ESTATÍSTICA BÁSICA – 2ª Edição – 1998

Índice

ESTATÍSTICA BÁSICA

"É o estudo de fenômenos relativos a uma população, observando seu comportamento e identificando sua tendência para o futuro". É sobretudo um método de observação, seja através de pesquisa direta, seja através de comparação de fenômenos semelhantes, identificando suas características.

Portanto, podemos concluir que a Estatística não é simples coleta de dados, ela estuda a probabilidade de ocorrência do fenômeno ou estuda seu comportamento no tempo. A Estatística deixou de ser simples catalogação de dados numéricos sobre fenômenos coletivos, para ser o estudo de como chegar a conclusões sobre o universo a partir da observação de amostras.

De acordo com a forma de observações das populações a pesquisa estatística pode ser: **Descritiva** e **Indutiva**.

Descritiva - é aquela que parte de um conjunto de dados e obtém conclusões sobre os mesmos, não passando do conjunto de conhecimentos fornecidos por estes dados.

Indutiva - é aquela que ultrapassa os limites do conjunto de conhecimentos fornecidos pelos dados observados, isto é, a partir de uma amostra estuda as características da população amostrada.

FASES DO TRABALHO ESTATÍSTICO

- Planejamento, Coleta, Crítica, Apuração, Apresentação e Análise dos dados.

- Planejamento - é o detalhamento de todo o processo de pesquisa, de acordo com o objetivo do trabalho.

- Coleta - é o levantamento de dados sobre o fato a ser analisado. Pode ser Direta e Indireta.

Direta **-** quando os dados são obtidos na fonte originária. (Questionários, cartas, entrevistas, observações, etc.)

Indireta **-** quando os dados são copilados de outras fontes que não a originária. (Anuários, publicações, etc.)

Críticas - é utilizada para avaliar a qualidade dos dados obtidos, rejeitando respostas falsas e. realizando possíveis correções para homogenização dos dados.

- Apuração - Manual, Mecânica e Informatizada.

- Apresentação - Numérica e Gráfica.

- Análise - Estudo e conclusões sobre a população investigada.

APRESENTAÇÃO TABULAR - uma tabela estatística compõe-se de elementos essenciais e complementares

Elementos Essenciais

Título - é a indicação que precede a tabela, contendo a designação do fato observado, local e época em que foi registrado.

Cabeçalho - é a parte superior da tabela que especifica o conteúdo das colunas.

Coluna Indicadora - é a parte da tabela que especifica o conteúdo das linhas.

Corpo - é o conjunto de colunas e linhas que contém, respectivamente, as informações sobre o fato observado.

Casa - é o cruzamento da coluna com a linha no corpo da tabela. Não existindo informações as casas devem ser preenchidas com algum sinal convencional:

— quando não existe a informação pesquisada.

... quando não se dispõe dos dados. Ex: falta de informações sobre as precipitações pluviométricas em Aracaju, no mês de junho.

? quando há dúvida sobre a exatidão do valor. É colocado a esquerda da informação.

& quando o dado retifica informação anteriormente publicada. É colocada a esquerda da informação.

x quando o dado for omitido a fim de evitar individualização da informação.

0 quando o valor numérico é muito pequeno para ser expresso pela unidade de medida utilizada.

Elementos Complementares (situam-se no rodapé da tabela).

Fonte - é a indicação da entidade responsável pelo fornecimento dos dados ou por sua elaboração.

Nota - são informações de natureza geral, destinadas a conceituar ou esclarecer o conteúdo das tabelas ou indicar a metodologia adotada no levantamento ou na elaboração dos dados.

Chamadas - são informações de natureza específica sobre determinada parte da tabela. São identificadas por algarismos arábicos ou asteriscos, sempre a esquerda de cada informação e a direita da coluna indicadora, em ordem crescente de cima para baixo e da esquerda para a direita.

UNIDADES DE MEDIDA **-** devem ser usadas sempre no singular e em letras minúsculas. Grama (g) Quilo (kg) Tonelada (t) Metro (m) Segundo (s) Minuto (min), etc.

Data de Referência dos Dados

- Anos civis consecutivos: 1960-70; 1890-910

- Anos civis não consecutivos: 1950-1965

- Período de doze meses diferente do ano civil: 1960/65

Principais normas para construção de tabelas

- As tabelas são delimitadas na parte superior e inferior, exclusive a esquerda e a direita.

- É Facultativo o emprego de traços para separar as colunas no corpo da tabela.

- Quando a tabela ocupar mais de uma página por excessiva altura, só será delimitada na última página. Por excessiva largura, divide-se a tabela em quantas partes forem necessárias. Em ambos os casos usa-se devidamente as expressões continua e conclusão.

- Quando uma tabela ocupar páginas confrontantes, todas as linhas devem ser numeradas na 1ª e última coluna. O disposto acima não sendo possível, deve-se desmembrar a tabela em seções uma embaixo da outra, separadas por um traço horizontal duplo.

- Quando a tabela tiver poucas colunas e muitas linhas poderá ser disposta em duas ou mais partes, lado a lado, separadas por um traço vertical duplo.

SÉRIES ESTATÍSTICAS

É um conjunto de dados numéricos relativo a fatos observados, expressando as variações qualitativas ou quantitativas através de números por ordem de grandeza.

Elementos Fundamentais da Série Estatística

- Época (tempo) - período relativo a coleta de dados.
- Região (espaço) - local onde se passam os fatos observados
- Fenômeno (fato observado) - variável pesquisada

TIPOS DE SÉRIES ESTATÍSTICAS

Conforme varie cada um dos elementos acima expostos, podemos classificar as séries estatísticas em quatro tipos fundamentais:

Série Histórica, Evolutiva, Temporária ou Cronológica - quando a variável é o tempo, permanecendo fixos o local e o fato observado.

Série Geográfica, Territorial ou Espacial - quando a variável é o local, permanecendo fixos o tempo, e o fenômeno observado.

Série Específica ou Categórica -quando a variável é o fenômeno descrito, permanecendo fixo o tempo e o local.

Distribuição e Freqüência ou Série de Freqüência - são aquelas que permanecem fixos o tempo, o local e o fenômeno descrito, sendo que este é apresentado através de gradações que concentram os resultados das observações.

Estudo de Distribuições de Freqüências

Variável - Símbolo suscetível de assumir valores dentro de um certo conjunto.

- *Variável Discreta* - aquela que assume valores em pontos, isto é, está associada ao processo de contagem. Ex.: número de alunos, número de crianças, número de prédios, etc.

- *Variável Contínua* - é aquela que pode assumir qualquer valor entre dois pontos. Esta variável resulta normalmente em mensuração. Ex.: Estatura, Peso, Idade, etc.

Dados Brutos - são aqueles que ainda não foram numericamente organizados. Ex: Índice Pluviométrico registrado num período de 42 dias, no município "X".

146	152	159	160	141	147	163	178	157	174	172
159	151	166	148	157	151	162	173	141	165	152
142	158	148	162	155	176	150	143	158	168	
154	179	153	145	170	174	168	172	150	164	

Rol - É a organização dos dados brutos, segundo sua grandeza. Índice Pluviométrico: (n=42)

141	145	148	151	154	158	160	164	168	173	178
141	146	150	152	155	158	162	165	170	174	179
142	147	150	152	157	159	162	166	172	174	
143	148	151	153	157	159	163	168	172	176	

Amplitude Total (A) - é a diferença entre o maior e o menor valor do Rol ou da Distribuição de Freqüência. Ex.: A=179 - 141 = 38

Intervalo de Classe (h) - é a diferença absoluta entre o limite inferior e superior de classes sucessivas. Corresponde aos intervalos em que o ROL foi subdividido.

Para populações ou Amostras com 200 ou mais unidades utiliza-se para cálculo do intervalo a fórmula empírica de Sturges:

$$h = \frac{A}{1 + 3{,}322 \log n} = \frac{38}{1 + 3{,}322 \log 42} \cong 5$$

Para pequenas amostras podemos utilizar a referida fórmula com prioridades, quando for possível, para intervalos múltiplos de 10, múltiplos de 5 e números pares.

Classes de Freqüências - são os grupamentos em que a série foi subdividida, de acordo com a dimensão do Intervalo de Classe. São constituídas pelos valores da variável x, que se enquadram entre seus extremos: Li — Ls; Li |— Ls; Li—| Ls; Li |—| Ls.

ÍNDICE PLUVIOMÉTRICO	FREQÜÊNCIA (f_i)
140 |—145	4
145 |—150	5
150 |—155	8
155 |—160	7
160 |—165	5
165 |—170	4
170 |—175	6
175 |—180	3
S	**42**

Ponto Médio (x_i) - são os valores eqüidistantes dos extremos de cada classe de freqüência. $x_i = L_i + \frac{h}{2}$

Freqüências de Classes

Freqüência Simples Absoluta (f_i) - é o número de vezes que se repete um determinado valor de uma variável no caso de distribuição de freqüências por valores, ou o número de valores de uma variável contida em cada classe de freqüência caso se trate de distribuição de freqüência por classes de valores.

Características de f_i

a) S $f_i = f_1 + f_2 + ... + f_n = n$
onde n é a freqüência total, correspondente ao número de observações.

b) $0 = f_i = n$
$f_i = 0 \Rightarrow$ não existe valores na classe i
$f_i = n \Rightarrow$ existe apenas uma classe de freqüência ou uma mesma variável, no caso de distribuição de freqüências por valores.

Freqüência Simples Relativa (fr_i) - é o número relativo de unidades em cada classe de freqüência. É obtida dividindo-se cada freqüência simples absoluta pela freqüência total.

Característica de fr_i

a) Se fr_i=0, significa que f_i=0 se fr_i=1, significa que f_i=1

$\sum fr_i = fr_1 + fr_2 + ... + fr_n$ $$\sum fr_i = \frac{f_1}{n} + \frac{f_2}{n} + ... + \frac{f_n}{n} \Rightarrow \sum fr_i = \frac{1}{n}.n = 1$$

Freqüência Acumulada de uma Classe - é a soma da freqüência simples (absoluta ou relativa) desta classe com as respectivas freqüências simples das classes anteriores. As freqüências acumuladas também são chamadas: Abaixo de (crescente) e Acima de (decrescente).

Principais Características

$F_1 = f_1$ e $F_n = \sum f_i = n$ $Fr_1 = fr_1$ e $Fr_n = f_n$

Freqüência Percentual - é o produto da freqüência relativa (simples ou acumulada) por 100
$f_i\% = fr_i \times 100 \Rightarrow \sum f_i\% = 100\%$ e $F_i\% = Fr_i \times 100$

APRESENTAÇÃO GRÁFICA

A representação gráfica de uma série estatística tem por finalidade a ilustração e a visão mais rápida e geral do fato observado com suas características particulares.

Principais Diagramas

- Curvas
- Colunas ou Barras
- Setores
- Polar
- Histograma
- Polígono de Freqüência
- Semilogarítmos
- Ogivas

Diagramas em Curvas: este tipo de gráfico presta-se para a representação de séries cujos dados se apresentam em função do tempo (séries cronológicas). O princípio geral do gráfico em curva é caracterizado pelo par (x, y) de coordenadas que podem ser representadas num sistema cartesiano. Determinados graficamente, todos os pontos da série, basta uni-los por segmentos de reta.

CONSUMO RESIDENCIAL DE ENERGIA ELÉTRICA EM ARACAJU - 1990-95

ANOS	CONSUMO (Mwh)
1990	170.489
1991	175.048
1992	223.903
1993	268.323
1994	217.018
1995	204.571

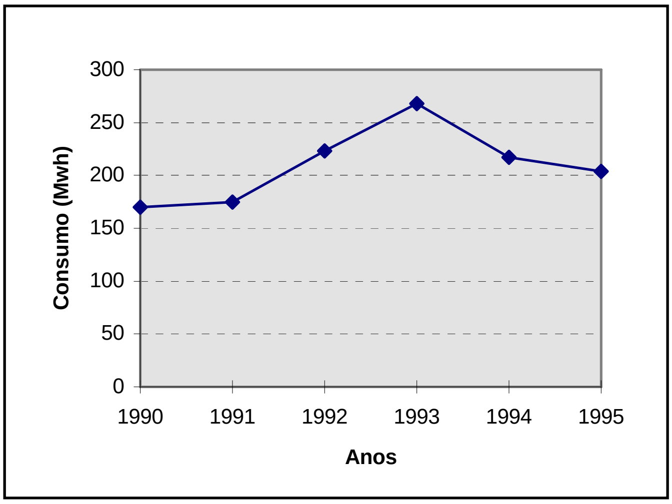


Diagrama Em Colunas - esta representação é feita por meio de retângulos com áreas proporcionais aos dados observados. A distância entre as colunas corresponde a 2/3 ou metade da base do retângulo. Sempre que possível as colunas devem ser construídas em ordem crescente ou decrescente. O mesmo critério é utilizado para o diagrama em barras. Pode ser utilizados em séries históricas, geográficas e específicas.

ÓBITOS MENORES DE UM ANO DA REGIÃO SUDESTE DO BRASIL - 1992

SUDESTE	Nº DE ÓBITOS
Minas Gerais	9.729
Espirito Santos	1.602
Rio de Janeiro	7.444
São Paulo	17.141
TOTAL	**35.916**

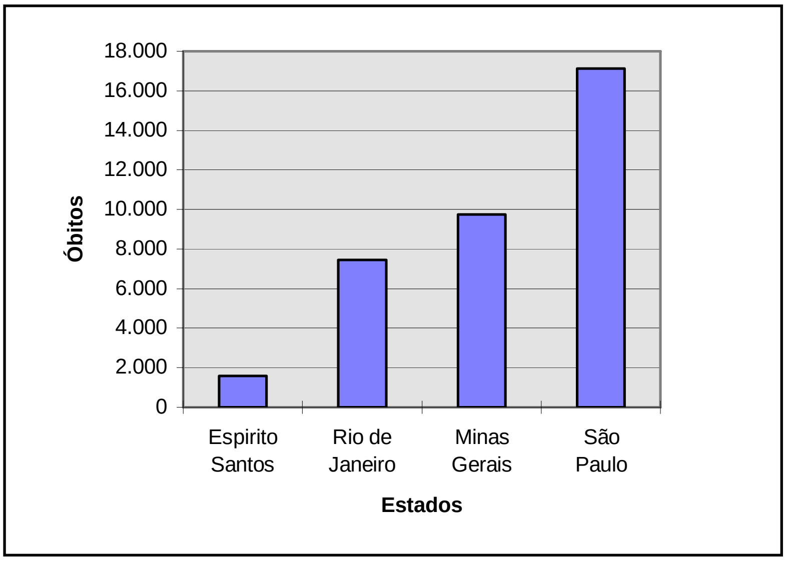
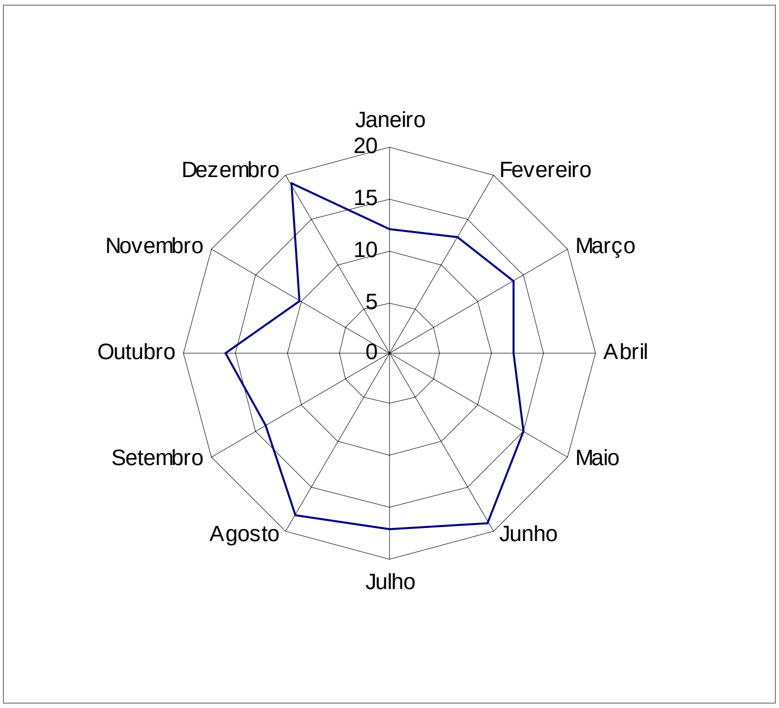


Diagrama Polar - é construído num círculo de raio arbitrário, dividido em partes iguais de acordo com o número de itens da série a ser representada. Marca-se, em cada raio vetor, o valor correspondente e, une-se os pontos por segmentos de reta. Tem grande aplicação na análise de séries mensais.

MOVIMENTO MENSAL DE COMPRAS DA EMPRESA DELTA
1995

MESES	VALORES (R$1.000,00)
Janeiro	12
Fevereiro	13
Março	14
Abril	12
Maio	15
Junho	19
Julho	17
Agosto	18
Setembro	14
Outubro	16
Novembro	10
Dezembro	19

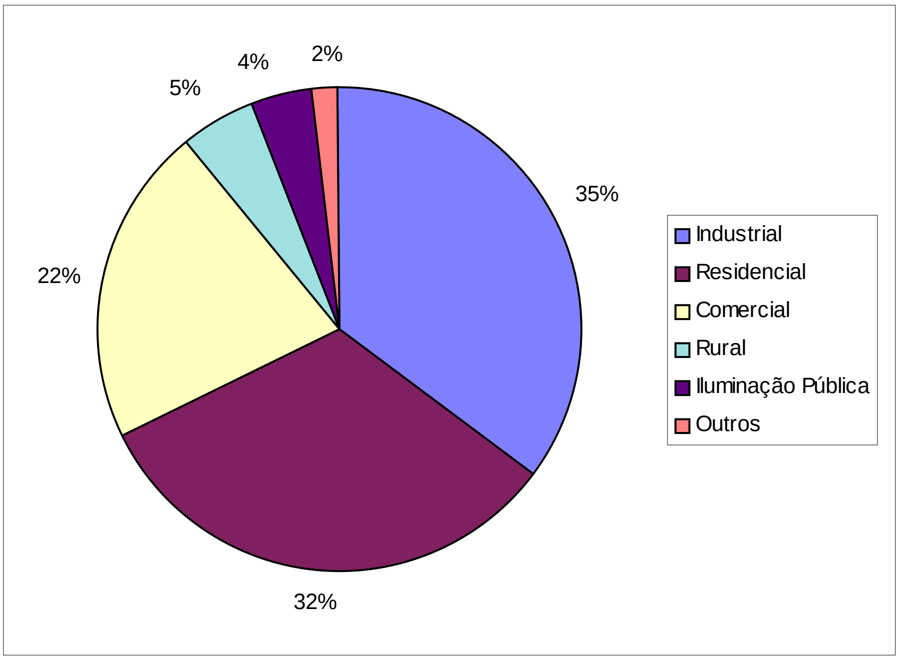


Cada ponto (P) do gráfico fica determinado pelo raio vetor e pelo ângulo polar (x, α). No gráfico Polar a coordenada angular (α) é constante, variando a coordenada linear x de acordo com o valor a ser representado.

Diagrama Em Setores - é construído num círculo de raio qualquer, com ângulos centrais (setoriais) proporcionais as parcelas dos dados observados. Pode ser aplicado em séries geográficas, específicas e históricas.

CONSUMO DE ENERGIA ELÉTRICA EM ARACAJU, POR CLASSE DE CONSUMIDOR - 1995

CLASSES	CONSUMO (Mwh)
Residencial	204.571
Comercial	139.626
Industrial	225.310
Iluminação Pública	25.059
Rural	32.055
Outros	11.360
TOTAL	**637.981**



No diagrama devem ser registrado os percentuais dos setores, que devem ser construídos em ordem decrescente no sentido horário, a partir do raio fixado acima do centro do círculo no sentido vertical.

Histograma - é um conjunto de retângulos adjacentes, com base no eixo das abcissas e áreas proporcionais às freqüências de classe.

Polígono de Freqüência - é constituído com base nos pontos médios das classes de freqüências da distribuição, ligados entre si por segmentos de retas.

Poligonal característica ou curva característica - é o contorno externo do histograma.

Curva de freqüência - é o limite do histograma e do polígono de freqüência, quando “n” cresce (grandes amostras), enquanto o intervalo de classe (h) tendem a zero.

Estes diagramas são utilizados em distribuição de freqüência.

Idade dos Alunos do Primeiro Grau da Escola Fonte do Saber - 1995

Idade	Nº de Alunos
$4 \vdash 6$	10
$6 \vdash 8$	15
$8 \vdash 10$	27
$10 \vdash 12$	20
$12 \vdash 14$	5
$\sum$	77

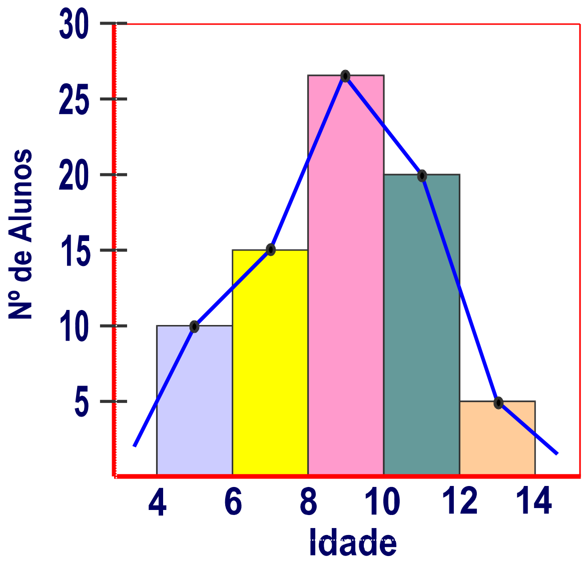


Diagrama Semilogaritmo - são usados para estudar as variações proporcionais (relativas), em vez das flutuações absolutas dos dados. É aplicado quando os valores da série estão em progressão geométrica, bem como em estudos comparativos de duas séries quando uma delas tem magnitude muito maior que a outra.

Ex.: Representar num gráfico Semilogaritmo, com "camadas" de 7 cm os dados da série abaixo:

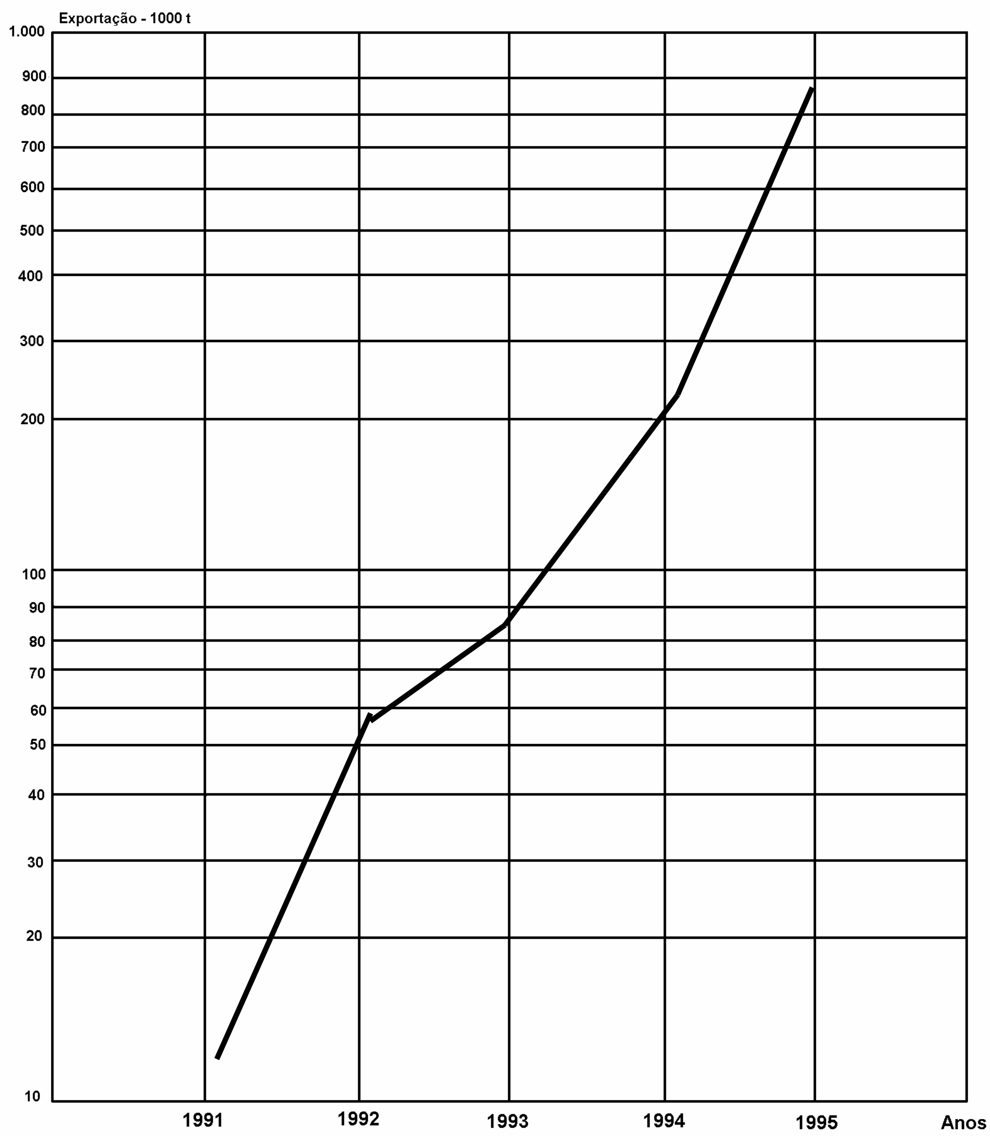
EXPORTAÇÃO DE TECIDO, TIPO A DO ESTADO DE SERGIPE
1991 - 95

ANO	EXPORTAÇÃO (t)	MANTISSA DOS NÚMEROS	MANTISSA X 7 cm (*)
1991	11.433	.0581	0,4
1992	56.694	.7535	5,3
1993	88.854	.9487	6,6
1994	225.936	.3539	2,5
1995	896.432	.9529	6,7

(*) - Altura dos pontos

Construção do Diagrama - altura das camadas semilogarítmicas (logα x 7cm)

log α	log α x 7 cm	log α	log α x 7 cm
log 1 = 0,00	0,0 cm	log 6 = 0,78	5,5 cm
log 2 = 0,30	2,1 cm	log 7 = 0,85	6,0 cm
log 3 = 0,48	3,4 cm	log 8 = 0,90	6,3 cm
log 4 = 0,60	4,2 cm	log 9 = 0,95	6,7 cm
log 5 = 0,70	4,9 cm	log 10 = 1,00	7,0 cm



MEDIDAS DE POSIÇÃO

Média Aritmética ($\overline{x}$)

Média Aritmética de uma série de valores $x_i = x_1, x_2, \ldots, x_n$, é a razão entre a soma de todos os valores e o número de termos da série

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} \quad e \quad \overline{x} = \frac{\sum x_i}{n}$$

Propriedades da Média Aritmética

1- A soma algébrica da diferença entre cada valor observado e a média aritmética é nula: $\sum_{i=1}^{n}\left(x_i - \overline{x} = 0\right)$

$$\sum_{i=1}^{n}\left(x_i - \overline{x}\right) \;=\; \sum_{i=1}^{n} x_i \;-\; \sum_{i=1}^{n} \overline{x} \;=\; \mathbf{n}\overline{x} \;\mathbf{-}\; \mathbf{n}\overline{x} \;=\; \mathbf{0}$$

2 - A média de uma constante é a própria constante.

x_i = a, a, a, a, ..., a $\qquad \overline{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{na}{n} = a$

3 - A soma dos quadrados dos desvios calculados em relação à média aritmética é um mínimo.

$$\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 < \sum_{i=1}^{n}\left(x_i - a\right)^2 \qquad \textbf{para} \qquad \mathbf{a} \neq \overline{x}$$

Tomemos: $\mathbf{S = \sum(x_i - \overline{x})^2 = S\, x_i^2 - 2\overline{x}\; S\, x_i + n\, \overline{x}^2 = S\, x_i^2 - n\overline{x}^2}$

$\mathbf{R = S\,(x_1 - a)^2 = S\, x_1^2 - 2an\overline{x} + na^2}$

Subtraindo R de S temos:

$\mathbf{R - S = n\,(a^2 - 2a\overline{x} + \overline{x}^2) = n(a - \overline{x})^2}$

Como n e $(a - x)^2$ são positivos implica que R - S > 0, sendo R > S

4 - A média aritmética de uma série de valores é igual a um número arbitrário somado algebricamente à média dos desvios tomados em relação a esse número.

$$\bar{x} = A_0 + \frac{\sum d_i}{n} \qquad di = x_i - A_0$$

$$\bar{x} = A_0 + \frac{S f_i a_i}{S f_i} \cdot h \qquad a_i = \frac{x_i - A_0}{h}$$

5 - A média aritmética de duas ou mais séries de valores é um parâmetro ponderado entre as médias de cada série e respectivos números de termos.

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + \ldots + nk\bar{x}k}{n_1 + n_2 + nk} \Rightarrow \bar{x} = \frac{\sum ni\bar{x}i}{\sum ni}$$

Se as séries tiverem os mesmos números de termos isto é, se $n_1 = n_2 = \ldots = n_k$

$$\bar{x} = \frac{\bar{x}_1 + \bar{x}_2 + \ldots + \bar{x}k}{k}$$

6 - Somando ou subtraindo uma constante a todos os valores da série, a média ficará aumentada ou diminuída, respectivamente, dessa constante (idem para multiplicação e divisão, exceto quando a constante for igual a zero).

Soma $\Rightarrow$ $\mathbf{x_i = (x_1 + a);\ (x_2 + a);\ \ldots\ (x_n + a)}$

Média $= \dfrac{\mathbf{(x_1 + a) + (x_2 + a) + \ldots + (x_n + a)}}{\mathbf{n}} = \bar{x} + \mathbf{a}$

Multiplicação $\Rightarrow$ $\mathbf{x_i = ax_1, ax_2, \ldots, ax_n}$

Média $\Rightarrow \dfrac{\mathbf{ax_1 + ax_2 + \ldots + ax_n}}{\mathbf{n}} = \mathbf{a}\bar{x}$

Média Aritmética Ponderada

Em algumas séries determinados valores são mais significativos que outros, merecendo por isto tratamento especial, de acordo o grau de importância de cada valor da variável. Neste caso a média é a relação entre a soma dos produtos de cada valor pelos respectivos pesos ($p_i = p_1, p_2, ..., p_n$) ou freqüências de valores ($f_i = f_1, f_2, ..., f_n$).

$$\bar{x} = \frac{\mathbf{x_1 p_1 + x_2 p_2 + ... + x_n p_n}}{\mathbf{p_1 + p_2 + ... p_n}} = \frac{\sum_{i=1}^{n} x_i p_i}{\sum_{i=1}^{n} p_i} \quad \textbf{ou} \quad \bar{x} = \frac{\sum \mathbf{x_i} f_\mathbf{i}}{\sum f_\mathbf{i}}$$

Controle de Charlier para a média aritmética - destina-se a verificar a exatidão dos cálculos efetuados.

$$f(\alpha_i + 1) = f\alpha_i + f \quad \Leftrightarrow \quad \sum f(\alpha_i + 1) = \sum f\alpha_i + n$$

Erro Padrão ou Erro da Média

Diferentes amostras retiradas da mesma população podem apresentar média diferentes.

A variação existente entre este conjunto de médias é estimado através do erro padrão, que corresponde ao desvio padrão das médias.

Representamos o erro padrão por. $s_{\bar{x}} = \frac{s}{\sqrt{n}}$ *ou* $s_{\bar{x}} = \sqrt{\frac{s^2}{n}}$

Erro Relativo máximo cometido no cálculo da média aritmética para dados agrupados em distribuição de freqüência.

$$E = \frac{\mathbf{h}}{2\bar{x} - \mathbf{h}}$$

Principais características da Média Aritmética

- É descritiva de todos os dados de uma série e de fácil compreensão.
- Depende de cada valor da série e qualquer alteração de um deles altera o seu valor.

- É influenciada por valores extremos, podendo, em alguns casos, não representar a série.
- Representa uma série cujos valores estão ou se aproximam de uma progressão aritmética.
- É das medidas de tendência central a de maior emprego.
- Tem grande aplicação nas distribuições simétricas.
- Não pode ser calculada para distribuições com limites indeterminados.
- Não necessariamente tem existência real, isto é, nem sempre é um elemento que faça parte do conjunto, para bem representá-lo, embora pertença, obrigatoriamente, ao intervalo entre a maior e a menor ocorrência.

Média Geométrica ($\bar{x}_g$)

Média Geométrica - é a raiz $n^{ésima}$ do produto de ***n*** termos positivos $x_i = x_1, x_2, ..., x_n$.

$$\bar{x}_g = \sqrt[n]{x_1 . x_2 \ldots x_n} \quad \Rightarrow \quad \bar{x}_g = \sqrt[n]{\pi x_i}$$

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

Média Geométrica Ponderada - dados n termos $x_i = x_1, x_2, ..., x_n$ com pesos ($p_i = p_1, p_2, ... p_n$) ou ($f_i = f_1, f_2, ... f_n$) a média geométrica é calculada pelas seguintes equações:

a) $\bar{x}_g = \sqrt[\Sigma pi]{x_1^{p1} . x_2^{p2} \ldots . x_n^{pn}} \quad \Rightarrow \quad \bar{x}_g = \sqrt[\Sigma pi]{\pi \, x_i^{pi}}$

b) $\bar{x}_g = \sqrt[\Sigma fi]{\pi \, x_i^{fi}}$ **ou** $\log \bar{x}_g = \dfrac{1}{\sum_{i=1}^{n} f_i} . \sum_{i=1}^{n} f_i \log x_i$

Principais característica da Média Geométrica

- A média geométrica de duas ou mais séries de valores pode ser obtida através das respectivas médias de cada série, isto é:

$$\bar{x}_g = \sqrt[n_1 + n_2 + \dots + n_k]{\bar{x}g_1^{n1}\,\bar{x}g_2^{n2}\ldots\bar{x}g_K^{nK}}$$

- Não tem significado quando houver valores nulos na série ou quando existir número ímpar de valores negativos;
- É menos influenciada pelos valores extremos do que a média aritmética;
- Adquire maior representatividade, nas séries em que os valores se escalonam em progressão geométrica;
- Tem boa aplicação no cálculo de:
  - Números Índices
  - Taxas de crescimento relativamente constante (população, capitais sujeitos a juros compostos etc.)
  - Valor médio em um conjunto de percentagens de variações mensais.

Média Harmônica ($\bar{x}_h$)

Dados n termos $x_i = x_1, x_2, \ldots x_n$, a média harmônica desses valores é o inverso da média aritmética dos inversos dos valores.

$$\bar{x}_h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

Média Harmônica Ponderada - utilizada quando os valores da variável estão ponderados ou agrupados em Distribuição de Freqüências.

$$\bar{x}_{\mathbf{h}} = \frac{\sum_{i=1}^{n} p_i}{\sum_{i=1}^{n} \frac{p_i}{x_i}} \qquad \textbf{ou} \qquad \bar{x}_{\mathbf{h}} = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{x_i}}$$

Principais Características da Média Harmônica

- Representa uma série de razões ou de grandezas inversamente proporcionais;
- Apesar de não ser muito empregada é recomendada para séries estatísticas recíprocas.
  - cálculo de velocidade média entre dois pontos
  - custo média de artigos comprados a uma quantia fixa
  - confronto entre média de velocidade e tempo (grandezas inversamente proporcionais).

Comparação entre as médias: $\bar{x}_{\text{h}} = \bar{x}_{\text{g}} = \bar{x}$

a) Demonstrar que $\bar{x}_{\text{g}} = \bar{x}$

$$\frac{a+b}{2} = \sqrt{a.b}$$

$$(a+b)^2 = 4ab$$

$$(a-b)^2 = 0$$

b) Demonstrar que $\bar{x}_{\text{h}} = \bar{x}_{\text{g}}$

$$\sqrt{a.b} = \frac{2}{\frac{1}{a} + \frac{1}{b}}$$

$$ab\,(a+b)^2 = 4\,a^2b^2 \Rightarrow (a-b)^2 = 0$$

Moda (M_o)

Dada uma série de n valores $x_i = x_1, x_2, \ldots, x_n$, denominamos Moda o valor que ocorre com maior freqüência.

Em relação à moda as séries estatísticas podem ser: **Amodal -** quando não possui moda. **Modal -** quando possui uma moda. **Bimodal -** quando possui duas modas. **Plurimodal -** quando possui várias modas.

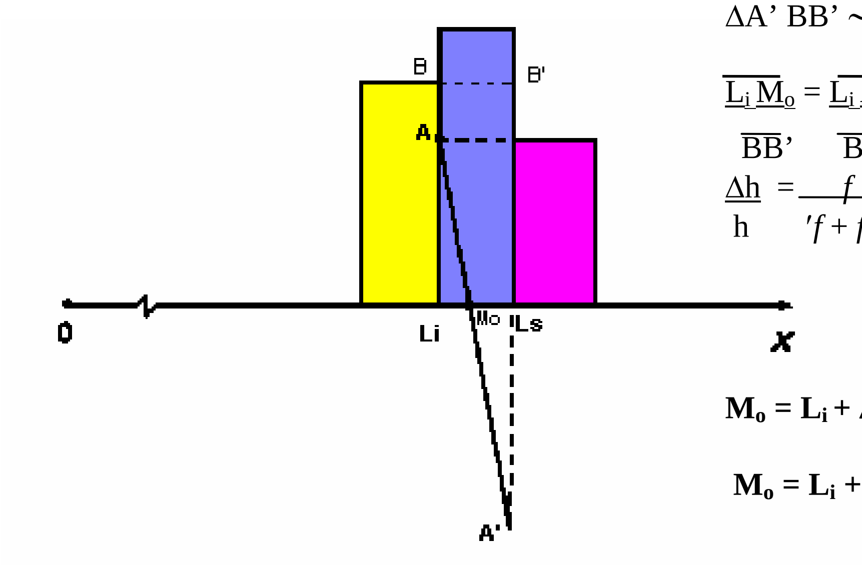
Moda para dados agrupados em Distribuição de Freqüências.

Classe Modal - é a classe que possui a maior freqüência.

Moda Bruta - é o ponto médio da classe modal.

Métodos utilizados para o cálculo da Moda de uma *Distribuição de Freqüência: King, Czuber* e *Pearson*.

PROCESSO DE KING



$\overline{L_iB} = \overline{L_sB'}$ e $\overline{L_iA'} = \overline{L_sA}$

$\Delta A'BB' \sim \Delta A'L_iM_o$

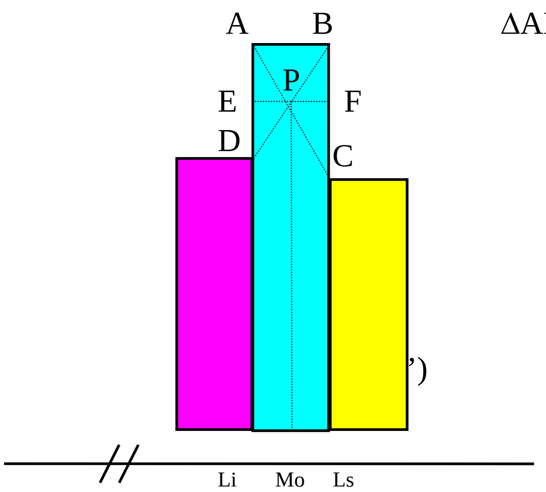
$$\frac{\overline{L_iM_o}}{\overline{BB'}} = \frac{\overline{L_iA'}}{\overline{BA'}}$$

$$\frac{\Delta h}{h} = \frac{f}{'f + f'}$$

$$\mathbf{M_o = L_i + \Delta h}$$

$$\mathbf{M_o = L_i + \frac{f'.h}{'f + f'}}$$

PROCESSO DE CZUBER



$\overline{AD} = \Delta_1 = f - 'f$ e $\overline{BC} = \Delta_2 = f - f'$

$\Delta ADP \sim \Delta BPC$. $\frac{\overline{EP}}{\overline{PF}} = \frac{\overline{AD}}{\overline{BC}}$

$$\frac{\Delta h}{h - \Delta h} = \frac{\Delta_1}{\Delta_2}$$

$$Mo = Li + \Delta h$$

$$Mo = Li + \frac{(f - 'f).h}{2f - ('f + f')}$$

Relação Empírica de Pearson $\Rightarrow$ $M_o = 3\ M_e - 2\bar{x}$ Aplicável em distribuições de fraca assimetria:

Principais características da Moda

- É medida de posição
- Tem boa aplicação quando se deseja assinalar o valor mais comum ou típico.
- É menos estável do que a média aritmética e apresenta grande instabilidade na amostra.
- Não se deixa influenciar pelas flutuações extremas, visto que não depende de todos os valores da série.
- A existência de limites indefinidos não impede que seja calculada a moda da distribuição.
- Apresenta a desvantagem de não se adaptar ao tratamento algébrico.
- Não perde seu significado mesmo nas distribuições assimétricas.
- Sempre tem existência real, ou seja, sempre é representada por um elemento do conjunto de dados, exceto quando os dados estão agrupados em distribuição de freqüência.

Mediana (M_e)

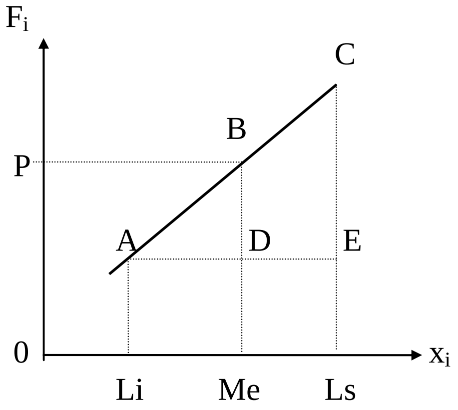
Mediana de uma série de "n" termos $x_i = x_1, x_2. .... x_n$, ordenados, é o elemento que separa a série em dois subconjuntos, de modo que seja precedido e seguido pelo mesmo número de ocorrências.

Posição da Mediana para dados não tabulados: $\mathbf{P = \dfrac{n+1}{2}}$.

Para série com número ímpar de termos a mediana é o termo central da série. Caso n seja par a mediana é a média aritmética entre os termos centrais da série.

Para dados agrupados em distribuição de freqüências, a Mediana encontra-se na classe mediana que é identificada pelas freqüências acumuladas. $\left(P = \dfrac{n}{2}\right)$

Cálculo da mediana para dados agrupados em distribuição de freqüência por classe.



$\overline{CE} = f_i$; $\overline{BD} = P - {'F_i}$ e $\overline{AE} = h$

$\overline{AD} = M_e - L_i = \Delta h$ $\Delta ACE \sim \Delta ABD$

$$\frac{\overline{CE}}{\overline{AE}} = \frac{\overline{BD}}{\overline{AD}} \Rightarrow \frac{f}{h} = \frac{P - {'F_i}}{M_e - L_i} \Rightarrow M_e = L_i + \frac{(P - {'F_i}) \cdot h}{f}$$

Principais Características da Mediana

- Tem boa aplicação em séries de valores que há resultados extremos que afetariam de maneira acentuada a média (Renda Nacional, etc.). Em distribuições que existem gradações: (Tenente, capitão, etc.)
- A soma dos valores absolutos de todos os desvios referente a mediana é mínima em relação aos desvios absolutos referentes a outro valor;
- Não se deixa influenciar pela magnitude dos termos extremos;
- É a abcissa de ordenada que divide a área da curva em duas partes iguais;
- É uma medida separatriz;
- Apresenta grande instabilidade na amostra. Se extrairmos amostras distintas de uma mesma população, a diferença que existe, em geral, entre as médias, é menor que a que se pode observar entre suas respectivas medianas;
- É influenciada pela posição ou localização dos valores e não pela magnitude deles.

Separatrizes

Quartil (Q), Decil (D) e Centil (C) são separatrizes que dividem a série em quatro (4), dez (10) e cem (100) partes iguais respectivamente.

A posição das separatrizes para dados não agrupados segue o mesmo critério da Mediana.

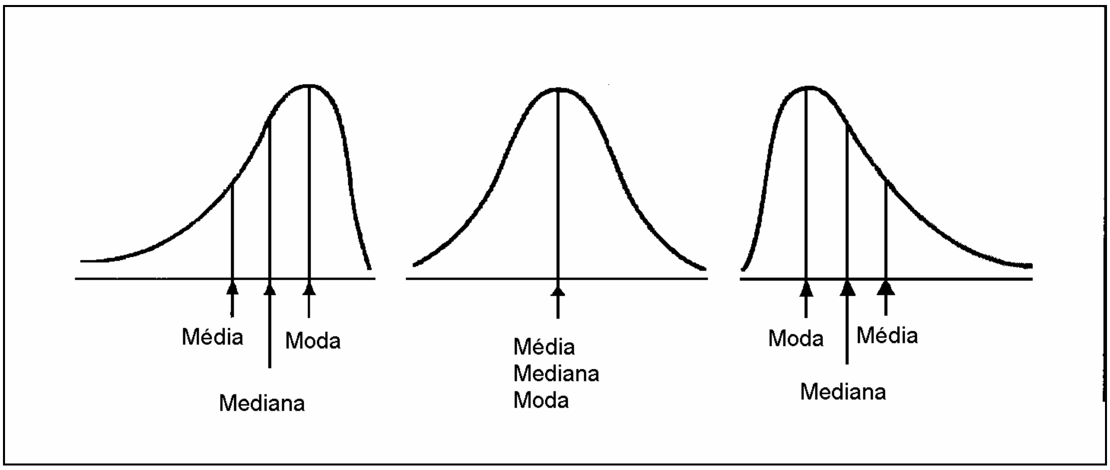
Para dados agrupados em Distribuição de Freqüências a posição das separatrizes é dada pela relação: $\mathbf{P = \frac{k\,.\,n}{i}}$, com classes identificadas pelas freqüências acumuladas.

Para cálculo do valor de qualquer Separatriz utiliza-se a equação da mediana, visto que ela própria é uma separatriz. $Me = Q_2 = D_5 = C_{50}$

Relação entre Média Aritmética, Moda e Mediana

Em geral a moda é menos empregada, sendo adequada para caracterizar situações onde estejam em causa os valores mais freqüentes. Por exemplo em estudos de mercado, o principal interesse pode ser pesquisar os produtos que mais se vendem.

Corretamente a escolha é feita entre a média e a mediana. Em alguns casos a mediana é mais resistente do que a média, principalmente nas séries assimétricas, por não ser influenciada pelos valores extremos. Já a média tem vantagens quando a curva de freqüência é mais ou menos simétrica, transformando-a num bom estimador de parâmetros populacionais, através de dados amostrais.



Quando a distribuição é simétrica, sua média, moda e mediana coincidem. Caso contrário a média e a mediana se deslocam mais em direção dos valores extremos, a média mais que a mediana. Portanto em qualquer distribuição:

- a moda é sempre a abcissa de ordenada máxima;
- a mediana se situa sempre entre a moda e a média aritmética;
- a média aritmética se situa sempre ao lado da cauda de distribuição.

1. EXERCÍCIOS RESOLVIDOS

1.01 - Dez medidas do diâmetro de um cilindro foram anotadas por um cientista como: 3,88; 4,09; 3,92; 3,97; 4,02; 3,95; 4,03; 3,92; 3,98 e 4,06 polegadas. Determinar a média aritmética.

$$\bar{x} = \frac{\sum xi}{n} = \frac{3{,}88 + 4{,}09 + 3{,}92 + 3{,}97 + 4{,}02 + 3{,}95 + 4{,}03 + 3{,}98 + 3{,}92 + 4{,}06}{10}$$

$$\bar{x} = \frac{39{,}82}{10} = 3{,}89 \text{ polegadas}$$

1.02 - Entre 100 números, vinte são 4, quarenta são 5, trinta são 6 e os restantes são 7. Determinar a média aritmética dos números.

Solução:

$$\bar{x} = \frac{\sum x_i \cdot f_i}{n} = \frac{80 + 200 + 180 + 70}{100} \Rightarrow \bar{x} = 5{,}30$$

1.03 - Calcular as médias aritmética, geométrica e harmônica entre os números 8, 17 e 22

Solução:

$$\bar{x} = \frac{\sum x_i}{n} \qquad \bar{x} = \frac{8 + 17 + 22}{3} = 15{,}7$$

x_i	$\log x_i$
8	0,90
17	1,23
22	1,34
	3,47

$$\log \bar{x}_g = \frac{\sum \log x_i}{n}$$

$$\log \bar{x}_g = \frac{3{,}47}{3} = 1{,}16 \Rightarrow \bar{x}_g = 14{,}45$$

$$\bar{x}_h = \frac{n}{\sum \frac{1}{}} \qquad \bar{x}_h = \frac{3}{1/8 + 1/17 + 1/22} = 13{,}1$$

x_i.

1.04 - Um homem viaja de A para B à velocidade média de 30 Km/h e volta de B para A, pelo mesmo caminho, à velocidade média de 60 km/h. Determinar a velocidade média para a viagem completa.

Solução:

Suponhamos que a distância entre A e B é de 60 km, apesar de que podemos considerar qualquer distância.

Tempo para se deslocar de A para B: $\dfrac{60\text{ km}}{30\text{ km/h}} = 2$ horas

Tempo de B para A = $\dfrac{60\text{ km}}{60\text{ km/h}} = 1$ hora

Velocidade média para a viagem completa:

$\dfrac{\text{distância total}}{\text{tempo total}} = \dfrac{120\text{ km}}{3\text{ h}} = 40$ km/h (média harmônica entre 30 e 60)

$$\bar{x}_h = \frac{n}{\sum \frac{1}{x_i}} = \frac{2}{\frac{1}{30} + \frac{1}{60}} = 40 \text{ km}$$

1.05 - Determinar a média geométrica entre: 2, 5, 7, 10 e 12

$$\bar{x}_g = \sqrt[n]{\prod_{i=1}^{n} x_i} \Rightarrow = \sqrt[5]{2x5x7x10x12} = \sqrt[5]{8.400}$$

$$\log \bar{x}_g = \frac{\log 8.400}{5} = \frac{3,92428}{5} = 0,78485 \Rightarrow \bar{x}_g = 6,09$$

1.06 - Determinar a mediana das distribuições abaixo:

x_i = 9; 11; 14; 15; 17; 19; 23; 28 e y_i = 2; 5; 7; 10; 13; 17; 20;

$$M_e = \frac{15 + 17}{2} = 16 \qquad M_e = 10$$

1.07 - Determinar a moda das distribuições abaixo:
x_i = 2; 3; 5; 5; 6; 6; 6; 6; 7 e y_i = 2; 2; 3; 4; 4; 4; 5; 7

$$M_o = 6 \qquad M_o = 4$$

1.08 - Achar a mediana dos seguintes valores:

46 49 54 47 58 55 65 62 46
59 64 52 41 61 67 48 41 72

$$Me = \frac{54 + 55}{2} = 54{,}5$$

$$P = \frac{n + 1}{2} = \frac{18 + 1}{2} = 10{,}5 \text{ (entre a posição } 10^a \text{ e } 11^a\text{)}$$

1.09 - Calcular a Média Aritmética, Mediana e Moda da idade relativa a um grupo de 350 pessoas:

Idade (anos)	**Nº de pessoas**	x_i	α_i	$f_i\alpha_i$	F_i	$f_i\alpha_i^2$
10 |—15	2	12,5	-2	-4	2	8
15 |—20	50	17,5	-1	-50	52	50
20 |—25	120	22,5	0	0	172	0
25 |—30	60	27,5	1	60	232	60
30 |—35	110	32,5	2	220	342	440
35 |—40	8	37,5	3	24	350	72
$\sum$	350	-	-	250	-	630

$$\bar{x} = A_0 + \frac{\sum f_i \alpha_i}{\sum f_i} \cdot h = 22{,}5 + \frac{250}{350} \cdot 5 \Rightarrow \bar{x} = 26{,}07$$

$$M_e = L_i + \frac{P - {'F_i}}{f_i} \cdot h = 25 + \frac{(175\text{-}172)}{60} \cdot 5 \Rightarrow \qquad M_e = 25{,}25$$

$$M_o = L_i + \frac{f - {'f}}{2f\text{-}({'f} + f')} \cdot h = 20 + \frac{(120 - 50)}{240 - 110} \cdot 5 \Rightarrow \qquad M_o = 22{,}69$$

1.10 - Calcular: Q_1, C_{10} e C_{90} da questão 1.09

$$Q_1 = L_i + \frac{(P - 'F)}{f} \cdot h = 20 + \frac{(87{,}5 - 52)}{120} \cdot 5 \Rightarrow \qquad Q_1 = 21{,}48$$

$$C_{10} = L_i + \frac{(P - 'F)}{f} \cdot h = 15 + \frac{(35 - 2)}{50} \cdot 5 \Rightarrow \qquad C_{10} = 18{,}30$$

$$C_{90} = L_i + \frac{(P - 'F)}{f} \cdot h = 30 + \frac{(315 - 232)}{110} \cdot 5 \Rightarrow \qquad C_{90} = 33{,}7$$

MEDIDAS DE DISPERSÃO

Dispersão - É o afastamento de todos os valores de uma série em relação a média aritmética ou mediana. De acordo com a grandeza dos afastamentos as séries estatísticas podem ser: homogêneas (fraca dispersão $\Rightarrow C_v < 30\%$) e heterogêneas (forte dispersão $\Rightarrow C_v \geq 30\%$)

x_i	-	5, 5, 5, 5, 5	$\Rightarrow$	dispersão nula
y_i	-	3, 4, 5, 6, 7	$\Rightarrow$	fraca dispersão
z_i	-	1, 2, 5, 8, 9	$\Rightarrow$	forte dispersão

Analisando as distribuições acima verificamos que todas possuem a mesma média aritmética, contudo revelam positivas diferenças entre elas, sendo a principal, a que se refere ao grau de concentração de valores, em torno de um valor médio, que é mensurado pela dispersão absoluta ou relativa.

Dispersão Absoluta: dimensiona o afastamento médio entre os termos de uma série e seu valor médio. Seu resultado é expresso na mesma unidade de medida dos dados pesquisados. **Amplitude Total, Desvio Quartil, Desvio Médio, Variância e Desvio Padrão.**

Dispersão Relativa: mede a relação entre a dispersão absoluta e o valor médio da série, com resultado expresso em porcentagens. Principais: Desvio Quartil Reduzido e Coeficiente de Variação de Pearson.
Amplitude Total - é a diferença entre os valores extremos da série:
A = Ls - Li

A amplitude total é pouco utilizada na análise estatística, porque sendo calculada apenas com os valores extremos da série, é muito instável, pela influência de valores excepcionalmente altos ou baixos.

Desvio Quartil - é a semidiferença entre o terceiro e primeiro quartil, cuja amplitude em torno da mediana abrange 50% dos valores centrais.

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

O Desvio Quartil é mais estável do que a amplitude, visto que utilizando 50% dos valores centrais, despreza apenas 25% dos valores mais baixos, eliminando, portanto as flutuações ao acaso.

Aplicação do Desvio Quartil

- quando a mediana for a medida da tendência central;
- quando houver valores esparsos ao extremos capazes de influenciar desproporcionalmente outra medida de dispersão;
- quando a concentração em torno da mediana for de interesse primordial.

Desvio Médio - é a média aritmética dos desvios absolutos tomados em relação à sua média aritmética.

$$Dm = \frac{\sum_{i=n}^{n} |xi - \overline{x}|}{n}$$

Para Dados Agrupados: $x_i = x_1, x_2........x_n$ e $f_1 = f_1, f_2.......f_n$

$$Dm = \frac{\sum |x_i - \overline{x}| f_i}{\sum f_i}$$

Característica do Desvio Médio

- Depende de todos os valores da distribuição;
- Pode ser calculado a partir da média ou da mediana;
- É mínimo quando calculado a partir da mediana.

Variância - é a média quadrática dos desvios tomados em relação à média aritmética. Para dados não agrupados em distribuição de freqüências a variância é obtida por:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} \text{ (população)} \quad \text{e} \quad s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \text{ (amostras)}$$

Para dados agrupados em distribuição de freqüências a variância é obtida pelas seguintes expressões:

Variável: $x_i = x_1, x_2 \ldots x_n$ Freqüências: $f_i = f_1, f_2 \ldots f_n \Rightarrow \sum f_i = n$

$$\sigma^2 = \frac{\sum (xi - \bar{x})^2 f_i}{\sum fi} \quad \text{ou} \quad \sigma^2 = h^2\left[\frac{\sum f_i \alpha_i^2}{\sum f_i} - \left(\frac{\sum f_i \alpha_i}{\sum f_i}\right)^2\right]$$

No caso de amostras a variância é representada por s^2 e o denominador da expressão por (n-1) no lugar de n ou do $\sum f_i$.

Desvio Padrão (σ ou s) $\Rightarrow$ é a raiz quadrada da variância.

Principais Propriedades do Desvio Padrão e da Variância

1 - O desvio padrão ou a variância de uma constante é zero.

$x_i = a, a, a \ldots \Rightarrow \bar{x} = a \Rightarrow$ Dispersão nula.

2 - Somando ou subtraindo uma constante a todos os valores da série, o desvio padrão e a variância não se alteram, enquanto a média aritmética ficará aumentada ou diminuída, respectivamente, desta constante.

$x_i = x_1, x_2, \ldots x_n$ Média $= \bar{x}$

$y_i = (x_1 + a), (x_2 + a), \ldots (x_n + a)$ Média $= \bar{x} + a$

Desvios de y_i em torno da média

$(x_1 + a) - (\bar{x} + a) = x_1 - \bar{x}$

$(x_2 + a) - (\bar{x} + a) = x_2 - \bar{x}$ $\Rightarrow$ variância e desvio padrão igual a série x_i

$\vdots \quad \vdots \quad \vdots$

$(x_n + a) - (\bar{x} + a) = x_n - \bar{x}$

3 - Multiplicando ou dividindo os valores da série por uma constante, o desvio padrão e a média aritmética ficarão multiplicados ou divididos, respectivamente, pela constante. A variância ficará multiplicada ou dividida pelo quadrado da constante.

		Desvio de y_i em torno da média aritmética
$x_i = x_1, x_2, ..., x_n$	Média = $\bar{x}$	$x_1a - a\bar{x} = a(x_1 - \bar{x})$
$y_i = x_1a, x_2a, ..., x_na$	Média = $\bar{x}a$	$x_2a - a\bar{x} = a(x_2 - \bar{x})$
		$\vdots \quad \vdots \quad \vdots$
		$x_na - a\bar{x} = a(\bar{x}_n - \bar{x})$

4 - A variância combinada de duas ou mais séries de valores é calculada pela seguinte expressão:

$$s^2 = \frac{(n_1 -1)[s^2_1 + (\bar{X} - \bar{X}_1)^2] + (n_2-1)[s^2_2 + (\bar{X} - \bar{X}_2)^2] + ... + (n_k -1)[s^2_k + (\bar{X} - \bar{X}_k)^2]}{(n_1 + n_2 + ... n_k) - k}$$

$$s^2 = \frac{S(n_i-1)[s_i^2 + (\bar{X} - \bar{X}_i)^2]}{n_i - k}$$

Para dados populacionais (n -1) é substituído por n.

5 - Devemos calcular o desvio padrão quando procuramos um parâmetro que se revista de um máximo de estabilidade, além de uma plena utilização no cálculo de outras estatísticas.

Controle de Charlier para o Desvio Padrão - destina-se a verificar a exatidão dos cálculos efetuados

$$S\, f(\alpha_i + 1)^2 = f\alpha^2_i + f\alpha_i + n$$

Relação entre desvio quartil, desvio médio e desvio padrão. $D_q < D_m < s$

- MEDIDAS RELATIVAS DE DISPERSÃO

Têm grande aplicação, visto que possibilita comparar dispersão de universos diferentes (em unidade de medida e/ou valor médio), já que seus resultados são expressos em percentagens devido a relação entre as medidas de dispersão absoluta e de tendência central. São medidas usadas para análises comparativas.

- *DESVIO QUARTIL REDUZIDO*

É a relação percentual entre o Desvio quartil e a mediana.

$$D_r = \frac{Q_3 - Q_1}{2M_e} \cdot 100$$

- Coeficiente de variação de *PEARSON* - é a relação percentual entre o desvio padrão e a média aritmética. $\mathbf{CV = \frac{s}{\bar{x}} \cdot 100}$

Variância Relativa: $\mathbf{s_r = \frac{s^2}{\bar{x}} \cdot 100}$

Coeficiente de Variação de *THORUDIKE*: $\mathbf{Cv_t = \frac{s}{M_e} \cdot 100}$

MOMENTOS, ASSIMETRIA E CURTOSE

Momentos - são desvios de uma série de valores calculados em relação a média aritmética ou a um valor arbitrário. A partir dos momentos centrados na média aritmética, calculamos as principais medidas de Dispersão, bem como de Assimetria e Curtose.

Momentos Auxiliares - são momentos centrados numa origem qualquer.

$$mr = \frac{\sum_{i=1}^{n}(x_i - A)^r}{n} \qquad mr = \frac{\sum_{i=1}^{n}(x_i - A)^r f_i}{\sum_{i=1}^{n} f_i}$$

r = 0, 1, 2 , 3 , 3 , 4

Quando A = 0, teremos o momento em relação a origem das coordenadas. O primeiro momento (r=1) em relação a origem é a média aritmética.

Momentos Centrais

Se A $=\overline{x}$, teremos o que se denomina "momento central", que tem por origem a própria média aritmética, definido pela relação:

$$mr = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^r}{n} \qquad r = 0, 1, 2, 3, 4 \ldots$$

Para dados agrupados em distribuição de freqüência:

$$mr = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2 f_i}{\sum f_i} \quad e \quad mr = h^r \frac{\sum f_i \alpha_i}{\sum f_i}$$

Em relação aos momentos centrais observa-se o seguinte:

a) O momento de ordem zero em relação à média é igual a 1.
b) O 1º momento em relação à média é igual a zero.
c) O 2º momento em relação à média é igual à variância.
d) O 3º momento em relação à média é igual a zero e estuda a assimetria.
e) O 4º momento em relação à média estuda a curtose

Os três primeiros itens ocorrem em qualquer tipo de distribuição.

O item "d" só é verdadeiro para as distribuições simétricas, isto é, para estas todos os momentos de ordem ímpar são nulos. A série 2, 4, 5, 10, 14 não é simétrica, visto que o momento central de 3ª ordem é igual a 42.

Momentos Centrais em função dos Momentos Auxiliares:

Os momentos centrais em função dos momentos auxiliares podem ser obtidos da seguinte relação: $M_r = (m - m_1)^r$ $\qquad r = 2, 3, 4, \ldots$

Para isto basta desenvolver o segundo membro e, posteriormente, transformar os expoentes de "m" em índices, e reduzir os termos semelhantes.

M_o - sempre é igual a 1 e M_1 - sempre é igual a zero
M_2 - $(m - m_1)^2 = m^2 - 2mm_1 + m_1^2 = m_2 - 2m_1m_1 + m_1^2 \Rightarrow M_2 = m_2 - m_1^2$
$M_3 = m_3 - 3m_2m_1 + 2m_1^3$
$M_4 = m_4 - 4m_3m_1 + 6m_2m_1^2 - 3m_1^4$

Controle de Charlier, para o cálculo dos momentos pelo processo abreviado (Momentos Auxiliares).

$\sum f(\alpha + 1) = \sum f\,\alpha + n$
$\sum f(\alpha + 1)^2 = \sum f\,\alpha^2 + 2\sum f\,\alpha + n$
$\sum f(\alpha + 1)^3 = \sum f\,\alpha^3 + 3\sum f\,\alpha^2 + 3\sum f\,\alpha + n$
$\sum f(\alpha + 1)^4 = \sum f\,\alpha^4 + 4\sum f\,\alpha^3 + 6\sum f\,\alpha^2 + 4\sum f\,\alpha + n$

Momentos sob forma abstrata - para evitar unidades particulares, podem ser definidos momentos abstratos centrados na média, designados geralmente de ***Coeficientes Alfas***, a fim de que possamos comparar dados heterogêneos. $\boldsymbol{\alpha_r = \frac{Mr}{s^r}}$

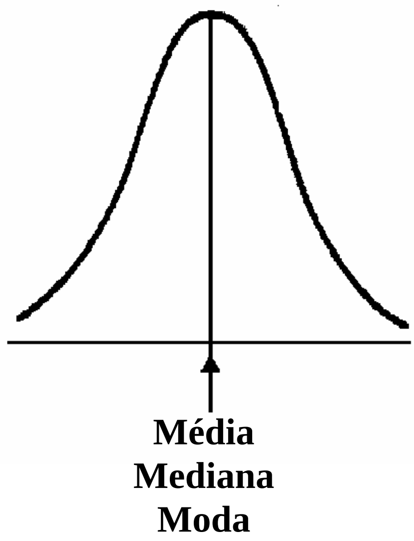
Em qualquer distribuição α_1 e α_2 serão sempre "zero" e "um", enquanto α_3 e α_4 revelam, respectivamente, a ***Assimetria*** e a ***Curtose*** da distribuição.

ASSIMETRIA

É o grau de desvio ou afastamento da simetria de uma distribuição. Quando a curva é simétrica, a média, a mediana e a moda coincidem, num mesmo ponto, de ordenada máxima, havendo um perfeito equilíbrio na distribuição. Quando o equilíbrio não acontece, isto é, a média, a mediana e a moda recaem em pontos diferentes da distribuição esta será assimétrica; enviesada a direita ou esquerda.

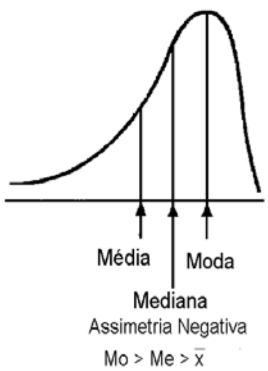
Distribuição Simétrica $\Rightarrow$ Assimetria (S) = 0

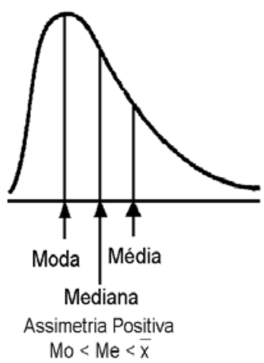
$M_o = M_e = \bar{x}$ e $Q_3 - M_e = M_e - Q_1$



Distribuição assimétrica ***Negativa*** ou enviesada a esquerda - quando os valores se concentram na extremidade superior da escala e se distribuem gradativamente em direção à extremidade inferior.

Distribuição assimétrica ***Positiva*** ou enviesada a direita quando os valores se concentram na extremidade inferior da escala e se distribuem gradativamente em direção à extremidade superior.





Cálculo Da Assimetria

Critério de **Bowley**

$Q_3 - M_e \neq M_e - Q_1 \quad \Rightarrow \quad Q_3 + Q_1 - 2M_e \neq 0$ (assimetria absoluta)

$$S = \frac{Q_3 + Q_1 - 2M_e}{Q_3 - Q_1}$$ (assimetria relativa) $\Rightarrow$ sendo: $-1 \leq S \leq 1$

Será ***Positivo*** a medida que o terceiro quartil se afasta da mediana, enquanto que o primeiro quartil se aproxima da mesma, tendo como limite: $Q_1 = Q_2$, quando a assimetria assume o valor máximo positivo: (S = 1).

$$\mathbf{S = \frac{Q_3 - Q_1}{Q_3 - Q_1} = 1}$$

Será ***Negativo*** a medida que o primeiro quartil afasta-se da mediana, enquanto o terceiro quartil aproxima-se da mesma, dando como limite: $Q_3 = Q_2$, quando a assimetria assume valor máximo negativo:

$$\mathbf{S = \frac{-Q_3 + Q_1}{Q_3 - Q_1} = -1}$$

Critério de ***Kelley*** - Para corrigir parte do inconveniente de se desprezar 50% das ocorrências (Critério de *Bowley*), *Kelley* aconselha o uso dos "**Centis**" eqüidistantes da mediana, tais como C_{10} e C_{90}, para cálculo da assimetria.

$$\mathbf{S = \frac{C_{90} + C_{10} - 2M_e}{C_{90} - C_{10}}}$$ Cujos limites variam também de [-1 a +1]

Coeficiente de Assimetria de ***Pearson*** - A medida que a distribuição deixa de ser simétrica, a média, a mediana e a moda vão se afastando, aumentando cada vez mais a diferença entre elas.

$$\alpha = \frac{3(\bar{x} - Me)}{s}$$

Coeficiente do Momento de Assimetria - Uma medida importante de assimetria que utiliza o terceiro momento central, expresso sob forma abstrata, ou relativa.

$$\alpha_3 = \frac{M_3}{s^3}$$

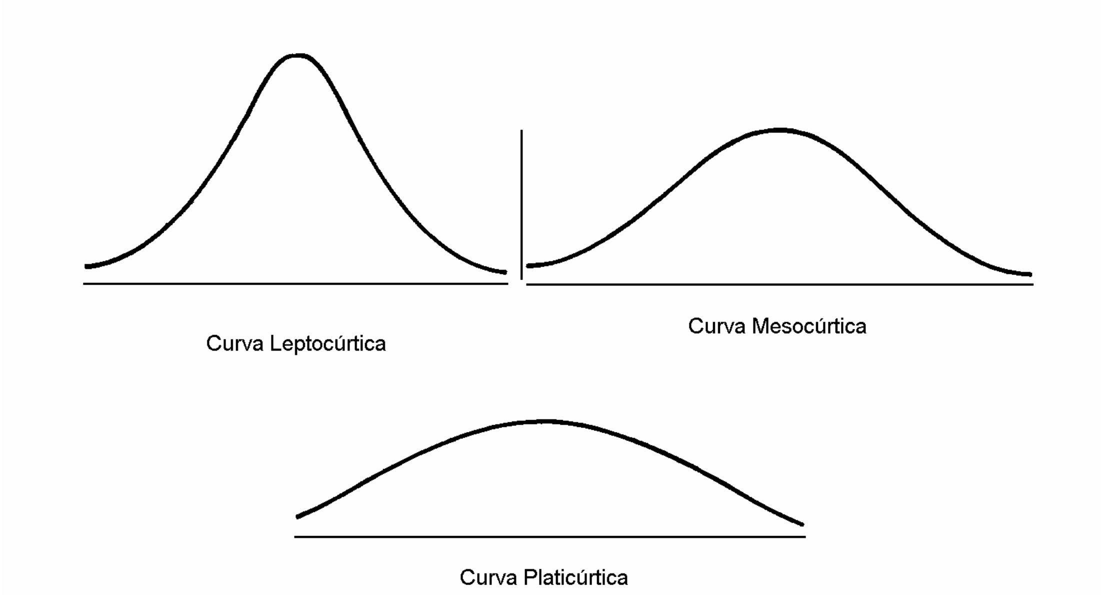
$\alpha_3 = 0$ - distribuição simétrica
$\alpha_3 < 0$ - distribuição assimétrica negativa

$\alpha_3 > 0$ - distribuição assimétrica positiva

CURTOSE

É o grau de achatamento de uma distribuição, em relação a distribuição normal. A curtose pode ser de três tipos:

- **Mesocúrtica** - quando a distribuição é normal.
- **Leptocúrtica** - quando a distribuição é mais pontiaguda que a normal
- **Platicúrtica** - quando a distribuição é mais achatada que a normal.



MEDIDAS DE CURTOSE

Coeficiente Percentílico de Curtose:

$$C = \frac{D_q}{C_{90} - C_{10}} \quad \Rightarrow \quad C = \frac{Q_3 - Q_1}{2(C_{90} - C_{10})}$$

Se C = 0,263 a distribuição é Mesocúrtica
Se C < 0,263 a distribuição é Leptocúrtica
Se C > 0,263 a distribuição é Platicúrtica
Coeficiente do momento de curtose (α_4) - utiliza o quarto momento central, relativo ao quadrado da variação sendo: $\boldsymbol{\alpha_4 = \frac{M_4}{s^4}}$

Para distribuição mesocúrtica: $\alpha_4 = 3$; leptocúrtica: $\alpha_4 > 3$ e Platicúrtica: $\alpha_4 < 3$.

Exercício Resolvido

2.1 Determinar a variância das amostras constituída dos seguintes elementos.

x_i = 7; 10; 12; 15; 16; 18; 20 e y_i = 12; 15; 18; 23; 25; 26; 28.

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

$$s^2 = \frac{49 + 16 + 4 + 1 + 4 + 16 + 36}{6} \qquad s^2 = \frac{81 + 36 + 9 + 4 + 16 + 25 + 49}{6}$$

$s^2 = 16$ e $s = 4$ $\qquad s^2 = 36{,}6667 \Rightarrow \quad s = 6{,}06$

2.2 Calcular a variância da questão 1.09

$$s^2 = h^2\left[\frac{\sum f_i\alpha_i^2}{\sum f_i - 1} - \left(\frac{\sum f_i\alpha_i}{\sum f_i - 1}\right)^2\right] = 25\left[\frac{630}{349} - \left(\frac{250}{349}\right)^2\right]$$

$$s^2 = 25 \cdot 1{,}2921 \Rightarrow s^2 = 32{,}3025 \Rightarrow s = 5{,}68$$

2.3 Calcular os primeiros momentos, em relação a origem dos números 1, 2, 3, e 4

$$mr = \frac{\sum(x_i - A)^r}{n}$$

$$m_0 = \frac{\sum x^0}{n} = \frac{1}{4}\left(1^0 + 2^0 + 3^0 + 4^0\right) = 1$$

$$m_1 = 2{,}5; \quad m_2 = 7{,}5; \quad m_3 = 25; \quad m_4 = 88{,}5$$

2.4 Calcular os primeiros momentos centrais dos números 1, 2, 3, e 4.

$$Mr = \frac{\sum (x_i - \bar{x})^r}{n}$$

$$M_0 = \frac{\sum (x - \bar{x})^0}{n} = \frac{(1-2{,}5)^0 + (2-2{,}5)^0 + (3-2{,}5)^0 + (4-2{,}5)^0}{4} = 1$$

$M_1 = 0$; $M_2 = 1{,}25$; $M_3 = 0$ $M_4 = 2{,}5625$

2.5 Determinar os quatros primeiros momentos centrados na média, da seguinte distribuição.

Alturas (cm)	f_i	$\propto_i$	$f_i \propto_i$	$f_i \propto_i^2$	$f_i \propto_i^3$	$f_i \propto_i^4$
150 ├— 158	5	-2	-10	20	-40	80
158 ├— 166	18	-1	-18	18	-18	18
166 ├— 174	42	0	0	0	0	0
174 ├— 182	27	1	27	27	27	27
182 ├— 190	8	2	16	32	64	128
$\sum$	100	-	15	97	33	253

Momentos auxiliares:

$$m_r = h^r \frac{\sum f_i \alpha_i^r}{\sum f_i}$$

$m_1 = 1{,}20$; $m_2 = 62{,}08$; $m_3 = 168{,}96$ e $m_4 = 10.362{,}88$
Momentos centrais

$M_0 = 1$; $M_1 = 0$; $M_2 = 60{,}64$; $M_3 = -51{,}07$ e $M_4 = 9.995{,}62$

2.6 Calcular a Assimetria e Curtose da questão 2.5.

Assimetria : $\propto_3 = \dfrac{M_3}{s_3} = \dfrac{-51{,}07}{489{,}13} \Rightarrow \propto_3 = -0{,}10$

Curtose: $\propto_4 = \dfrac{M_4}{s_4} = \dfrac{9.995{,}62}{3.853{,}93} \Rightarrow \propto_4 = 2{,}59$

2.7 Com os dados da distribuição de frequência abaixo calcular: Desvio Padrão, Assimetria e Curtose a partir dos momentos.

X_i	f^i	$\propto^i$	$f^i \propto_i$	$f_i \propto_i^2$	$f_i \propto_i^3$	$f_i \propto_i^4$
0 ⊢ 5	5	-3	-15	45	-135	405
5 ⊢ 10	7	-2	-14	28	-56	112
10 ⊢ 15	11	-1	-11	11	-11	11
15 ⊢ 20	15	0	0	0	0	0
20 ⊢ 25	18	1	18	18	18	18
25 ⊢ 30	13	2	26	52	104	208
30 ⊢ 35	9	3	27	81	243	729
35 ⊢ 40	5	4	20	80	320	1280
$\sum$	83	-	51	315	483	2763

$m_r = h^r \dfrac{\sum f_i \alpha_i^r}{\sum f_i}$ $m_1 = 3{,}07$; $m_2 = 94{,}88$; $m_3 = 727{,}41$ e $m_4 = 20.805{,}72$

Momentos centrais

$M_2 = m_2 - m_1^2 \Rightarrow M_2 = 85{,}46 \Rightarrow$ Desvio padrão (s) = 9,244
$M_3 = m_3 - 3m_2 m_1 + 2m_1^3 \Rightarrow M_3 = -127{,}58 \Rightarrow$ Assimetria (α_3) = -0,16

$M_4 = m_4 - 4m_3m_1 + 6m_2m_1 - 3m_1^4 \Rightarrow M_4 = 20.805{,}72 \Rightarrow$Curtose ($\alpha_4$)= 1,83

EXERCICIOS PROPOSTOS

01. Completar os dados que faltam para as seguintes distribuição:

x_i	f_i	F_i	fr_i
1	4		0,04
2	8		
3		30	0,18
4	27		0,27
5	15	72	
6		83	
7	10	93	0,10
8			
$\sum$			

02. Encontrar a freqüência correspondente à terceira classe da distribuição abaixo, sabendo-se que a média é igual a 11,50.

x_i	5	8	13	18	25
f_i	4	5	...	3	1

03. Sejam os seguintes níveis de Glicose no sangue, em jejum, de 50 crianças:

51 61 57 63 60 57 82 68 59 72 68 73 75 66 61 76 72 75 64 69 68 69 55 65 67
59 66 76 65 64 71 72 73 54 81 60 80 66 52 71 75 74 73 65 74 84 56 62 83 62

Obter uma distribuição de freqüência por classes e calcular as freqüências relativas (simples e acumulada).

04. Um colégio de 1º grau funciona em três turnos e possui alunos nas quatro séries em todos os turnos. Preparar uma tabela, com o título, onde possamos revelar o nº de alunos desse estabelecimento, por série e ao mesmo tempo por turno.

05. Usando o processo abreviado, calcule a média e a variância amostral:

x_i	30.000	30.002	30.004	30.006	30.008	30.010
F_i	10	22	36	46	50	52

06. As notas finais de um estudante em matemática, estatística, português e história são respectivamente 82, 86, 91e 71, atribuindo-se a essas disciplinas, respectivamente, os pesos 3, 5, 3, 1 a média aritmética será.

07. As estaturas de dois grupos de indivíduos, o primeiro com 300 integrantes e o segundo com 124, as médias aritméticas foram respectivamente 166,82cm e 161,92cm. Qual a média aritmética do conjunto?

08. Trinta e seis estudantes foram submetidos a um exame de biologia, obtendo as seguintes notas:

93 98 72 84 88 90 78 80 89 94 95 77 81 75 72 76 70
76 83 99 73 83 87 91 83 92 90 92 77 86 86 74 87 81
98 94

a. Construa uma distribuição de freqüência por classe e calcule as freqüências relativas (simples e acumulada)
b. Determine a média, a mediana e o desvio padrão dessas notas.

09. Os valores abaixo referem-se ao número de itens defeituosos por lote de 40. Disponha os dados numa distribuição de freqüência e determine a média, a mediana e a moda da distribuição . Faça um histograma.

6 3 4 7 10 8 5 12 9 1 10 11 8 7 3 2 4 0 5 1
7 3 5 8 4 0 7 3 8 0 9 8 2 3 5 3 6 2 3 5

10. Determinar a média e o desvio padrão relativo ao nº de acidentes por dia em certa rodovia:

Nº de acidentes	0	1	2	3	4
Nº de dias	15	16	12	9	8

11. Calcule o salário médio horário de 100 funcionários de uma clínica, onde 60 empregados recebem R$ 0,80 / h; 30 recebem R$ 1,20 /h e os demais ganham R$ 2,00 / h.

12. Dado o conjunto de números : x_i = 4, 4, -3, 5, 5, 2, 12, 4, -5 e 7. Calcule a média aritmética.

13. Em certa Empresa trabalham 4 analistas de mercado, 2 supervisores, 1 chefe de seção e 1 gerente, que ganham respectivamente: Cr$ 130.000,00; Cr$ 160.000,00; Cr$ 175.000,00 e Cr$ 250.000,00. Qual o valor do salário médios desses funcionários.

14. Gastamos em janeiro $ 10.000,00 para comprar um produto que custou $ 100,00 a unidade. em fevereiro gastamos $ 24.000,00 para comprar o mesmo produto ao preço unitário de $ 120,00. Qual o custo médio do artigo neste dois meses.

15. Os custos médios de produção dos produtos A, B e C foram, respectivamente R$10,00 , R$12,00 e R$15,00. Determine o custo médio dos três produtos, sabendo que em um mês, foram produzidas 1.000 unidades do produto A, 500 do produto B e 500 unidades do produto C.

16. Uma amostra de 20 volumes transportados pelo correio aéreo, apresentou os seguintes pesos: 1 volume com 6 kg, 3 com 3 kg, 5 com 2kg, 3 com 1,3 kg, 6 com 1kg e 2 com 0,5kg. Qual o peso médio desses volumes?

17. Dez alunos gastaram 72, 80, 75, 90, 85, 74, 70, 83, 86 e 78 minutos, respectivamente, para responder todas as questões de uma prova. qual o tempo médio gasto pelo grupo?

18. A soma dos pesos dos componentes de uma equipe de Vôlei corresponde a 648kg. A relação entre o número de componentes e a média aritmética é de 1/8. Determinar o número de componentes do grupo e a média dos pesos.

19. Em uma prova de estatística a turma A de 30 alunos obteve média de rendimento igual a 6 e a turma B, de 42 alunos, média de 5. A média de rendimento das duas turmas foi:

20. Para obter peso médio de um conjunto de ratos tanto se pode pesar cada rato por vez como pesar todos os ratos de uma vez. No entanto obtemos mais informações se pesarmos um rato por vez. Por que?

21. De acordo com alguns, Estatística é uma ciência engraçada porque, se nós formos ao restaurante e eu comer um frango, enquanto você não come nenhum, em média comemos, cada um, meio frango e estamos muito satisfeitos. Discuta a validade desta média.

22. Se: $z_1 = x_1 + y_1$, $z_2 = x_2 + y_2$, . . . , $z_n = x_n + y_n$ então z é:

23. O salário médio anual de todos os empregados de uma loja foi de R$ 5.000,00. Os salários médios pagos aos empregados do sexo masculino e feminino da loja foram R$ 5.200,00 e R$ 4.200,00 respectivamente. Qual a percentagem de empregados de cada sexo.

24. Para um dado concurso, 60% dos candidatos eram do sexo masculino e conseguiram média de 70 pontos em determinada prova. Sabendo-se que a média geral dos candidatos (independentemente do sexo) foi de 64 pontos, qual a média dos candidatos do sexo feminino?

25. A nota média de uma turma mista de 50 alunos foi 5,3 , sendo 5,0 a média dos meninos e 8,0 das meninas. Quantos meninos e meninas havia na turma?

26. O salário médio pago aos empregados da firma é R$ 7.100,00 . Os salários médios pagos aos empregados especializados e não especializados são respectivamente R$ 8.000,00 e R$ 5.000,00 . Determinar a porcentagem dos empregados especializados e não especializados da firma.

27. Os desvios tomados em relação à média arbitrária $A_0 = 9$ de um conjunto de números são {-4; -1; 2; 0; 3; -3; 5; 1}. A média aritmética do conjunto será:

28. Um pesquisador da rádio XY aborda 40 transeuntes ao acaso e pergunta-lhes a idade obtendo as seguintes informações:

38 21 16 32 39 35 26 39 25 39 22 23 27 44 30 32 35 28 17 26
22 28 39 18 37 42 40 39 22 21 40 21 15 26 43 40 37 33 24 30

Apresente os dados na forma de uma distribuição de freqüências, por classes e calcule a média e o desvio padrão amostral.

29. Foram encontrados os seguintes valores para a transaminase da alamina em indivíduos normais:

12; 15; 18; 12; 17; 17; 16; 16; 14; 12; 14; 17; 16; 15; 16; 18; 16; 12; 18; 18.
Calcule a média e o desvio padrão dos dados.

30. Cronometrando o tempo para várias provas de uma gincana automobilística, encontramos para duas das equipes participantes:

- Equipe I (40 provas): tempo médio = 45s e variância = $400s^2$
- Para a Equipe II, compomos o seguinte quadro:

Tempo(s)	20	40	50	80	Total
Nº de provas	10	15	30	5	60

a) Qual a média, o desvio padrão da equipe II
b) Qual a equipe que apresentou resultados mais homogêneos? Justificar.

31. Uma prova consta de três questões com pesos iguais a 3,2 e 1, respectivamente, para a primeira, a segunda e terceira questão. Se um aluno obteve 8,0 na prova 8,5 na primeira questão e 6,5 na segunda, que grau ele conseguiu na terceira questão?

32. Dados os conjuntos de números:
A = {100, 101, 102, 103, 104, 105} e B = {0; 1; 2; 3; 4; 5} o que podemos afirmar sobre a média aritmética de A em relação a B.

33. A evolução das vendas dos últimos três meses apresentou os seguintes índices: 122,31; 132,42 e 115,32 . Determinar o aumento médio percentual no período.

34. A contagem de bactérias, em certa cultura, aumentou de 500 para 2.000 em três dias. Qual foi a porcentagem média de acréscimo por dia?

35. calcule a média geométrica para as séries:

a) x_i = 8, 15, 10, 12 b) x_i = 3, 4, 5, 6, 7, 8,

c)

x_i	8	9	10	11	12
f_i	12	10	7	5	3

36. Encontre a média harmônica para as séries:

a) 5, 7, 12, 15

b)

x_i	2	3	4	5	6
f_i	3	4	6	5	2

37. Encontrar dois números cuja média aritmética é 50 e a média harmônica é 32.

38. A média geométrica dos preços de dois produtos , A e B é R$ 7,20; enquanto a média aritmética é R$ 9,00. Determinar os preços dos produtos A e B.

39. Calcule a média geométrica e harmônica dos dados da questão 17.

40. Determine a média geométrica dos números 2, 6, e 18 .

41. A média geométrica de dois números é 8 e a média aritmética é 10. Quais são esses números?

42. Qual a média harmônica dos números 2, 6 e 18? Compare o resultado obtido com a média aritmética do inverso desses números .

43. As cidades A, B e, C, são eqüidistantes uma das outras. Um motorista viaja de A para B a 30 km/h, de B para C a 40 km/h e de C para A a 50 km/h. Qual a velocidade média desenvolvida?

44. Sendo A, B, e C respectivamente, médias aritmética, geométrica e harmônica de um mesmo conjunto de números, é correto afirmar que:
a) $A \geq B \geq C$; b) $A \leq B \leq C$ c) $A \geq B \leq C$

d) $A \leq B \geq C$ e) a e b corretas.

45. Sabe-se que o consumo em kg de ração/semana de frangos com idade de 1 a 9 semanas é: 108 189 287 371 462 637 833 924 959 .
Calcular: média aritmética, desvio padrão e coeficiente de variação

46. Os pesos (kg), ao nascer, de 12 bezerros machos de raça Charolesa são: 47; 46; 45; 47; 37; 34; 45; 25; 48; 40; 40; 41.

Calcular: média aritmética, desvio padrão e coeficiente de variação.

47. Os dados abaixo são comprimento de crianças ao nascer. Calcular: média aritmética, variância, desvio padrão, coeficiente de Variação e erro da média.

46; 47; 51; 46; 50; 50; 52; 49; 48; 51; 49; 52; 48; 47; 49.

48. Calcular o desvio padrão e coeficiente de variação dos dados referentes ao peso (kg) de ovinos: 37; 35; 38; 36; 35; 40; 42; 44; 41; 30; 37.

49. Em determinada Escola foram aplicados testes a dois grupos (I e II) de alunos. Qual dos grupos possui menor dispersão.

I -	8,	7,	6,	6,	9,	10,	8,	10
II -	6,	6,	8,	9,	9,	7,	5,	6

50. Uma amostra de 2.000 universitários apresentou a seguinte distribuição de alturas.

ALTURAS (cm)	Nº DE ALUNOS
142|---148	46
148|---154	88
154|---160	332
160|---166	676
166|---172	614
172|---178	158
178|---184	65
184|---190	21
$\sum$	2.000

a) Determine a média aritmética, o desvio padrão e o coeficiente de variação.

b) Aumente em 5 cm todas as alturas e determine a média aritmética, o desvio padrão e o coeficiente de variação da nova distribuição.

c) Compare e discuta os resultados dos itens: a e b.

51. Anotou-se a altura de 80 pessoas e constatou-se que a altura média foi de 170cm e a variância de 7,4 cm. Posteriormente verificou-se que o

instrumento para medir, usado na experiência tinha 9 cm a menos. Retifique a média e a variância encontradas.

52. A variância de 2 números é 1 e a média aritmética é 8 . Quais são esses dois números?

53. Dados os conjuntos de números A={ 1.000, 1.001, 1.002, 1.003, 1.004, 1.005 } e B={ 0, 1, 2, 3, 4, 5 }, o que podemos afirmar sobre o desvio padrão da série A em relação a série B?

54. Seja s^2 a variância do conjunto $x_i = x_1, x_2 \ . \ . \ . \ x_n$. Caso multipliquemos todos os x_i por uma constante K, a variância dos novos valores obtidos será igual a:

55. O coeficiente de variação de uma distribuição é 0,083%, e o desvio é de 2,49 centímetros. Determine a média aritmética dessa distribuição.

56. O conjunto $x_i = \{ x_1, x_2, \ . \ . \ . \ x_n \}$ possui média aritmética 10 e desvio padrão 3. considerando um outro conjunto y, cujos elementos estão relacionados com os de x, mediante a equação $y_i = 4x_i + 2$, qual será a média aritmética e o desvio padrão do conjunto y?

57. Admitindo-se que em uma distribuição de freqüência, perfeitamente simétrica, contendo 100 dados observados, o somatório do produto de cada freqüência absoluta pelo valor absoluto dos desvios da respectiva observação em relação à média aritmética é igual a 5.500. Determinar o valor do desvio padrão nesta distribuição.

58. Sabendo-se que um conjunto de dados apresenta para média aritmética e para desvio padrão, respectivamente 18,3 e 1,47. Calcule e interprete o coeficiente de variação.

59. Um fabricante de caixas de cartolina fabrica três tipos de caixa. Testa-se a resistência de cada caixa, tomando uma amostra de 100 caixas e determinando a pressão necessária para rompimento da mesma. Que

tipo de caixa apresenta a menor variação absoluta na pressão de ruptura? Qual delas apresenta maior variação relativa?

Tipos de Caixa	A	B	C
Pressão média de ruptura (bária)	150	200	300
Desvio padrão das pressões (bária)	40	50	60

60. Consideremos a distribuição de freqüência, com dados hipotéticos, referente aos salários recebidos por 40 operários de uma empresa X.. Determine o salário médio, modal e mediano dos 40 operários.

SALÁRIOS (R$)	Nº DE EMPREGADOS
300 |--- 400	4
400 |--- 500	8
500 |--- 600	10
600 |--- 700	15
700|--- 800	3
TOTAL	40

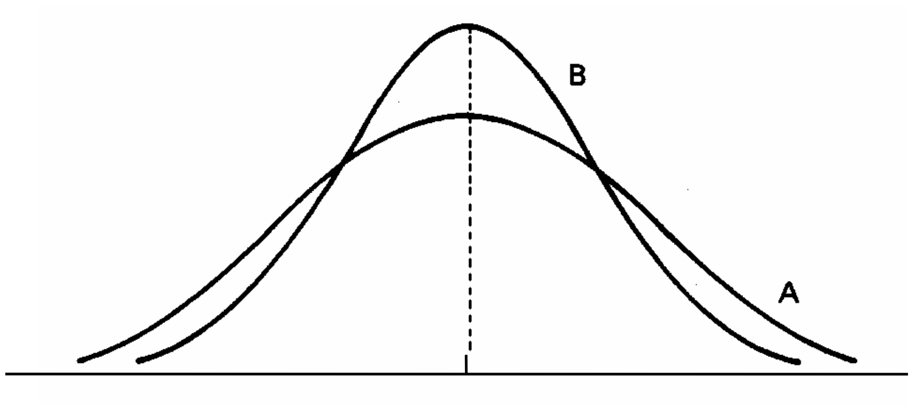
61. Em uma distribuição, a média aritmética é 36 e a mediana é 38, com desvio padrão igual a 2. Para analisar a distribuição:
a) conhecer a média e a mediana é dispensável.
b) através da média e da mediana conhecemos a assimetria.
c) através do desvio padrão encontramos o segundo quartil.
d) é dispensável o conhecimento do desvio padrão

62. Das características abaixo, uma delas é da moda:
a) é influenciada pela grandeza dos afastamentos extremos.
b) é demasiadamente afetada pelos valores extremos da série.
c) a soma dos afastamentos é um mínimo.
d) é desprovida de significação estatística a menos que a distribuição compreenda um grande número de dados.

63. As medidas estatísticas que permitem uma análise dos dados mais aprofundados são:

a) média e desvio padrão.
b) mediana e desvio padrão.
c) quartil e mediana.
d) média e mediana.
e) mediana e moda.

64. Examinando a figura abaixo, podemos dizer:



a) ( ) O desvio padrão da distribuição A é maior do que o da distribuição B, e as médias são iguais.

b) ( ) O desvio padrão de A é menor do que o de B e as médias são diferentes.

c) ( ) O desvio padrão de A é igual ao de B, independentemente do valor da média.

d) ( ) As distribuições possuem o mesmo coeficiente de variação.

65. Realizou-se uma prova de matemática para duas turmas. Os resultados foram os seguintes:

Turma A: $\bar{x} = 5$ e $s = 2{,}5$
Turma B: $\bar{x} = 5$ e $s = 2$

Com esses resultados, podemos afirmar:
a) ( ) A turma B apresentou maior dispersão absoluta.
b) ( ) A dispersão relativa é igual à dispersão absoluta.

c) ( ) Tanto a dispersão absoluta quanto a relativa são maiores para a turma B.
d) ( ) A dispersão absoluta de A é maior do que a de B, mas em termos relativos as duas turmas não diferem quanto ao grau de dispersão das notas.

66. Estudando o grau de achatamento de uma determinada distribuição, um pesquisador concluiu que essa era mesocúrtica.
Marque as alternativas corretas

a) ( ) Esta distribuição apresenta Mo = $\bar{x}$ > Me e C = 0,263:
b) ( ) O grau de assimetria é nulo enquanto C. 0,263:
c) ( ) A Mo = Me = $\bar{x}$ (assimetria nula) e C = 0,263:
d) ( ) Esta distribuição é simétrica com Mo = $\bar{x}$ = Me, dividindo-a em duas partes iguais.

67. Sabendo-se que em uma distribuição a média aritmética é a metade da moda e que a mediana é igual a moda menos um então o valor da média, moda e mediana respectivamente são:

68. Calcular a média aritmética ,variância, assimetria e curtose relativo aos pontos do teste de habilitação dos alunos da UFS - 1996.

PONTOS	NÚMERO DE ALUNOS
27 ├— 32	1
32 ├— 37	5
37 ├— 42	8
42 ├— 47	12
47 ├— 52	8
52 ├— 57	10
57 ├— 62	6
$\sum$	50

69. Suponha-se que a estrutura etária dos alunos de uma determinada unidade escolar se apresente segundo a distribuição:

IDADE (anos)	NÚMERO DE ALUNOS
7 |— 9	197
9 |— 11	372
11 |— 13	527
13 |— 15	114
15 |— 17	49
17 |— 19	25
19 |— 21	3
$\sum$	1.287

Calcular: média aritmética, desvio padrão, mediana, moda e assimetria.

70. Calcule os quatro primeiros momentos naturais dos conjuntos.

a) $x_i = \{1, 3, 5, 7, 9\}$ b) $y_i = \{0, 2, 4, 6, 8\}$

71. Calcule os quatro primeiros momentos centrados na média, dos conjuntos.

a) $x_i = \{1, 3, 5, 7, 9\}$ b) $y_i = \{0, 2, 4, 6, 8\}$

72. Calcule os quatro primeiros momentos centrados na origem 7, para o conjunto.

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

73. Calcule os quatro primeiros momentos centrados na origem 6, para o conjunto.

$y_i = \{0, 2, 4, 6, 8\}$

74. Determine o valor do desvio padrão, utilizando-se dos momentos obtidos no exercício 71.

75. Determine os coeficientes de assimetria e de curtose para os conjuntos do exercício 71.

76. Em uma distribuição os momentos de segunda e terceira ordem, centrados na média, são respectivamente 4 e 3,2 . Em uma outra distribuição, os mesmos momentos são 16 e 32. Qual das duas distribuições é mais assimétrica?

77. Nas mesmas distribuições do problema anterior, os momentos de 4ª ordem, centrados na média, são 51,2 e 742,4 , respectivamente. Qual das duas apresenta maior índice de agudeza?

78. Qual das duas distribuições (dos exercícios 76 e 77) apresenta maior semelhança com a curva normal?

a) Quanto à assimetria b) Quanto à agudeza.

79. Determine os quatros primeiros momentos centrados na média e os coeficientes de assimetria e curtose para as distribuições de freqüências abaixo. Comente os resultados.

a) SALÁRIO/HORA (R$)	Nº DE FUNCIONÁRIO	b) IDADE MÉDIA (anos)	Nº DE PESSOAS
0,5 |--- 1,5	18	10 |--- 13	13
1,5 |--- 2,5	22	13 |--- 16	12
2,5 |--- 3,5	57	16 |--- 19	14
3,5 |--- 4,5	68	19 |--- 22	0
4,5 |--- 5,5	35	22 |--- 25	11
$\sum$	200	$\sum$	50

80. Se o primeiro momento centrado no número 2 é igual a 5, qual o valor da média aritmética?

81. O desvio padrão de uma distribuição simétrica é 3. Qual deveria ser o valor do quarto momento centrado na média para que a distribuição fosse leptocúrtica.

a) Igual a 243 c) Maior que 243 e) a e c estão corretas.
b) Menor que 243 d) Qualquer valor acima de 157

82. Os segundos momentos centrados na média de duas distribuições são 9 e 16 enquanto que os terceiros momentos referidos a mesma origem são -8,1 e -12,8 respectivamente. A distribuição mais desviada a esquerda é:

RESPOSTAS

01⇒18, 11 e 07	02⇒07	05⇒30.003,69 e 7,29
06⇒85	07⇒165,39	08⇒85 ; 85 e 7,95
09⇒5,60; 5,25 e 3,43	10⇒1,65 e 1,35	11⇒R$ 1,04/h
12⇒3,5	13⇒9 meses e 24 dias	14⇒113,33
15⇒R$ 11,75	16⇒1,8 Kg	17⇒79,3 min
18⇒9 e 72	19⇒5,42	

20⇒Os pesos individuais permitem calcular a variância, que mede a dispersão.

21⇒Temos: $\bar{x}$ = 0,5 e s = 0,7 ⇒ a média não tem significado porque a variabilidade é muito grande.

22⇒$\bar{x} + \bar{y}$	23⇒80% e 20%	24⇒55
25⇒45 e 5	26⇒70% e 30%	27⇒9,375
28⇒31,00 e 8,09	29⇒15,4 e 2,14	30⇒45 e 15,00

31⇒9,5

32⇒A média de A é igual a média de B mais a constante de 100

33⇒23,15% 34⇒58,74% 35⇒10,95; 5,22 e 9,29

36⇒8,12 e 3,53 37⇒80 e 20 38⇒R$14,40 e R$ 3,60

39⇒79,052mm e 78,86 mm 40⇒6 41⇒4 e 16

42⇒4,15 43⇒38,3 Km/h 44⇒a

45⇒530,00 ; 340,96 e 64,33% 46⇒41,25 ; 6,74 e 16,34

47⇒49,00 ; 2,00 e 4,08% 48⇒3,9 kg e 10,33%

49⇒1,60 e 20,00% ⇔ 1,15 e 21,57%

50⇒ a) 164,70cm; 7,48cm e 4,77% b) 169,70cm; 7,48cm e 4,63%

51⇒161cm e 7,4cm 52⇒7 e 9

53⇒O desvio padrão de A é igual ao desvio padrão de B.

54⇒K^2s^2 55⇒3000cm 56⇒42 e 12 57⇒68,75

58⇒8% a distribuição é bastante homogênea, isto é, os xis estão bastante próximos de $\bar{x}$, tornando a medida típica da série.

59⇒A e A 60⇒562,50; 629,40 e 580,00 61⇒b

62 ⇒ d 63 ⇒ a 64 ⇒ a

65 ⇒ d 66 ⇒ d 67 ⇒ $\frac{3}{2}$; 3 e 2

68⇒ 47,00; 63,25; α_3 = - 0,135 e α_4 = 2,1

69⇒11,27; 2,20; 11,28; 11,55 e α_3 = 0,69

70⇒a)5; 33; 245 e 1933,8 b)4; 24; 160 e 1132,8

71 ⇒ a) zero; 8; zero e 108,8 b) zero; 8; zero e 108,8

72 ⇒ -2; 12; -56; e 336 73 ⇒ -2; 12; -56; e 336

74 $\Rightarrow$ 2,8   75 $\Rightarrow$ zero e 1,7   76 $\Rightarrow$ a segunda

77 $\Rightarrow$ a segunda   78 $\Rightarrow$ a) primeira   b) segunda

79 $\Rightarrow$ a) zero; -1,35, -0,774 e 4,6068 $\propto_3$ = -0,494 e $\propto_4$ =2,528
b) zero; 18,52; 41,32 e 694,84 $\propto_3$ = 0,520 e $\propto_4$ =2,026

80 $\Rightarrow$ 7   81 $\Rightarrow$ c   82 $\Rightarrow$ 2ª