E

6. (TCDF – 95) Em relação aos tipos de gráficos, assinale a opção correta.

a) Uma série categórica é melhor representada por um gráfico de linha.
b) Uma série cronológica é melhor representada por um gráfico de setores.
c) Se uma distribuição de freqüências apresenta intervalos de tamanhos desiguais, o melhor gráfico para representá-la é um polígono de freqüências.
d) O gráfico de barras é usado somente para séries geográficas.
e) O gráfico de setores é usado para comparar proporções.
Resposta: Alternativa E

MEDIDAS DE TENDÊNCIA CENTRAL E SEPARATRIZES

1 – Média Aritmética ($\overline{X}$)

É o quociente entre a soma dos valores observados e o seu número total.

Exemplo:

x = { 1, 3, 5, 7, 9}

$$\bar{x} = \frac{1+3+5+7+9}{5} \Rightarrow \frac{25}{5} \Rightarrow 5$$

2 – Média Aritmética Ponderada

É o quociente entre a soma dos produtos dos valores da variável pelos respectivos pesos e a soma dos pesos.

Exemplo:
Quatro provas com pesos 1, 2, 3, 4.
Notas 8, 7, 9 e 9

$$\overline{x} = \frac{(8 \text{ x } 1) + (7 \text{ x } 2) + (9 \text{ x } 3) + (9 \text{ x } 4)}{1 + 2 + 3 + 4}$$

Se tivermos uma representação tabular podemos considerar:

pesos ≡ freqüências absolutas

Exemplo:

x_i	f_i	$x_i f_i$
4	1	4 X 1 = 4
5	5	5 X 5 = 25
6	6	6 X 6 = 36
7	5	7 x 5 = 35
8	3	8 x 3 = 24
SOMA	20	124

$\overline{X} = \frac{124}{20}$ → **x = 6,2**

Se tivermos valores agrupados em classe, calculamos os pontos médios das classes, e estes representarão os valores de x_i das classes.

CLASSES	**F_i**	x_i	$x_i f_i$
2 ├ 4	3	3	9
4 ├ 6	5	5	25
6 ├ 8	10	7	70
8 ├ 10	5	9	45
10 ├ 12	3	11	33
SOMA	26		182

$\overline{x} = \frac{182}{26}$ → **x = 7**

3 – Propriedades da Média Aritmética

1º) Propriedade: A soma algébrica dos desvios tomados em relação à média aritmética é igual **a zero.**

$\Sigma d_i = 0$, onde $d_i = x_i - \overline{x}$ → conjunto de valores

Dados tabulados

x_i	f_i	x_if_i
1	2	2
3	4	12
5	7	35
7	4	28
9	2	18
SOMA	19	95

com média $\bar{x} = \frac{95}{19} \rightarrow \bar{x} = 5$

Fazendo os desvios, teremos:

x_i	f_i	$d_i = x_i - \bar{x}$	d_if_i
1	2	1 – 5 = – 4	– 8
3	4	3 – 5 = – 2	– 8
5	7	5 – 5 = 0	0
7	4	7 – 5 = 2	8
9	2	9 – 5 = 4	8
			$\Sigma = 0$

Observação:

Para dados tabulados agrupados em classes o procedimento é análogo ao apresentado, devendo-se calcular os pontos médios das classes(x_i) para a inicialização dos cálculos.

2ª) Propriedade: A soma dos quadrados dos desvios tomados em relação à média aritmética é **um mínimo**.

Exemplos:

Conjunto de valores

~~x~~= {1, 3, 5, 7, 9}, onde x = 5

x_i	$d_i = x_i - \bar{x}$	$\mathbf{d_i^2}$
1	– 4	16
3	– 2	4
5	0	0
7	2	4
9	4	16
SOMA		$\mathbf{\Sigma d_i^2 = 40}$

Se escolhermos um valor arbitrário x_0 e tomarmos esta média arbitrária para o cálculo dos desvios, teremos:

x_i	$d_i = x_i - \bar{x}_0$	$d_i^2 = (x_i - \bar{x})^2$
1	1 – 3 = – 24	
3	3 – 3 = 0	0
5	5 – 3 = 0	4
7	7 – 3 = 4	16
9	9 – 3 = 6	36
SOMA	Σ = ...	**Σ = 60**

para $\bar{x}_0 = 3$

Vemos que 60 é maior do que 40 da tabela anterior, assim como qualquer outra soma de desvios ao quadrado (tomados em relação a valores arbitrários) será maior do que a tomada em relação à média.

Observação:
Para dados tabulados agrupados em classes é válida a mesma observação feita na propriedade anterior.

3ª) Propriedade: Média das médias. Se n_1 valores têm média $\bar{x}_1$; n_2 valores têm média $\bar{x}_2$; ... n_m valores têm média $\bar{x}_m$, a média do conjunto formado por todos os valores é dada pela relação:

Exemplo:

$$x = \frac{\bar{x}_1 n_1 + \bar{x}_2 n_2 + ... + \bar{x}_m n_m}{n_1 + n_2 + ... + n_m}$$

Exemplos:

$x_1 = \{1, 3, 5, 7, 9\} \rightarrow n_1 = 5$ $\bar{X}_1 = 5$

$x_2 = \{2, 4, 6, 8, 10, 12, 14\} \rightarrow n_2 = 7$ $\bar{X}_2 = 8$

$$x = \frac{5.5 + 7.8}{5 + 7} \rightarrow \frac{25 + 56}{12} \rightarrow \frac{81}{12} \rightarrow 6{,}75$$

4ª) Propriedade: se somarmos (ou subtrairmos) um valor constante e arbitrário **K** a cada um dos elementos de um conjunto de valores, a média aritmética fica somada (ou subtraída) dessa constante.

Exemplo:

X = {1, 3, 5, 7, 9} com $\bar{x} = 5$

fazendo K = 10 e somando, teremos:

Y = {11, 13, 15, 17, 19} com $\bar{y} = 15$

Podemos ver que $\bar{y} = 15$ é igual a $\bar{x} = 5$ acrescida de K = 10, ou seja, $\bar{y} = 5 + 10 = 15$.

5º) Propriedade: Se multiplicarmos (ou dividirmos) cada um dos elementos de um conjunto de valores por um valor **K** (constante, arbitrário e ≠ 0), a média aritmética fica multiplicada (ou dividida) por essa constante.

Exemplo:

X = {100, 200, 300, 400, 500} com $\overline{x} = 300$

Fazendo K = 10 e dividindo, teremos:

Y = {10, 20, 30, 40, 50} com $\overline{y} = 30$

Da mesma forma, podemos ver que $\overline{y} = 30$ é igual a $\overline{x} = 300$ dividido por K = 10, ou seja, y = $\frac{300}{10}$ 30.

4 – Média Geométrica Simples ($\overline{Xg}$)

Ela, como a média aritmética, também pode ser simples ou ponderada.

5 – Média Geométrica Simples:

É a raiz índice **n** do produto dos **n** números.
Exemplo:

X = {1, 3, 5, 7, 9), com 5 números, portanto n = 5.

$\overline{Xg} = \sqrt[5]{1 . 3 . 5 . 7 . 9} \rightarrow Xg = \sqrt[5]{945} \rightarrow Xg = 3{,}936$

Y = {2, 4 e 8}

$\overline{Yg} = \sqrt[3]{2 . 4 . 8} \rightarrow Yg = \sqrt[3]{64} \rightarrow Yg = 4$

Observação:

Em sua representação literal o produto é designado pelo símbolo **Π,** que representa o "produtório", assim temos:

$$\overline{x}g = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

5 – Média Geométrica Ponderada

Neste caso devem ser utilizadas as freqüências com que os números aparecem na amostra, pois elas fornecem a ponderação (peso).

Exemplo:

x_i	f_i
1	2
3	4
5	3
7	1
TOTAL	10

$n = \Sigma f_i$ e k = 4

$\overline{X} = \sqrt[10]{1^2 \cdot 3^4 \cdot 5^3 \cdot 7^1} = 3{,}055.$

assim, temos que : $\bar{x}g = \sqrt[n]{\prod_{i=1}^{n} x_i^{\,f_i}}$

Propriedades da Média Geométrica

1ª) O produto dos quocientes de cada valor de um conjunto de números pela média geométrica do conjunto é igual a 1.

Exemplo:

X = {4,9} .:

$\bar{x}g = \sqrt{4 \times 9} = 6$

Aplicando a propriedade, teremos:

$$\frac{4}{6} \times \frac{9}{6} = \frac{36}{36} = 1$$

2ª) Séries que possuem, o mesmo número de elementos, com a mesma **soma,** apresentam a mesma média aritmética e séries que possuem o mesmo número de elementos com o mesmo **produto** têm a mesma média geométrica.

Exemplo:

X_1 = {1, 3, 9} → $\bar{x} = \frac{1+3+9}{3} = \frac{13}{3} = 4{,}33$

X_2 = {2, 4, 7} → $\bar{x} = \frac{2+4+7}{3} = \frac{13}{3} = 4{,}33$

São iguais

&

Y_1 = {1, 2, 28} → $Yg_1 = \sqrt[3]{1.2.28} \Rightarrow Yg_1 = \sqrt[3]{56} \Rightarrow Yg_1 = 3{,}825$

Y_2= {2, 4, 7} → $Yg_2 = \sqrt[3]{2.4.7} \Rightarrow Yg_2 = \sqrt[3]{56} \Rightarrow Yg_2 = 3{,}825$

São iguais

3ª) $\overline{X}_g \le \overline{X}$, ou seja, a média geométrica é menor ou igual à média aritmética.

4ª) Para valores próximos, teremos $\overline{X}_g$ próximo de $\overline{X}$, enquanto que para valores afastados, teremos $\overline{X}_g$ afastado de $\overline{X}$.

Exemplo:

X = {4 e 6}, valores próximos

$$\left.\begin{array}{l} \bar{x} = \dfrac{4+6}{2} \Rightarrow \dfrac{10}{2} \Rightarrow 5 \\[2em] \bar{x}g = \sqrt{4.6} \Rightarrow \sqrt{24} \Rightarrow 4{,}898 \end{array}\right\} \text{Médias Próximas}$$

Se tivermos:

X = {2, 50}, valores afastados

$$\left.\begin{array}{l} \bar{x} = \dfrac{2+50}{2} = \dfrac{52}{2} = 26 \\[2em] \bar{x}g = \sqrt{2.50} = \sqrt{100} = 10 \end{array}\right\} \text{Médias Afastadas}$$

5 – Média Harmônica ($\bar{X}_h$)

“É o inverso da média aritmética dos inversos dos números.”

5.1 – Média Harmônica Simples

Exemplo:

X = {1, 4, 9} → $\bar{x}_h = \dfrac{1}{\dfrac{\frac{1}{1}+\frac{1}{4}+\frac{1}{9}}{3}} = \dfrac{3}{\frac{1}{1}+\frac{1}{4}+\frac{1}{9}} = \dfrac{3}{\frac{49}{36}} = \dfrac{3}{1} \times \dfrac{36}{49} = \dfrac{108}{49} = 2{,}2$

Se generalizarmos, teremos:

$$\bar{x}_h = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

5.2 – Média Harmônica Ponderada

As freqüências com que os números aparecem na amostra devem ser utilizadas, em virtude de fornecerem a ponderação (peso).

Exemplo:

x_i	f_i	$f_i : x_i$
1	2	2 : 1 = 2
3	4	4 : 3 = 1,33
5	3	3 : 5 = 0,6
7	1	1 : 7 = 0,143
TOTAL	10	4,073

$$\boxed{\mathbf{X_h} = \frac{10}{4{,}073} \mathbf{= 2{,}455}}$$

Se generalizarmos, teremos:

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

Observação:

Neste caso $n = \Sigma f_i = 10$ e $k = 4$

Propriedades da Média Harmônica

1ª) $\bar{X}_h \leq \bar{X}_g \leq \bar{X}$, ou seja, a média harmônica é menor ou igual à média geométrica e esta é menor ou igual à média aritmética.

2ª) Valores próximos têm as três médias próximas e valores afastados têm as três médias afastadas.

MÉDIA PELO MÉTODO DA VARIÁVEL TRANSFORMADA

1º Devemos elaborar a coluna da variável transformada.

- $$\left(\frac{PM - 1^o\,PM}{Amplitude\,(classe)}\right)$$

Onde:

PM = valores dos pontos médios;
1º PM = valor do 1º ponto médio encontrado;

14. (AFRF-2002) Em um ensaio para o estudo da distribuição de um atributo financeiro (X) foram examinados 200 itens de natureza contábil do balanço de uma empresa. Esse exercício produziu a tabela de freqüências abaixo. A coluna *Classes* representa intervalos de valores de X em reais e a coluna *P* representa a freqüência relativa acumulada. Não existem observações coincidentes com os extremos das classes.

Classes	P (%)
70-90	5
90-110	15
110-130	40
130-150	70
150-170	85
170-190	95
190-210	100

Assinale a opção que corresponde à média da distribuição.

a) 124   b) 127   c) 136   d) 138   e) 140

TESTES SOBRE MÉDIA

(AFRF-2000) Freqüências Acumuladas de Salários Anuais, em Milhares de Reais, da Cia. Alfa

Classes de Salário	Freqüências Acumuladas
(3 ; 6]	12
(6 ; 9]	30
(9 ; 12]	50
(12 ; 15]	60
(15 ; 18]	65
(18 ; 21]	68

Quer-se estimar o salário médio anual para os empregados da Cia. Alfa. Assinale a opção que representa a aproximação desta estatística calculada com base na distribuição de freqüências.

a) 9,93 b) 15,00 c) 13,50 d) 10 e) 12,50

8. (FISCAL DE TRIBUTOS DE MG-96) A estatura média dos sócios de um clube é 165cm, sendo a dos homens 172cm e a das mulheres 162cm. A porcentagem de mulheres no clube é de:

a) 62% b) 65% c) 68% **d) 70%** e) 72%

Para efeito das quatro próximas questões, considere os seguintes dados:

DISTRIBUIÇÃO DE FREQÜÊNCIAS DAS IDADES DOS FUNCIONÁRIOS DA EMPRESA ALFA, EM 1º/1/90

Classes de Idades (anos)	Freqüências (fi)	Pontos Médios (PM)	$\frac{PM-37}{5}=di$	di.fi	$di^2.fi$	$di^3.fi$	$di^4.fi$
19,5 !--- 24,5	2	22	-3	-6	18	-54	162
24,5 !--- 29,5	9	27	-2	-18	36	-72	144
29,5 !--- 34,5	23	32	-1	-23	23	-23	23
34,5 !--- 39,5	29	37	---	---	---	---	---
39,5 !--- 44,5	18	42	1	18	18	18	18
44,5 !--- 49,5	12	47	2	24	48	96	192
49,5 !--- 54,5	7	52	3	21	63	189	567
Total	n=100			16	206	154	1106

10. (AFTN-96) Marque a opção que representa a média das idades dos funcionários em 1º/1/90.

a) 37,4 anos **b) 37,8 anos** c) 38,2 anos d) 38,6 anos e) 39,0 anos

Para efeito das duas questões seguintes, sabe-se que o quadro de pessoal da empresa continua o mesmo em 1º/1/96.

12. (AFTN-96) Marque a opção que representa a média das idades dos funcionários em 1º/1/96.

a) 37,4 anos b) 39,0 anos c) 43,4 anos **d) 43,8 anos** e) 44,6 anos

9. (Auditor do Tesouro Municipal - Recife – 2003) Em uma amostra, realizada para se obter informação sobre a distribuição salarial de homens e mulheres, encontrou-se que o salário médio vale R$ 1.200,00. O salário médio observado para os homens foi de R$ 1.300,00 e para as mulheres foi de R$ 1.100,00. Assinale a opção correta.

a) O número de homens na amostra é igual ao de mulheres.
b) O número de homens na amostra é o dobro do de mulheres.
c) O número de homens na amostra é o triplo do de mulheres.
d) O número de mulheres é o dobro do número de homens.
e) O número de mulheres é o quádruplo do número de homens.

(Oficial de Justiça Avaliador TJ CE 2002 / ESAF) A tabela abaixo apresenta a distribuição de freqüências do atributo salário mensal medido em quantidade de salários mínimos para uma amostra de 200 funcionários da empresa X. As três próximas questões referem-se a essa tabela. Note que a coluna **Classes** refere-se a classes salariais em quantidades de salários mínimos e que a coluna **P** refere-se ao percentual da freqüência acumulada relativo ao total da amostra. Não existem observações coincidentes com os extremos das classes.

Classes	P
4 – 8	20
8 – 12	60
12 – 16	80
16 – 20	98
20 – 24	100

20. (TJ CE 2002 / ESAF) Assinale a opção que corresponde ao salário médio amostral calculado a partir de dados agrupados.

a) 11,68 b) 13,00 c) 17,21 d) 16,00 e) 14,00

6 – Moda (M_0)

É uma medida de tendência central que se caracteriza pelo valor mais freqüente (maior freqüência absoluta simples).

6.1. – Moda em Valores Não-Tabulados

X_1 = {2, 4, 5, 5, **6**, **6**, **6**, 7, 8, 9} **M_0 = 6**

CONJUNTO AMODAL, não existe moda → X_2 = {2, 2, 3, 3, 4, 4}

CONJUNTO BIMODAL → X_3 = {1, **2, 2, 2**, 3, 4, 4, **5, 5, 5**, 6}

M_0 = 2
M_0 = 5

OBS.: Três modas em diante o conjunto é multimodal.

6.2. – Moda em Valores Tabulados

Podemos encontrar os seguintes casos:

a) valores discretos → DETERMINAÇÃO IMEDIATA

Exemplo:

x_i	f_i
1	3
2	5
3	9
4	4
5	1

b) valores contínuos ou agregados em classes: devemos usar alguns métodos

Métodos

I. Moda Bruta;
II. Método de King;
III. Método de Czuber.

OBS.: Quando o enunciado não especificar o método a ser usado, usaremos o método de Czuber.

I - Moda Bruta

- determinação da classe modal;
- moda bruta = ponto médio da classe modal.

Exemplo:

CLASSES	**F_i**
2 ├ 4	3
4 ├ 6	5
6 ├ 8	**9**
8 ├ 10	4
10 ├ 12	2

Classe Modal : 6 ├ 8
Moda Bruta = 7

II - Método de King:

$$M_o = L_i + c.\frac{f_{post}}{f_{ant} + f_{post}}$$

onde: L_i = limite inferior da classe modal;
c = intervalo da classe modal;
f_{ant} = freqüência absoluta da classe anterior à modal;
f_{post} = freqüência absoluta da classe posterior à modal.
Exemplo:

CLASSES	F_i
2 ├ 4	3
4 ├ 6	5
6 ├ 8	**7**
8 ├ 10	4
10 ├ 12	2

1º Identifica-se a classe modal pela classe de maior freqüência→Classe Modal: 6 ├ 8

2º Emprega-se a fórmula:

$$Mo = 6 + 2.\frac{4}{5+4} \quad \rightarrow \quad 6+2\ .\ 0{,}44 \quad \rightarrow \quad 6+0{,}88 \quad \rightarrow \quad \mathbf{6{,}88}$$

M_0 = 6,88 aprox**.**

III - Método de Czuber

É o método considerado mais preciso.

$$Mo = L_i + \frac{\Delta 1}{\Delta 1 + \Delta 2} \cdot c$$

Onde:

L_i = Limite inferior da classe modal;
$\Delta 1$ = Diferença entre a freqüência da classe modal e a freqüência imediatamente inferior;
$\Delta 2$ = Diferença entre a freqüência da classe modal e a freqüência imediatamente posterior;
c = Intervalo da classe modal

Exemplo:

CLASSES	F_i
2 ├ 4	3
4 ├ 6	5
6 ├ 8	7
8 ├ 10	4
10 ├ 12	2

1º Identifica-se a classe modal pela classe de maior freqüência→Classe Modal: 6 ├ 8

2º Emprega-se a fórmula:

$$Mo = 6 + \frac{2}{3+2} \cdot 2 \Rightarrow 6 + 0{,}4 \cdot 2 \Rightarrow 6 + 0{,}8 \Rightarrow 6{,}8$$

Mo = 6,8

TESTES SOBRE MODA

DISTRIBUIÇÃO DE FREQÜÊNCIAS DAS IDADES DOS FUNCIONÁRIOS DA EMPRESA ALFA, EM 1º/1/90

Classes de Idades (anos)	Freqüências (fi)	Pontos Médios (PM)	$\frac{PM-37}{5}=di$	di.fi	di^2.fi	di^3.fi	di^4.fi
19,5 !--- 24,5	2	22	-3	-6	18	-54	162
24,5 !--- 29,5	9	27	-2	-18	36	-72	144
29,5 !--- 34,5	23	32	-1	-23	23	-23	23
34,5 !--- 39,5	29	37	---	---	---	---	---
39,5 !--- 44,5	18	42	1	18	18	18	18
44,5 !--- 49,5	12	47	2	24	48	96	192
49,5 !--- 54,5	7	52	3	21	63	189	567
Total	n=100			16	206	154	1106

10. (AFTN-96) Marque a opção que representa a moda das idades dos funcionários em 1º/1/90.

a) 35,97 anos **b) 36,26 anos** c) 36,76 anos d) 37,03 anos e) 37,31anos

Para efeito das duas questões seguintes, sabe-se que o quadro de pessoal da empresa continua o mesmo em 1º/1/95.

12. (AFTN-96) Marque a opção que representa a média das idades dos funcionários em 1º/1/95.

a) 41,26 anos b) 39,03 anos c) 43,42 anos d) 43,81 anos e) 44,67 anos

3. (AFRF-2002.2) O atributo do tipo contínuo X, observado como um inteiro, numa amostra de tamanho 100, obtida de uma população de 1000 indivíduos, produziu a tabela de freqüências seguinte:

Classes	Freqüência (f)
29,5-39,5	4
39,5-49,5	8
49,5-59,5	14
59,5-69,5	20
69,5-79,5	26
79,5-89,5	18
89,5-99,5	10

Assinale a opção que corresponde ao valor modal do atributo X, no conceito de Czuber.

a) 69,50 **b) 73,79** c) 71,20 d) 74,53 e) 80,10

12(AFCE) Com relação ao valor modal da distribuição abaixo responda:

Vendas anuais de Pequenos Comércios (em Milhar)	Fac
0 ├10	9
10 ├20	24
20 ├30	52
30 ├40	69
40 ├50	80

a) É um valor inteiro entre 24 e 26;
b) É menor que 24
c) É maior que 27
d) É um decimal entre 25 e 26
e) É igual ao limite inferior da terceira classe.

(Oficial de Justiça Avaliador TJ CE 2002 / ESAF) A tabela abaixo apresenta a distribuição de freqüências do atributo salário mensal medido em quantidade de salários mínimos para uma amostra de 200 funcionários da empresa X. As três próximas questões referem-se a essa tabela. Note que a coluna **Classes** refere-se a classes salariais em quantidades de salários mínimos e que a coluna **P** refere-se ao percentual da freqüência acumulada relativo ao total da amostra. Não existem observações coincidentes com os extremos das classes.

Classes	P
4 – 8	20
8 – 12	60
12 – 16	80
16 – 20	98
20 – 24	100

21. (TJ CE 2002 / ESAF) Assinale a opção que corresponde ao salário modal no conceito de Czuber.

a) 6 b) 8 c) 10 d) 12 e) 16

7 – Mediana (M_d)

É uma medida de tendência central que divide uma **série ordenada** (ROL) de tal maneira que pelo menos a metade ou 50% dos itens sejam iguais ou maiores que ela. Desta maneira é também uma **separatriz**, dividindo a série em partes iguais.

Podemos encontrar os seguintes casos na determinação de seu valor:

7.1 – Mediana em valores não-tabulados

a) Quando tivermos um número ímpar de valores

Exemplo:

X = {1, 3, 5, 7, 9, 11, 12, 13, 14}, onde n = 9 (ímpar)

Determina-se o elemento central:

$$E = \frac{n+1}{2} = \frac{9+1}{2} = 5.$$

Verifica-se, então, na seqüência ordenada o valor correspondente à posição 5 indicada por E. Vemos, assim, que a mediana será o valor, $M_d = 9$

b) Quando tivermos um número par de valores

Exemplo:

X = {1, 3, ⑤, ⑦, 9, 11}

Determinam-se os elementos centrais: E = n : 2 = 6 : 2 = 3

Verifica-se, então, na seqüência ordenada os valores correspondentes às posições 3 e 4 (posição seguinte), indicadas por E. Para a determinação da Mediana, calculamos a **média aritmética dos dois valores centrais.**

$$\mathbf{M_d = \frac{5+7}{2} = 6 \rightarrow M_d = 6}$$

7.2 – Mediana em valores tabulados

a) Quando se tratarem de dados discretos (não agrupados em classes)

Exemplo:

x_i	f_i
2	5
4	10
6	15
8	12
10	5
12	3
	Σ = 50

$\Sigma f_i = n = 50 \rightarrow$ número **par**

Calculamos E para um número par de valores:

E = n : 2 = 50 : 2 = 25

Em seguida calcula-se a freqüência acumulada

x_i	f_i	f_{ac}
2	5	5
4	10	15
⑥	15	30
8	12	42
10	5	47
12	3	50
	Σ = 50	

Comparamos o valor 25 com as freqüências acumuladas, verificando-se que ele está após o 15 e antes do 30.

Como o conjunto de valores é par, calculamos a mediana pela média aritmética entre os valores (25º e 26º), que no caso é o próprio 6.

Assim:

$$Md = \frac{6+6}{2} = 6 \quad \rightarrow \quad \textbf{Md = 6}$$

Observação:

Se **n** for ímpar, teremos $E = \frac{n+1}{2}$ e procedendo de maneira análoga chegaremos à mediana, lembrando que neste caso teremos um único valor central, o que dispensa o cálculo da média aritmética.

b) Quando se tratarem de dados contínuos (agrupados em classes)

$$Md = L_i + \frac{\frac{n}{2} - f_{ant.acumulada}}{f_{md}} \cdot c$$

onde:

L_i → limite inferior da classe mediana

$\frac{n}{2}$ → será sempre o número de elementos dividido por 2

f_{Md} → freqüência absoluta da classe mediana

$f_{ant\ ac}$ → freqüência acumulada anterior à classe mediana

c = intervalo da classe mediana

Exemplo:

CLASSES	F_i	f_{ac}
2 ├ 4	3	3
4 ├ 6	5	8
6 ├ 8	7	15
8 ├ 10	4	19
10 ├ 12	1	20
TOTAL	20	

← 10

Classe Mediana: 6 ├ 8 → $E = \frac{20}{2} = 10$

O valor 10 está colocado na coluna de freqüências acumuladas depois do 8 e antes do 15. Assim, teremos:

$$Md = 6 + \frac{\frac{20}{2} - 8}{7} \cdot 2 \Rightarrow 6 + \frac{10-8}{7} \cdot 2 \Rightarrow 6 + \frac{2}{7} \cdot 2 \Rightarrow 6 + \frac{4}{7} \Rightarrow 6 + 0{,}57 \Rightarrow 6{,}57$$

M_d = 6,57 aproximadamente

TESTES SOBRE MEDIANA

15. (AFRF-2000) Freqüências Acumuladas de Salários Anuais, em Milhares de Reais, da Cia. Alfa

Classes de Salário	Freqüências Acumuladas
(3 ; 6]	12
(6 ; 9]	30
(9 ; 12]	50
(12 ; 15]	60
(15 ; 18]	65
(18 ; 21]	68

Quer-se estimar o salário mediano anual da Cia. Alfa. Assinale a opção que corresponde ao valor aproximado desta estatística, com base na distribuição de freqüências.

a) 12,50 **b) 9,60** c) 9,00 d) 12,00 e) 12,10

Para efeito das quatro próximas questões, considere os seguintes dados:

DISTRIBUIÇÃO DE FREQÜÊNCIAS DAS IDADES DOS FUNCIONÁRIOS DA EMPRESA ALFA, EM 1º/1/90

Classes de Idades (anos)	Freqüências (fi)	Pontos Médios (PM)	$\frac{PM-37}{5}=di$	di.fi	$di^2.fi$	$di^3.fi$	$di^4.fi$
19,5 !--- 24,5	2	22	-3	-6	18	-54	162
24,5 !--- 29,5	9	27	-2	-18	36	-72	144
29,5 !--- 34,5	23	32	-1	-23	23	-23	23
34,5 !--- 39,5	29	37	---	---	---	---	---
39,5 !--- 44,5	18	42	1	18	18	18	18
44,5 !--- 49,5	12	47	2	24	48	96	192
49,5 !--- 54,5	7	52	3	21	63	189	567
Total	n=100			16	206	154	1106

11. (AFTN-96) Marque a opção que representa a mediana das idades dos funcionários em 1º/1/90.

a) 35,49 anos b) 35,73 anos c) 35,91 anos **d) 37,26 anos** e)38,01 anos

Para efeito das duas questões seguintes, sabe-se que o quadro de pessoal da empresa continua o mesmo em 1º/1/96.

13. (AFTN-96) Marque a opção que representa a mediana das idades dos funcionários em 1º/1/96.

a) 35,49 anos b) 36,44 anos c) 41,49 anos d) 41,91 anos **e) 43,26 anos**

Para a solução das duas próximas questões utilize o enunciado que segue. O atributo do tipo contínuo X, observado como um inteiro, numa amostra de tamanho 100 obtida de uma população de 1000 indivíduos, produziu a tabela de freqüências seguinte:

Classes	Freqüência (f)
29,5-39,5	4
39,5-49,5	8
49,5-59,5	14
59,5-69,5	20
69,5-79,5	26
79,5-89,5	18
89,5-99,5	10

18. (AFRF-2002.2) Assinale a opção que corresponde à estimativa da mediana amostral do atributo X.

a) 71,04 b) 65,02 c) 75,03 d) 68,08 e) 70,02

9 – Quartis, Decis e Centis (para dados agrupados em classes)

São separatrizes que dividem o conjunto de valores ou distribuição em 4 partes iguais, 10 partes iguais e 100 partes iguais.

Assim, teremos:

Quartis → Q_1, i = 1, 2, 3

Decis → D_1, i = 1, 2, 3...9

Centis → C_1, i = 1, 2, 3....99

Para fins de cálculo podemos fazer:

a) **Quartis**

Calculamos o elemento do quartil

$Eqi = \frac{in}{4}$ **,** i **=** 1, 2, 3.

Calculamos o quartil desejado:

$$Qi = L_i + \frac{\frac{i.n}{4} - f_{ant.ac.}}{fQi} \cdot c$$

Observação: Este cálculo é análogo ao usado para a determinação da mediana.

b) **Decis**

Calculamos o elemento do decil

Ed**i=** $\frac{in}{10}$ **, i = 1, 2, 3, 4, ... , 9**

Calculamos o decil desejado

$$Di = L_i + \frac{\frac{i.n}{10} - f_{ant.ac.}}{fDi} \cdot c$$

c) **Centis**

Calculamos o elemento do centil

E_{Ci} = $\frac{in}{100}$ **, i = 1, 2, 3....99.**

Calculamos o centil desejado

$$Ci = L_i + \frac{\frac{i.n}{100} - f_{ant.ac.}}{fCi} \cdot c$$

Exemplo:

Ao aplicar uma prova de Estatística a uma turma de 120 alunos, encontrou-se o resultado da tabela a seguir:

CLASSES	NÚMERO DE ALUNOS	f_{ac}
30 ├ 40	1	1
40 ├ 50	3	4
50 ├ 60	11	15
60 ├ 70	21	36
70 ├ 80	43	79
80 ├ 90	32	111
90 ├ 100	9	120
TOTAL	120	

Calcule:

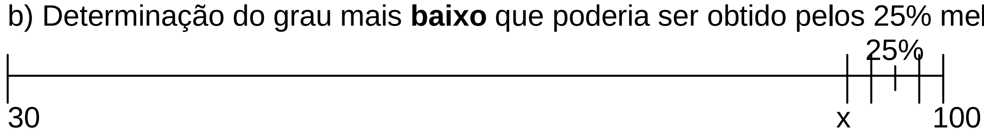
a) os quartis da distribuição dada;
b) o grau mais baixo que poderia ser obtido pelos 25% melhores alunos da turma;
c) o grau mais alto que é possível ser obtido pelos 20% piores alunos da turma.

Solução:

a) Determinação dos quartis:

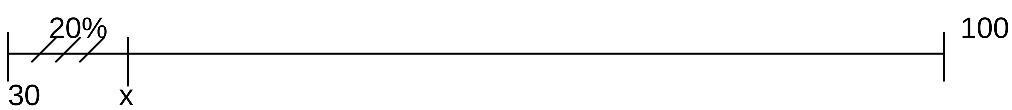
1º quartil → Q_1 = 67,14 aproximadamente
2º quartil → Q_2 = 75,58 aproximadamente
3º quartil → Q_3 = 83,43 aproximadamente

b) Determinação do grau mais **baixo** que poderia ser obtido pelos 25% melhores alunos da turma.



Então, x é o valor do 3º quartil = 83,43

c) Determinação do grau mais **alto** que é possível ser obtido pelos 20% piores alunos da turma.



Então, **x** é o valor do 2º Decil (D_2) ou vigésimo percentil (C_{20}).

Para calcular D_2.

Então, D_2 = 64,28 ou 64 aproximadamente.

TESTE SOBRE DECIL

Em um ensaio para o estudo da distribuição de um atributo financeiro (X) foram examinados 200 itens de natureza contábil do balanço de uma empresa. Esse exercício produziu a tabela de freqüências abaixo. A coluna *Classes* representa intervalos de valores de X em reais e a coluna *P* representa a freqüência relativa acumulada. Não existem observações coincidentes com os extremos das classes.

Classes	P (%)
70-90	5
90-110	15
110-130	40
130-150	70
150-170	85
170-190	95
190-210	100

17. (AFRF-2002) Assinale a opção que corresponde à estimativa do quinto decil da distribuição de X.

a) 138,00 b) 140,00 c) 136,67 d) 139,01 e) 140,66

MEDIDAS DE DISPERSÃO

1 – Amplitude Total (A_t)

É igual a diferença entre o maior valor e o menor valor dos dados apresentados.
Na sua determinação podemos considerar as seguintes situações:

- dados não tabulados:

X = {①, 3, 5, 7, ⑨}

$A_t = 9 - 1 = 8 \rightarrow A_t = 8$

- dados tabulados não agregados em classes (dados discretos):

x_i	f_i
1	10
3	20
5	40
7	20
9	10

$A_t = 9 - 1 = 8 \rightarrow A_t = 8$

- dados tabulados agregados em classes (dados contínuos):

CLASSES	f_i
2 ├ 4	10
4 ├ 6	20
6 ├ 8	40
8 ├ 10	20
10 ├ 12	10

Admite-se duas soluções:

1ª) diferença entre o limite superior da última classe e o limite inferior da 1ª classe. No caso, teremos:

A_t = 12 – 2 = 10 → A_t = 10

2ª) diferença entre o ponto médio da última classe e o ponto médio da 1ª classe

A_t = 11 – 3 = 8 → A_t = 8

2 – Desvio Médio Absoluto (DMA)

$$DM = \frac{\Sigma|d_i| \cdot f_i}{\Sigma f_i}$$

É igual a média dos valores absolutos dos desvios calculados em relação à média do conjunto de valores.

- dados não tabulados:

X = {1, 3, 5, 7, 9}, $\bar{x}$ = 5 e n = 5

x_i	$d_i = x_i - \bar{x}$	$\|d_i\|$
1	1 – 5 = – 4	4
3	3 – 5 = – 2	2
5	5 – 5 = 0	0
7	7 – 5 = 2	2
9	9 – 5 = 4	4
SOMA		12

Assim o DMA = $\frac{12}{5}$ = 2,4 → DMA = 2,4

- dados tabulados não agregados em classes (dados discretos):

x_i	f_i	$d_i = x_i - \bar{x}$	$\|d_i\|$	$\|d_i\| f_i$
1	10	– 4	4	40
3	20	– 2	2	40
5	40	0	0	0
7	20	2	2	40
9	10	4	4	40
TOTAL	100			Σ = 160

Assim o DMA = $\frac{160}{100}$ = 1,6 → DMA = 1,6

Assim de uma maneira geral o desvio médio é calculado por: $DM = \frac{\Sigma|d_i| \cdot f_i}{\Sigma f_i}$

- dados tabulados agrupados em classes (dados contínuos):

O seu cálculo é análogo ao dos dados discretos, devendo iniciá-lo pela determinação dos pontos médios das classes, que assim representarão os valores x_i do conjunto de dados.

3 – Amplitude Semi-Interqualítica (desvio quarlítico)

É a metade da diferença entre o terceiro quartil (Q_3) e o primeiro quartil (Q_1).

$$Dq = \frac{Q_3 - Q_1}{2}$$

4 – Desvio Padrão (S)

É a medida de dispersão mais usada. É a raiz quadrada da média aritmética dos quadrados dos desvios.

$s = \sqrt{\frac{\Sigma d_i^{\,2}}{n}}$ para dados não tabulados

$s = \sqrt{\frac{\Sigma d_i^{\,2} \cdot f_i}{n}}$ para dados tabulados

Desvio Padrão Populacional

Uma outra fórmula para o desvio padrão populacional:

$$s = \sqrt{\frac{1}{n} \times \left[\sum xi^2 - \frac{\left(\sum xi\right)^2}{n} \right]}$$

Onde:

$$d_i = x_i - \overline{X}$$

$s = \sqrt{\frac{\Sigma d_i^{\,2}}{n-1}}$ para dados não tabulados

$s = \sqrt{\frac{\Sigma d_i^{\,2} \cdot f_i}{n-1}}$ para dados tabulados

Desvio Padrão Amostral **(Mais Usado)**

Uma outra fórmula para o desvio padrão amostral:

$$s = \sqrt{\frac{1}{n-1} \times \left[\sum xi^2 - \frac{\left(\sum xi\right)^2}{n} \right]}$$

5 – Propriedades do Desvio Padrão

1ª) Propriedade Somando-se (ou subtraindo-se) a cada elemento de um conjunto de valores uma constante arbitrária, o desvio padrão **não se altera**.
Exemplo:

$X = \{ x_1, x_2, x_3, x_4, x_5\}$ com $\bar{x} = a$

Se somarmos (ou subtrairmos) um valor constante, o grau de dispersão permanece o mesmo e o desvio padrão, que é uma medida de dispersão, utilizará o mesmo número para medi-la.

$X = \{ 6, 8, 10, 10 \}$, com $S_x = 1{,}9$

Somando 100 a cada elemento, teremos o conjunto:

$Y = \{ 106, 108, 110, 110 \}$, cujo desvio padrão será o mesmo 1,9 → $S_y = 1{,}9$

2ª)Propriedade Multiplicando-se (ou dividindo-se) cada elemento de um conjunto de valores por um valor constante, arbitrário e diferente de zero, o desvio padrão fica multiplicado (ou dividido) por esta constante.

Exemplo:
$X = \{6, 8, 10, 10\}$, com $S_x = 1{,}9$

Se multiplicarmos por 10 cada elemento, teremos o conjunto:
$Y = \{60, 80, 100, 100\}$, cujo desvio padrão S_y será obtido pela multiplicação do desvio padrão S_x por 10.
Assim $S_y = S_x = 1{,}9 \times 10 = 19$ → $S_y = 19$.

6 – Variância (S^2)

É a média aritmética do quadrado dos desvios.

$$s = \frac{\Sigma d_i^{\,2}}{n}$$ para dados não tabulados

$$s = \frac{\Sigma d_i^{\,2} \cdot f_i}{n}$$ para dados tabulados

Variância Populaciomal

Uma outra fórmula para variância populacional:

$$s = \frac{1}{n} \times \left[\sum xi^2 - \frac{\left(\sum xi\right)^2}{n} \right]$$

Onde:

$$d_i = x_i - \overline{X}$$

$s = \frac{\Sigma d_i^{\,2}}{n-1}$ para dados não tabulados

$s = \frac{\Sigma d_i^{\,2} \cdot f_i}{n-1}$ para dados tabulados

Variância Amostral **(Mais Usado)**

Uma outra fórmula para variância amostral:

$$s = \frac{1}{n-1} \times \left[\sum xi^2 - \frac{\left(\sum xi\right)^2}{n} \right]$$

7 – Propriedades da Variância

Observa-se que a 1ª propriedade é idêntica à do desvio padrão, quanto à soma (ou subtração), no entanto, a **2ª propriedade** (multiplicação ou divisão) a apresenta a seguinte modificação: "**a variância fica multiplicada (ou dividida) pelo quadrado da constante"**.

8 – Dispersão Relativa – Coeficiente de Variação

Dentre as medidas de dispersão relativa o coeficiente de variação mais utilizado é o de Pearson (CV_P), que é o quociente entre o desvio padrão e a média aritmética do conjunto de dados.

$$CV_p = \frac{S}{\overline{X}}$$

Quanto menor for este valor, mais **homogêneo** será o conjunto de dados.

TESTES SOBRE MEDIDAS DE DISPERSÃO

1. (AFC-94) Entre os funcionários de um órgão do governo, foi retirada uma amostra de dez indivíduos. Os números que representam as ausências ao trabalho registradas para cada um deles, no último ano, são: 0, 0, 0, 2, 2, 2, 4, 4, 6, e 10. Sendo assim, o valor do desvio padrão desta amostra é:

a) $\sqrt{3}$ b) $\sqrt{9}$ c) $\sqrt{10}$ d) $\sqrt{30}$

2. (AFC-94) Uma empresa que possui 5 máquinas copiadoras registrou em cada uma delas no último mês (EM 1000 UNIDADES) : 20, 23, 25, 27 e 30 cópias, respectivamente. O valor da variância desta população é:

a) 5 b) 11,6 c) 14,5 d) 25

3. (AFC-94) A média e a variância do conjunto dos salários pagos por uma empresa eram de R$ 285 000,00 e $1{,}1627 \times 10^{10}$ respectivamente. O valor da variância do conjunto dos salários, após o corte de três zeros na moeda é:

a) $1{,}1627 \times 10^{7}$ b) $1{,}1627 \times 10^{6}$ c) $1{,}1627 \times 10^{5}$ d) $1{,}1627 \times 10^{4}$

4. (AFC-94) Seja X uma variável aleatória com média aritmética $\overline{X}$ =10 e desvio-padrão S=3. Considere as variáveis: y=2x+1 e z=2x. A única afirmação errada é:

a) as variáveis y e z tem a mesma média aritmética;
b) o desvio-padrão de y é 6;
c) as variáveis y e z têm o mesmo desvio-padrão;
d) a média de y é 21;
e) as variáveis x e z têm o mesmo coeficiente de variação.

5. (BACEN-94) Em certa empresa, o salário médioera de R$ 90 000,00 e o desvio-padrão dos salários era de R$ 10 000,00. Todos os salários receberam um aumento de 10%. O desvio-padrão dos salários passou a ser de:

a) 10 000,00 b) 10 100,00 c) 10 500,00 d) 10 900,00 e) 11 000,00

6. (TCU-91) Doze fichas de funcionários de uma empresa foram selecionadas ao acaso; foram anotados os números de dependentes, na ordem de seleção, a saber: 3, 0, 5, 2, 3, 6, 4, 1, 3, 2, 4, e 3. Para a variável número de dependentes, resolva a expressão: "média+moda+mediana+variância+1,5".

R: 13

7. (TCU-93) O quadro abaixo apresenta a renda mensal per capitã das localidades A e B:

LOCALIDADE	MÉDIA	DESVIO-PADRÃO
A	50	10
B	75	15

Assinale a opção correta.
a) o intervalo semi-interquartílico é dado por [10,15];
b) a renda da localidade A é mais homogênea que a renda na localidade B;
c) o coeficiente de variação é $\frac{50}{75}$;
d) a renda da localidade B é mais homogênea que a renda na localidade A;
e) os coeficientes de variação de renda nas localidades A e B são iguais.

(AFC-94) Para a solução das três próximas questões, considere os dados da tabela abaixo que representa a distribuição de freqüências das notas em uma prova de estatística aplicada em três turmas de 100 alunos cada.

Classes de Notas	Freqüência das Notas na Prova de Estatística		
	TURMA 01	TURMA 02	TURMA 03
0 ├ 2	20	10	5
2 ├ 4	40	15	10
4 ├ 6	30	50	70
6 ├ 8	6	15	10
8 ├ 10	4	10	5
TOTAL	100	100	100

8. (AFC-94) Assinale a afirmação correta.

a) Moda T2 <Moda T3;
b) Média T1 > Média T2;
c) Média T2 < Média T3;
d) Mediana T1 < Mediana T2;
e) Mediana T2 > Mediana T3

9. (AFC-94) A única opção errada é:

a) 1º quartil T1 > 1º quartil T3;
b) desvio-padrão T2 > desvio-padrão T3;
c) coeficiente de variação T2 > coeficiente de variação T3;
d) na turma 3: média= mediana = moda.

10. (AFC-94) A distribuição de notas é simétrica em relação a média aritmética:

a) nas três turmas;
b) nas turmas 1 e 2;
c) nas turmas 1 e 3;
d) somente na turma 1;
e) nas turmas 2 e 3.

11. (FISCAL DE TRIBUTOS DE MG-96) No conjunto de dados A={3, 5, 7, 9, 11}, o valor do desvio médio é:

a) 2,1 b) 2,4 c) 2,6 d) 3,2 e) 3,6

12. (FISCAL DE TRIBUTOS DE MG-96) O desvio-padrão do conjunto de dados A={2, 4, 6, 8, 10}, é aproximadamente:

a) 2,1 b) 2,4 c) 2,8 d) 3,2 e) 3,6

28. (AFTN-98) Os dados seguintes, ordenados do menor para o maior, foram obtidos de uma amostra aleatória, de 50 preços (Xi) de ações, tomada numa bolsa de valores internacional. A unidade monetária é o dólar americano.
4, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 8, 8, 8, 8, 8, 8, 8, 8, 8, 9, 9, 9, 9, 9, 9, 10, 10,10, 10, 10, 10, 10, 10, 11, 11, 12, 12, 13, 13,14, 15, 15, 15, 16, 16, 18, 23
Os valores seguintes foram calculados para a amostra:

$$\sum Xi = 490 \text{ e } \sum Xi^2 - \frac{\left(\sum Xi\right)^2}{50} = 668$$

Assinale a opção que corresponde à mediana e à variância amostral, respectivamente (com aproximação de uma casa decimal).

a) (9,0 13,6) b) (9,5 14,0) c) (8,0 15,0) d) (8,0 13,6) e) (9,0 14,0)

39. (TJ CE – 2002) Aplicando a transformação z = (x - 14)/4 aos pontos médios das classes (x) obteve-se o desvio padrão de 1,10 salários mínimos. Assinale a opção que corresponde ao desvio padrão dos salários não transformados.

a) 6,20 b) 4,40 c) 5,00 d) 7,20 e) 3,90

19. (AFRF-2000) Numa amostra de tamanho 20 de uma população de contas a receber, representadas genericamente por X, foram determinadas a média amostral M = 100 e o desvio-padrão S = 13 da variável transformada (X-200)/5. Assinale a opção que dá o coeficiente de variação amostral de X.

a) 3,0% b) 9,3% c)17,0% d) 17,3% e)10,0%

22. (AFC-94) Entre os funcionários de um órgão do governo, foi retirada uma amostra de dez indivíduos. Os números que representam as ausências ao trabalho registradas para cada um deles, no último ano, são: 0, 0, 0, 2, 2, 2, 4, 4, 6 e 10. Sendo assim, o valor do desvio padrão desta amostra é:

a) $\sqrt{3}$ b) $\sqrt{9}$ c) $\sqrt{10}$ d) $\sqrt{30}$

27. (FTE MG-96) No conjunto de dados A={3, 5, 7, 9, 11}, o valor do desvio médio é:

a) 2,1 b) 2,4 c) 2,6 d) 2,8 e) 3,1

37. (FISCAL INSS - 2002) Dada a seqüência de valores 4, 4, 2, 7 e 3 assinale a opção que dá o valor da variância. Use o denominador 4 em seus cálculos.

a) 5,5 b) 4,5 c) 3,5 d) 6,0 e) 16,0

39. (TJ CE – 2002) Aplicando a transformação z = (x - 14)/4 aos pontos médios das classes (x) obteve-se o desvio padrão de 1,10 salários mínimos. Assinale a opção que corresponde ao desvio padrão dos salários não transformados.

a) 6,20 b) 4,40 c) 5,00 d) 7,20 e) 3,90

36. (FTE-PA-2002) Um certo atributo W, medido em unidades apropriadas, tem média amostral 5 e desvio-padrão unitário. Assinale a opção que corresponde ao coeficiente de variação, para a mesma amostra, do atributo Y = 5 + 5W.

a) 16,7% b) 20,0% c) 55,0% d) 50,8% e) 70,2%

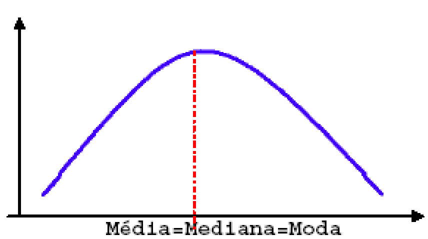
40. (AFRF–2003) O atributo Z=(X-2)/3 tem média amostral 20 e variância amostral 2,56. Assinale a opção que corresponde ao coeficiente de variação amostral de X.

a) 12,9% b) 50,1% c) 7,7% d) 31,2% e) 10,0%

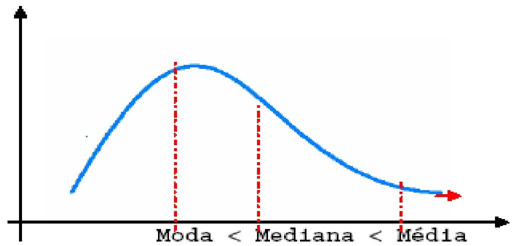
9 – Assimetria

Estuda a forma do gráfico de uma distribuição. Esta forma depende dos valores da média e moda. Assim, em uma distribuição em forma de sino, temos:

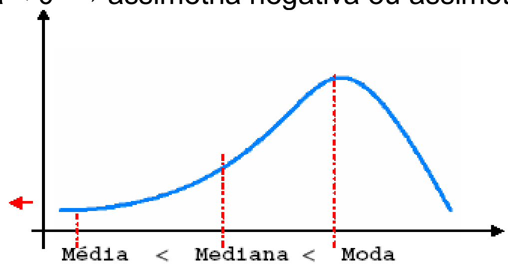
- Média – Moda = 0 → assimetria nula ou simétrica



- Média – Moda > 0 → assimetria positiva ou assimétrica à direita



- Média – Moda < 0 → assimetria negativa ou assimétrica à esquerda



9 – Coeficientes de Assimetria:

Índice Quartílico de assimetria $$As = \frac{Q_3 + Q_1 - 2md}{Q_3 - Q_1}$$

Índice Percentílico de assimetria $$As = \frac{P_{90} + P_{10} - 2.P_{50}}{P_{90} - P_{10}}$$

1º coeficiente de Pearson $$A = \frac{\overline{X} - Mo}{S}$$

2º coeficiente de Pearson $$As = \frac{3 \cdot (\bar{x} - Md)}{S}$$

Se $0{,}15 < |As| < 1$, a assimetria é considerada moderada; se As>1, é considerada forte.

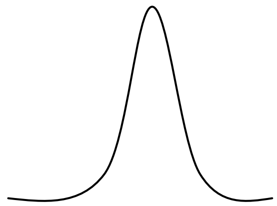
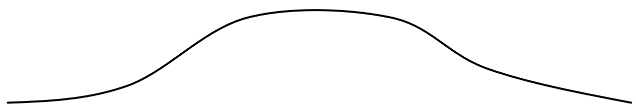
- As = 0 a assimetria é nula
- As > 0 a assimetria é positiva
- As < 0 a assimetria é negativa

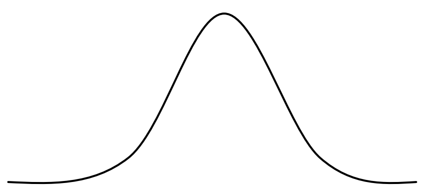
No segundo coeficiente de Pearson observa-se ainda que:

- Se X = Md a assimetria é nula, ou seja, a curva é simétrica.
- Se x > Md a assimetria é positiva, ou seja, a curva é assimétrica à direita.
- Se x < Md a assimetria é negativa, ou seja, a curva é assimétrica à esquerda.

10 – Curtose:

Denominamos **curtose** o grau de achatamento de uma distribuição em relação a uma distribuição padrão, denominada **curva normal.**

- Quando a distribuição apresenta uma curva de freqüência mais fechada que a normal, ela recebe o nome de Leptocúrtica.
- 
- Quando a distribuição apresenta uma curva de freqüência mais aberta que a normal, ela recebe o nome de Platicúrtica.
- 
- Quando a distribuição apresenta uma curva de freqüência normal, ela recebe o nome de Mesocúrtica.



11 – Coeficiente de Curtose:

$$C = \frac{Q_3 - Q_1}{2 \cdot (P_{90} - P_{10})} \quad \text{ou} \quad C = \frac{Q_3 - Q_1}{2 \cdot (D_9 - D_1)}$$

Assim temos:

- Se C = 0,263 $\rightarrow$ curva mesocúrtica;
- Se C < 0,263 $\rightarrow$ curva leptocúrtica;
- Se C > 0,263 $\rightarrow$ curva platicúrtica.

TESTES SOBRE ASSIMETRIA E CURTOSE

47. (TCU-93) Os montantes de venda a um grupo de clientes de um supermercado forneceram os seguintes sumários: média aritmética=$1,20 , mediana=$0,53 e moda=$0,25. Com base nestas informações, assinale a opção correta:

a) A distribuição é assimétrica à direita.
b) A distribuição é assimétrica à esquerda.
c) A distribuição é simétrica.
d) Entre os três indicadores de posição apresentados, a média aritmética é a melhor
medida de tendência central.
e) O segundo quartil dos dados acima é dado por $0,25.

45. (AFRF-2002) Entende-se por curtose de uma distribuição seu grau de achatamento em geral medido em relação à distribuição normal. Uma medida de curtose é dada pelo quociente

$$k = \frac{Q}{P_{90} - P_{10}}$$

onde Q é a metade da distância interquartílica e P_{90} e P_{10} representam os percentis de 90% e 10%, respectivamente. Assinale a opção que dá o valor da curtose ê para a distribuição de X.

a) 0,263 b) 0,250 c) 0,300 d) 0,242 e) 0,000

49. (FISCAL DO INSS-2002) A tabela abaixo dá a distribuição de freqüências de um atributo X para uma amostra de tamanho 66. As observações foram agrupadas em 9 classes de tamanho 5. Não existem observações coincidentes com os extremos das classes.

Classes	Freqüências
4-9	5
9-14	9
14-19	10
19-24	15
24-29	12
29-34	6
34-39	4
39-44	3
44-49	2

Sabe-se que o desvio padrão da distribuição de X é aproximadamente 10. Assinale a opção que dá o valor do coeficiente de assimetria de Pearson que é baseado na média, na mediana e no desvio padrão.

a) -0,600 b) 0,191 c) 0,709 d) 0,603 e) -0,610

51. (FISCAL DO INSS-2002) Considere a tabela de freqüências seguinte correspondente a uma amostra da variável X. Não existem observações coincidentes com os extremos das classes.

Classes	Freqüências Acumuladas (%)
2.000 – 4.000	5
4.000 – 6.000	16
6.000 – 8.000	42
8.000 – 10.000	77
10.000 – 12.000	89
12.000 – 14.000	100

Assinale a opção que corresponde ao valor do coeficiente de assimetria percentílico da amostra de X, baseado no 1º, 5º e 9º decis.

a) 0,024 b) 0,300 c) 0,010 d) -0,300 e) -0,028

46. (AFRF-2002.2) Para a solução da próxima questão utilize o enunciado que segue. O atributo do tipo contínuo X, observado como um inteiro, numa amostra de tamanho 100 obtida de uma população de 1000 indivíduos, produziu a tabela de freqüências seguinte:

Classes	Freqüência (f)
29,5-39,5	4
39,5-49,5	8
49,5-59,5	14
59,5-69,5	20
69,5-79,5	26
79,5-89,5	18
89,5-99,5	10

Assinale a opção que dá o valor do coeficiente quartílico de assimetria.

a) 0,080 b) -0,206 c) 0,000 d) -0,095 e) 0,300

44. (AFRF-2002) Em um ensaio para o estudo da distribuição de um atributo financeiro (X) foram examinados 200 itens de natureza contábil do balanço de uma empresa. Esse exercício produziu a tabela de freqüências abaixo. A coluna *Classes* representa intervalos de valores de X em reais e a coluna *P* representa a freqüência relativa acumulada. Não existem observações coincidentes com os extremos das classes.

Classes	P (%)
70-90	5
90-110	15
110-130	40
130-150	70
150-170	85
170-190	95
190-210	100

Seja S o desvio padrão do atributo X. Assinale a opção que corresponde à medida de assimetria de X como definida pelo primeiro coeficiente de Pearson.

a) 3/S b) 4/S c) 5/S d) 6/S e) 0

NÚMEROS ÍNDICES

1 - Conceito

Os números índices (ou apenas índices) são indicadores que medem alterações entre grandezas do mesmo tipo ou variações entre grandezas diferentes e aplicam-se no campo da produção, evolução dos preços, custo de vida, salários, registos demográficos, etc.
Como medem variações no tempo e no espaço, permitem sintetizar e apresentar de forma eficaz a natureza das alterações numa ou várias variáveis, sendo mais fácil identificar flutuações referentes a sub-períodos que se repetem ao longo do tempo.
Os números índices podem ser usados para vários propósitos tais como para medir as variações de preços de bens de consumo, a quantidade física de mercadorias produzidas ou vendidas; ou podem relacionar-se à conceitos tais como produtividade, eficiência inteligência entre outras.

2 - Índice relativo

É a relação entre o preço de uma única utilidade, em período determinado, e o de outro período, denominado básico ou de referência.

2.1- Índice relativo Simples

Representado por **0** o período base e por **n** o período atual temos:

- Relativo de preços: $Ip = \frac{p_n}{p_0} \cdot 100$
- Relativo de quantidades $Iq = \frac{q_n}{q_0} \cdot 100$
- Relativo de Valor $Iv = \frac{p_n \cdot q_n}{p_0 \cdot q_0} \cdot 100$

Onde:
p_0 = preço no período base
p_n = preço no período atual (período que se deseja o índice)
q_0 = quantidade no período base
q_n = quantidade no período atual (período que se deseja o índice)

Exemplo1:
O preço de certo produto, em 1998, era de R$ 12,00, em 2000 era de R$ 13,80.
Tomando como base o ano de 1998, calcule o índice relativo de aumento de preço para o ano de 2000.

$$Ip = \frac{p_n}{p_0} \cdot 100 \Rightarrow \frac{13{,}80}{12{,}00} \cdot 100 \Rightarrow 1{,}15 \cdot 100 \Rightarrow 115\%$$

Portanto houve um aumento de 15%

Exenplo2:

Uma empresa produziu 45 toneladas de aço em 1998, e 68 toneladas em 1999. qual a quantidade relativa em 1999?

$$Iq = \frac{q_n}{q_0} \cdot 100 \Rightarrow \frac{68}{45} \cdot 100 \Rightarrow 1{,}51 \cdot 100 \rightarrow 151\%$$

Portanto houve um aumento na produção de 51% com relação a 1998.

Exemplo3:

Uma empresa vendeu, em 1970, 1000 unidades de um artigo ao preço unitário de Cr$ 500,00. Em 1971 vendeu 2000 unidades do mesmo artigo ao preço unitário de Cr$ 600,00.O valor relativo da venda em 1971 foi:

$$Iv = \frac{600 \cdot 2000}{500 \cdot 1000} \cdot 100 \Rightarrow \frac{6 \cdot 2}{5 \cdot 1} \cdot 100 \Rightarrow 2{,}4 \cdot 100 \Rightarrow 240$$

Em 1971, o valor das vendas foi 140% superior ao de 1970.

Propriedade Circular: $I_{0,T} = I_{0,1} \cdot I_{1,2} \cdot I_{2,3} \cdot I_{3,4} \cdot ...I_{T-1,T}$

Ex.: Um artigo foi vendido em fevereiro com um acréscimo de 10% em relação a janeiro, em março com acréscimo de 5% em relação a fevereiro, e em abril com uma queda de 4% em relação a março.calcule o percentual de preço em abril referente a janeiro.

$I_{jan,abr} = I_{jan,fev} \cdot I_{fev,mar} \cdot I_{mar,abr}$ → 1,1 . 1,05 . 0,96 →1,1088 ou seja 110,88%

Reversão do Fatores: É quando o índice de preço **vezes (×)** o índice de quantidade é igual ao índice de valor.

$$Ip \times Iq = Iv \qquad \frac{Pt}{Po} \times \frac{Qt}{Qo} = \frac{Pt \times Qt}{Po \times Qo} = \frac{Vt}{Vo}$$

OBS.: A reversão de fatores não se aplica aos índices agregativos ponderados de Laspeyre e Paasche.

Mudança de Base: Dividimos cada índice da série original pelo número índice da nova base. Esse processo nos dá uma aproximação bastante satisfatória.

Ex.:

Índice Base 1957

1957	1958	1959	1960	1961	1962
100	104	97	112	120	124

Mudar a base para 1960:

Solução: É só dividir todos os índices por 112 (índice da nova base)

3 – Emprego da Média Aritmética Simples

- De relativo de preços:

$$\overline{P}_{0,t} = \frac{\Sigma p_{0,t}}{n}$$

Onde: $p_{0,t} = \frac{p_t}{p_0}$ e n = número de preços

- De relativo de quantidade:

$$\overline{Q}_{0,t} = \frac{\Sigma q_{0,t}}{n}$$

Onde: $q_{0,t} = \frac{q_t}{q_0}$ e n = número de quantidade

4 – Emprego da Média Harmônica Simples

- De relativo de preços: $\overline{P}^{h}{}_{0,t} = \frac{n}{\Sigma\left(\frac{1}{p_{0,t}}\right)}$

- De relativo de quantidade: $\overline{Q}^{h}{}_{0,t} = \frac{n}{\Sigma\left(\frac{1}{q_{0,t}}\right)}$

5 – Emprego da Média Geométrica Simples

- De relativo de preços: $\overline{P}^{g}{}_{0,t} = \sqrt{\prod p_{0,t}}$

- De relativo de quantidade: $\overline{Q}^{g}{}_{0,t} = \sqrt{\prod q_{0,t}}$

6 – EMPREGO DE ÍNDICES (AGREGATIVOS) PONDERADOS

Como vimos, os índices simples apresentam algumas desvantagens, em especial à se refere à inexistência de pesos diferentes para cada utilidade que os compõe de acordo com sua importância relativa. No caso dos índices ponderados, além da fórmula a ser usada para interpretar as variações de preço e de quantidade dos bens, há o problema do critério para a fixação dos pesos relativos de cada um deles. A ponderação proposta pelos métodos mais usados baseia-se na participação de cada bem no valor transacionado total e é feita, em geral, segundo dois critérios: peso fixo na época básica ou peso variável na época atual.

6.1 - Índice de Laspeyres ou Método da época Básica

- De preço: $L^{p}{}_{0,t} = \dfrac{\sum p_t q_0}{\sum p_0 q_0}$

- De quantidade: $L^{q}{}_{0,t} = \dfrac{\sum p_0 q_t}{\sum p_0 q_0}$

6.2 - Índice de Paasche ou Método da Época Atual

- De preço: $P^{p}{}_{0,t} = \dfrac{\sum p_t q_t}{\sum p_0 q_t}$

- De quantidade: $P^{q}{}_{0,t} = \dfrac{\sum q_t p_t}{\sum q_0 p_t}$

Veja o exemplo a seguir:

Considerando 1970 como base,determinar um índice de preço e de quantidade, usando o índice de Laspeyres e Paasche.

Artigos	1970		1971	
	Preço	Qtde	Preço	Qtde
A	2	4	2	5
B	3	3	4	2
C	5	2	6	5

Solução:

Laspeyres:

Índice de Preço: $L^{p}{}_{0,t} = \dfrac{\sum p_{71} q_{70}}{\sum p_{70} q_{70}} \Rightarrow \dfrac{(2.4)+(4.3)+(6.2)}{(2.4)+(3.3)+(5.2)} \Rightarrow \dfrac{32}{27} \Rightarrow 1{,}185$ ou 118,5%

Isto significa que houve uma variação no preço de 18,5% no decorrer desses anos.

Índice de Quantidade: $L^{q}{}_{0,t} = \dfrac{\sum p_{70} q_{71}}{\sum p_{70} q_{70}} \Rightarrow \dfrac{(5.2)+(2.3)+(5.5)}{(2.4)+(3.3)+(5.2)} \Rightarrow \dfrac{41}{27} \Rightarrow 1{,}518$ ou 151,8%

Isto significa que houve uma variação na quantidade de 51,8% no decorrer desses anos.

Paasche:

Índice de Preço: $P^{p}{}_{0,t} = \frac{\sum p_{71}q_{71}}{\sum p_{70}q_{71}} \Rightarrow \frac{(2.5)+(4.2)+(6.5)}{(2.5)+(3.2)+(5.5)} \Rightarrow \frac{48}{41} \Rightarrow 1{,}171$ ou 117,1%

Isto significa que houve uma variação no preço de 17,1% no decorrer desses anos.

Índice de Quantidade: $P^{q}{}_{0,t} = \frac{\sum q_{71}p_{71}}{\sum q_{70}p_{71}} \Rightarrow \frac{(2.5)+(4.2)+(6.5)}{(4.2)+(3.4)+(2.6)} \Rightarrow \frac{48}{27} \Rightarrow 1{,}777$ ou 177,7%

Isto significa que houve uma variação na quantidade de 77,7% no decorrer desses anos.

6.3 - Índice de Fischer (Índice Ideal)

O índice de Fischer, também conhecido corno forma ideal, é a média geométrica dos números-índices de Laspeyres e de Paasche. Sob o aspecto da ponderação, esse índice envolve os dois sistemas anteriormente adotados. A proposta de Fischer fundamenta-se no fato de os índices que o compõem não atenderem ao critério de decomposição das causas, além de um deles tender a superestimar e o outro a subestimar o verdadeiro valor do índice. Esse verdadeiro valor tenderá a ser um número superior ao fornecido pela fórmula de Paasche e inferior ao apresentado pela fórmula de Laspeyres, o que acontece com a média geométrica entre esses dois índices. Entretanto, o índice de Fischer, apesar de ser chamado de ideal, nisso pode ser considerado "perfeito". A necessidade de modificar pesos, em dada época comparada, em decorrência do cálculo do índice de Paasche, constitui uma restrição não desprezível ao seu emprego. Além disso, não parece ser possível determinar especificamente o que o índice de Fischer mede, bem como estabelecer o verdadeiro valor de um Índice perfeito, o qual serviria de elemento de referência.

- Índice de Preço: $I^{p}{}_{0,t} = \sqrt{L^{p}{}_{0,t} \times P^{p}{}_{0,t}}$

- Índice de Quantidade: $I^{q}{}_{0,t} = \sqrt{L^{q}{}_{0,t} \times P^{q}{}_{0,t}}$

Com base no exemplo anterior vamos calcular o índice ideal (Fischer) de preço e de quantidade:

- Índice de Preço: $I^{p}{}_{0,t} = \sqrt{1{,}185 \times 1{,}171} \Rightarrow \sqrt{1{,}387635} \Rightarrow 1{,}178$ ou 117,8%
  Isto significa que houve uma variação no preço de 17,8% no decorrer desses anos.

- Índice de Quantidade: $I^{q}{}_{0,t} = \sqrt{151{,}8 \times 177{,}8} \Rightarrow \sqrt{26990{,}04} \Rightarrow 164{,}28\%$
  Isto significa que houve uma variação na quantidade de 64,28% no decorrer desses anos.

TESTES SOBRE NÚMEROS ÍNDICES

60. (AFRF-2003) Dadas as três séries de índices de preços abaixo, assinale a opção correta.

Ano	S1	S2	S3
1999	50	75	100
2000	75	100	150
2001	100	125	200
2002	150	175	300

a) As três séries mostram a mesma evolução de preços.
b) A série S2 mostra evolução de preços distinta das séries S1 e S3.
c) A série S3 mostra evolução de preços distinta das séries S1 e S2.
d) A série S1 mostra evolução de preços distinta das series S2 e S3.
e) As três séries não podem ser comparadas pois têm períodos-base distintos.

55. (AFTN-1998) A tabela seguinte dá a evolução de um índice de preço calculado com base no ano de 1984.

Ano	1981	1982	1983	1984	1985	1986
Índice	75	88	92	100	110	122

No contexto da mudança de base do índice para 1981 assinale a opção correta:

a) Basta dividir a série de preços pela média entre 0,75 e 1,00
b) Basta a divisão por 0,75 para se obter a série de preços na nova base
c) Basta multiplicar a série por 0,75 para se obter a série de preços na nova base
d) O ajuste da base depende do método utilizado na construção da série de preços, mas a divisão por 0,75 produz uma aproximação satisfatória.
e) Basta multiplicar a série de preços pela média entre 0,75 e 1,00

Para efeito das duas próximas questões, considere os seguintes dados:

Artigos	Quantidades (1000t)			Preços (R$/t)		
	1993	1994	1995	1993	1994	1995
A1	12	13	14	58	81	109
A2	20	25	27	84	120	164

52. (AFTN-1996) Marque a opção que representa os índices de Laspeyres de preços, no período de 1993 a 1995, tomando por base o ano de 1993.

a) 100,0; 141,2; 192,5
b) 100,0; 141,4; 192,8
c) 100,0; 141,8; 193,1
d) 100,0; 142,3; 193,3
e) 100,0; 142,8; 193,7

53. (AFTN-1996) Marque a opção que representa os índices de Paasche de preços, no período de 1993 a 1995, tomando por base o ano de 1993.

a) 100,0; 141,3; 192,3
b) 100,0; 141,6; 192,5
c) 100,0; 141,8; 192,7
d) 100,0; 142,0; 193,3
e) 100,0; 142,4; 193,6

56. (AFRF-2000) Uma empresa produz e comercializa um determinado bem X. A empresa quer aumentar em 60% seu faturamento com X. Pretende atingir este objetivo aumentando o preço do produto e a quantidade produzida em 20%. Supondo que o mercado absorva o aumento de oferta e eventuais acréscimos de preço, qual seria o aumento de preço necessário para que a firma obtenha o aumento de faturamento desejado?

a) 25,3% b) 20,5% c) 33,3% d) 40,0% e) 35,6%

(AFRF-2002/2) No tempo t0+2 o preço médio de um bem é 30% maior do que em t0+1, 20% menor do que em t0 e 40% maior do que em t0+3. Assinale a opção que dá o relativo de preços do bem em t0+3 com base em t0+1.

a) 162,5% d) 092,9%
b) 130,0% e) 156,0%
c) 120,0%

57. (AFRF-2000) Um índice de preços com a propriedade circular, calculado anualmente, apresenta a seqüência de acréscimos ä1 = 3 %, ä2 = 2% e ä3 = 2 %, medidos relativamente ao ano anterior, a partir do ano to . Assinale a opção que corresponde ao aumento de preço doperíodo to + 2 em relação ao período to – 1.

a) 7,00% b) 6,08% c) 7,16% d) 9,00% e) 6,11%

REFERÊNCIAS BIBLIOGRÁFICAS

- Crespo, Antonio Arnot: Estatística Fácil, 13ª Edição – São Paulo – Saraiva, 1995.
- Mandin, Daniel: Estatística Descomplicada, 4ª Edição – Brasília – Vestcon, 2001.
- Carvalho Filho, Sérgio de – 2ª Ed. – Estatística Básica: Teoria e 150 questões – Rio de Janeiro:Elsevier 2005