ESTATÍSTICA

ÍNDICE

1 - ALGUNS CONCEITOS DA ESTATÍTICA

1.1 - INTRODUÇÃO

Cada vez mais se usa a Estatística em qualquer atividade da vida moderna, como bem pode ser comprovado através de uma rápida observação da mídia. Portanto, é imprescindível estudá-la.

Existem duas concepções da palavra ESTATÍSTICA:

- No plural (estatísticas) indica qualquer coleção de dados numéricos reunidos com a finalidade de informar sobre uma determinada atividade.

- No singular, segundo Fisher, é um ramo da Matemática Aplicada dedicado à análise de dados de observação.

Aqui trabalharemos com a segunda concepção.

1.2 - TIPOS DE ESTATÍSTICA

- **Estatística Descritiva**: quando o número de dados é suficientemente grande de modo a dificultar a absorção da informação que se quer estudar, elas devem ser reduzidas até ao ponto de se poder interpretá-las claramente.. Para isso, usa-se os números-síntese, conhecidos como estatísticas descritivas ou simplesmente estatísticas. Sendo assim, **a estatística descritiva é um número que sozinho descreve uma característica de um conjunto de dados**.

- **Estatística Indutiva** ou Inferência estatística: é o processo de generalização que se faz a partir de resultados particulares.

1.3 - FENÔMENO ESTATÍSTICO

Segundo Paulo Cezar Ribeiro da Silva, fenômeno estatístico é qualquer evento que se pretenda analisar, cujo estudo pode ser feito através da aplicação do método estatístico. Eles são divididos em três grupos:

Fenômenos de massa ou coletivo: são aqueles que não podem ser definidos por uma simples observação. A estatística dedica-se ao estudo desses fenômenos.

Exemplo: O índice de natalidade no Brasil.

Fenômenos individuais: são aqueles que irão compor os fenômenos de massa.

Exemplos: Cada nascimento em Belo Horizonte.

Cada preço de cerveja no Espírito Santo.

Fenômenos de multidão: quando as características observadas para a massa não se verificam para o particular.

1. 4 - SOBRE MÉTODOS

1.4.1 - MÉTODO

Segundo Alberto Mesquita Filho**,** método, entre outras coisas, significa *caminho para chegar a um fim ou pelo qual se atinge um objetivo.*

1.4.2 - MÉTODO CIENTÍFICO

De acordo com Mauro Pennafort, método científico é o "*usado nas ciências (exatas e até mesmo em algumas humanas) que consiste em estudar um fenômeno da maneira mais racional possível, de modo a evitar enganos, sempre buscando evidências e provas para as idéias, conclusões e afirmações.* Ou seja, é um "*conjunto de abordagens, técnicas e processos para formular e resolver problemas na aquisição objetiva do conhecimento*."

O método científico se compõe das seguintes fases:

1. Observação do Fenômeno: O fenômeno é observado e desenvolve-se a curiosidade em relação a ele.

2. Experimentação e Medição: Provoca-se o mesmo fenômeno várias vezes, registrando-se todas as possíveis variações e valores relacionados a ele. Nessa fase são feitas cuidadosas medições.

3. Estabelecimento de Leis Científicas: Depois da análise dos dados da experimentação, conclui-se uma *Lei Científica*, que é uma generalização que relaciona os dados que foram estudados. É importante notar que a Lei Científica não é a explicação do por quê daquilo, mas apenas uma descrição (de preferência matemática) do fenômeno.

4. Criação de Hipóteses: Imagina-se explicações para o fenômeno e sua lei. A procura da explicação (do por quê) leva, muitas vezes à criação de um Modelo. A hipótese ou modelo mais simples e elegante é escolhido como provável explicação para o fenômeno estudado.

Um modelo é uma descrição formal de um fenômeno, uma maneira de entender o fenômeno, que é capaz de fazer predições.

5. Teste das Hipóteses: A hipótese escolhida deve explicar novas observações e novos fenômenos. O modelo relacionado a esta hipótese deve ser capaz de fazer previsões sobre fenômenos que ainda vão ocorrer.

Se a hipótese estiver errada, dependendo do grau de erro, ela deve ser melhorada, parcialmente, corrigida ou abandonada (trocada por outra hipótese).

6. Estabelecimento de uma Tese: Se a hipótese é comprovada pelos testes, ela se torna uma tese. (Uma tese é uma hipótese comprovada). A partir de teses também se cria modelos.

7. Criação de uma Teoria: Uma teoria é um conjunto de teses que explicam um mesmo fenômeno ou alguns fenômenos relacionados entre si e que já foi testada e comprovada em um grande número de experiências.

1.4.3 - MÉTODO EXPERIMENTAL

Denomina-se método experimental àquele em que **as variáveis são manipuladas de maneira preestabelecida e seus efeitos suficientemente controlados e conhecidos pelo pesquisador para observação do estudo**. O método experimental possibilita a demonstração dos dados coletados o que valida a pesquisa realizada. A coleta dos dados, no método experimental, é feita de forma a conduzir à respostas claras e diferenciadas em função de uma hipótese que envolve relações de causa e efeito.

A principal função deste método é a demonstrabilidade. No caso dos resultados apresentados pelo método experimental deve-se divulgá-los tal e qual se apresentam, mesmo que tenham ocorrido fatos imprevistos durante sua demonstração, e devem estar isentos de qualquer tipo de opinião pessoal.

O Método Experimental baseia-se no método científico e **apóia-se na observação e controle de três variáveis distintas.** Estas variáveis são os elementos que constituem a situação de estudo relativamente à alteração ou manifestação de um comportamento cuja natureza se desconhece.

1.4.4 - MÉTODO ESTATÍSTICO

O método estatístico é um processo para se obter, apresentar e analisar características ou valores numéricos para uma melhor tomada de decisão em situações de incerteza.

O desenvolvimento desse método envolve as seguintes fases:

1. Definição do problema.

A primeira fase de uma pesquisa estatística é a formulação de um problema de estudo. Além de considerar detidamente esse problema, o pesquisador deve também levantar os trabalhos já realizados nesse campo ou em campos análogos ao do seu problema, pois ele poderá ali encontrar informações pertinentes para sua pesquisa.

2. Formulação de um plano para a coleta de dados.

Essa fase consiste em determinar o procedimento necessário para resolver o problema e como levantar as informações sobre seu objeto de pesquisa. Que dados deverão ser obtidos? Como eles poderão ser obtidos? Essas são algumas perguntas pertinentes para se elaborar um plano de trabalho.

Uma das maiores preocupações que o pesquisador deve ter é na escolha e formulação das perguntas, independente da modalidade de coleta de dados.

É necessário também construir um cronograma de atividades, onde serão fixados os prazos para cada uma das diversas fases do estudo, os custos envolvidos, o exame das informações disponíveis, a delimitação da amostra e a forma como serão recolhidos os dados.

3. Coleta de dados.

A coleta de dados é a obtenção, reunião e registro sistemático das informações, com um determinado objetivo.

Há dois tipos de fontes de obtenção de dados: as que dão origem aos dados primários e as que dão origem aos dados secundários.

Chamam-se **dados primários** os que são publicados ou comunicados pela própria pessoa ou pela organização que os recolheu. Por exemplo, as tabelas do Censo Demográfico.

Os **dados secundários** são os publicados ou comunicados por outra organização. Por exemplo, estatística extraídas de várias fontes e publicadas em um jornal.

Sempre que possível deve-se trabalhar com fontes primárias, pois elas são mais fidedignas e, normalmente, trazem mais informações.

A coleta de dados pode ser direta ou indireta.

Coleta de dados direta: é obtida diretamente da fonte.

Coleta de dados indireta: é obtida através da inferência de dados obtidos a partir de uma coleta direta, ou através do conhecimento de outros fenômenos que, de algum modo, se relacionam com o fenômeno a ser estudado.

4. Apuração dos dados

Consiste em resumir os dados através de sua contagem e agrupamento. O objetivo dessa fase é a obtenção de um conjunto compacto de números o qual possibilita ao pesquisador uma melhor compreensão do comportamento do fenômeno na sua totalidade.

Esse processo tem como desvantagem a perda dos detalhes, já que se trata de uma sintetização.

5. Apresentação dos dados.

Há duas formas de apresentação dos dados, sendo que elas não são mutuamente excludentes: a apresentação tabular e a apresentação gráfica.

Apresentação Tabular: é a apresentação numérica dos dados, distribuídos em linhas e colunas segundo algumas regras práticas adotadas pelos diversos sistemas estatísticos. Essas regras foram fixadas pelo Conselho Nacional de Estatísticas e podem ser encontradas em publicações obtidas em qualquer agência do IBGE.

Apresentação gráfica: é uma apresentação geométrica (gráficos), que traz a vantagem de proporcionar uma visualização rápida, fácil e clara do fenômeno.

Existem diversos tipos de gráficos e cada um deles será tratado mais detalhadamente em outro tópico dessa apostila.

6. Análise e interpretação dos dados.

O objetivo dessa fase é possibilitar ao pesquisador tirar conclusões que o ajudem a resolver o problema. A análise dos dados está ligada essencialmente ao cálculo de medidas. Resumindo, análise dos dados pode ser expressa por números-resumo que expressam as características particulares do fenômeno.

1.5 - POPULAÇÃO OU UNIVERSO ESTATÍSTICO

É o conjunto da totalidade dos indivíduos sobre os quais se faz uma inferência.

Exemplo: Em uma pesquisa sobre a intenção de votos para governador de um estado, a população seria o conjunto de todos os eleitores desse estado.

1.6 - AMOSTRA

É um subconjunto da população, ou seja, é uma parte selecionada da população através da qual se faz uma inferência sobre as características da mesma.

Exemplo: Quer-se analisar a qualidade de uma carga de sacos de arroz a ser exportada. Para isso analisa-se o arroz de alguns sacos (amostra) para inferir qual a qualidade da carga toda.

1.7 - VARIÁVEL

É aquilo que está sendo pesquisado na amostra ou na população.

As variáveis podem ser **qualitativa** – o que se pesquisa é um atributo como, por exemplo, a cor do cabelo – ou **quantitativa**, por exemplo, o número de filhos de uma determinada amostra.

Tipos de variáveis

	Qualitativa Apresentam como possíveis valores uma qualidade ou atributo. Por exemplo, cor do cabelo, esporte favorito, etc.	**Ordinal** Existe uma ordem nos seus valores. Por exemplo, grau de instrução (fundamental, médio, superior, etc.)
		Nominal Não existe uma ordem nos seus possíveis valores. Por exemplo, “esporte favorito”.

Variável	**Quantitativa** Os possíveis valores da variável são números. Por exemplo, idade, número de irmãos, etc.	**Discreta ou Descontínua** Quando se trata de contagem (números inteiros). Por exemplo, números de irmãos (0, 1, 2, ...)
		Contínua Quando se trata de medida (números reais). Por exemplo, “altura”.

2 - ARREDONDAMENTO DE DADOS

Arredondamento por falta: quando o dígito situado mais à esquerda entre os que irão ser eliminados for igual ou menor que 4, não deve ser alterado o dígito remanescentes.

Exemplos: Arredondamento por falta para décimos do número 2,7**35** = 2,7

Arredondamento por falta para inteiros do número 5, **432** = 5

Arredondamento por falta para centésimos do número 1, 32**41** = 1,32.

Arredondamento por excesso: quando o primeiro dígito após aquele que será arredondado for maior ou igual a 5 seguido por dígitos maiores que zero, o dígito remanescente será acrescido de uma unidade.

Exemplos: Arredondamento por excesso para inteiros do número 3, **5483** = 4

Arredondamento por excesso para décimos do número 2, 13**76** = 2,14

Arredondamento de dígitos seguidos do 5: quando o dígito mais à esquerda dos que serão eliminados for cinco ou cinco seguido somente de zeros, o último dígito remanescente **não se alterará se ele for par** e será **acrescido de uma unidade se for ímpar.**

Exemplos: Arredondamento para décimos do número 2, 3**5** = 2,4

Arredondamento para centésimos do número 1, 54**500** = 1,54

Arredondamento de soma: quando se tem uma soma, arredonda-se primeiro o total e depois as parcelas.

a) Se a soma das parcelas da série arredondada for superior ao total, deve-se retornar à série inicial, arredondando-se, por falta, tantas parcelas quantas forem as unidades excedentes.

Exemplo:

Série original	Série arredondada	Série corrigida
6,51	7	7
7,50	8	8
14,63	15	15
20,10	20	20
24,73	25	24 (parcela corrigida)
26,52	27	26 (parcela corrigida)
99,99	102>100	100

b) Se a soma das parcelas da série arredondada for inferior ao total retorna-se à série original, arredondando-se por excesso tantas parcelas quantas forem necessárias.

Exemplo

Série original	Série arredondada	Série corrigida
5,34	5	5
7,45	7	7
18,50	18	19 (parcela corrigida)
19,90	20	20
22,37	22	22
26,43	26	27(parcela corrigida)
99,99	98 <100	100

3 - FREQUÊNCIA

Frequência é o número de observações ou repetições de um valor ou modalidade.

III – 1 TIPOS DE FREQUÊNCIAS

a. Freqüência simples absoluta (f_j) – é o número de vezes que um valor da variável é citado.

Exemplo

O resultado de uma pesquisa sobre a nacionalidade de um grupo de turistas foi: Roberto, brasileiro, Emília, brasileira, Carlos, espanhol, Juan, espanhol, Luiz, brasileiro, Eduardo, brasileiro, Marisa, brasileira, Lídia, espanhola, Marcela, brasileira e Manolo, argentino.

Assim, a freqüência simples absoluta da nacionalidade brasileira é 6, da espanhola é 3 e da argentina é 1.

b. Freqüência simples relativa (fr_j) – é a freqüência absoluta em relação ao total de citações. Ela pode ser expressa em termos de fração, decimal ou porcentagem. Assim, no exemplo anterior tem-se:

Exemplo

Distribuição de viajantes, segundo a nacionalidade

Nacionalidade	Freqüência simples absoluta (f_j)	Freqüência simples relativa (fr_j)

Brasileira	6	60%
Espanhola	3	30%
Argentina	1	10%
Total	10	100%

c. Freqüência simples acumulada "Abaixo de" (F_j) - a freqüência acumulada "abaixo de" de uma classe ou de um valor individual é a soma das freqüências absolutas dessa classe ou desse valor com as freqüências absolutas das classes ou dos valores anteriores.

Exemplos

Distribuição de peças com defeito, segundo os lotes analisados

Número do lote	Freqüência Simples (f_j)	Freqüência acumulada (F_j)
1	5	5
2	10	15
4	12	27
5	15	42
6	8	50

Distribuição de indivíduos de uma vida, segundo a idade

Classes	Freqüência Simples (f_j)	Freqüência acumulada(F_j)
1 ├ 5	250	250
5 ├ 10	150	400
10 ├ 15	120	520
15 ├ 20	130	650
20 ├ 25	350	1000

d. Frequência Simples Acumulada "Acima de"(F_j) – a freqüência absoluta acumulada "acima de" uma classe ou de um valor individual é o número obtido através da adição da freqüência simples absoluta da classe ou do valor com as freqüências simples das classes ou valores individuais posteriores

Exemplos

Distribuição de peças com defeito, segundo os lotes analisados

Número do lote	Freqüência Simples (f_j)	Freqüência acumulada (F_j)
1	5	50
2	10	45
4	12	35
5	15	23
6	8	8

Distribuição de indivíduos de uma vida, segundo a idade

Classes	Freqüência Simples (f_j)	Freqüência acumulada(F_j)
1 ├ 5	250	1000
5 ├ 10	150	750
10 ├ 15	120	500
15 ├ 20	130	480
20 ├ 25	350	350

e. Frequência relativa acumulada "Abaixo de"(Fr_j) de uma classe ou valor individual é o número obtido pela adição da freqüência relativa da classe ou valor individual com as freqüências relativas das classes ou valores individuais anteriores.

Exemplos

Distribuição de peças com defeito, segundo os lotes analisados

Número do lote	Freqüência Relativa Simples (F_j)	Freqüência acumulada (Fr_j)
1	10	10
2	20	30
4	24	54
5	30	84
6	16	100

Distribuição de indivíduos de uma vida, segundo a idade

Classes	Freqüência Relativa Simples (fr_j) (%)	Freqüência acumulada (F_j) (%)

1 ├ 5	25	25
5 ├ 10	15	40
10 ├ 15	12	52
15 ├ 20	13	65
20 ├ 25	35	100

f. Frequência relativa acumulada "Acima de" (Fr_j) de uma classe ou valor individual é o número obtido pela adição da freqüência relativa da classe ou valor individual com as freqüências relativas das classes ou valores individuais posteriores.

Exemplos

Distribuição de peças com defeito, segundo os lotes analisados

Número do lote	Freqüência Relativa Simples (F_j)	Freqüência acumulada (Fr_j)
1	10	100
2	20	90
4	24	70
5	30	46
6	16	16

Distribuição de indivíduos de uma vida, segundo a idade

Classes	Freqüência Relativa Simples (F_j) (%)	Freqüência acumulada (F_j) (%)
1 ├ 5	25	100
5 ├ 10	15	75
10 ├ 15	12	60
15 ├ 20	13	48
20 ├ 25	35	35

4 - SÉRIES ESTATÍSTICAS

Não é conveniente apresentar os dados da forma como foram coletados, quando se faz uma pesquisa, pois muitas vezes o conjunto de valores é extenso e desorganizado, dificultando o entendimento do fenômeno.

Sendo assim é necessário utilizar-se das séries estatísticas. **Uma série estatística é uma sucessão de dados estatísticos referidos a caracteres qualitativos** e **uma seriação é uma sucessão de dados estatísticos referentes a caracteres quantitativos**.

a. Série Temporal – o fator que varia é um fator cronológico.

Exemplo

Distribuição de carros fabricados por uma montadora no primeiro semestre de um determinado ano

Meses	Número de carros fabricados
Janeiro	23 000
Fevereiro	18 000
Março	22 000
Abril	22 100
Maio	23 600
Junho	26 000

b. Série Geográfica – o fator que varia é um fator geográfico.

Exemplo

Distribuição de carros produzidos por uma montadora, segundo os estados

Estado	Número de carros
Minas Gerais	40 000
Paraná	22 000
Rio de Janeiro	42 000
São Paulo	58 000

c. Série Específica – o que varia é o fenômeno

Exemplo

Distribuição de vendas de carro de uma montadora, conforme o modelo

Modelo	Número de carros vendidos
A	64 500
B	93 100
C	15 750
Total	159 170

d. Distribuição de freqüências – Seriação – é uma série na qual o fenômeno apresenta gradações ou subdivisões, isto é, os dados são reunidos de acordo com sua magnitude.

Exemplo

Distribuição dos empregados de uma empresa, segundo o salário recebido.

Valor do salário	Número de empregados
Até 1 salário mínimo	3 000
De 1 a 2 salários mínimos	4 500
De 2 a 3 salários mínimos	2 500
Acima de 3 salários mínimos	4 000
Total	14 000

5 - TABELAS

Tabelas de freqüências são representações nas quais os valores se apresentam relacionadas às suas repetições, evitando assim que um mesmo valor apareça mais de uma vez.

Os tipos de tabelas podem ser

a. Tabela de distribuição de freqüências de dados não agrupados em classe – os valores da variável se apresentam individualmente. Esse tipo de tabela é usado quando se trabalha com variável discreta ou descontínua.

Exemplo

Distribuição das famílias de uma cidade conforme o número de filhos

Número de filhos	Número de famílias (Freqüência simples ou absoluta)
1	8 000
2	15 000
3	12 000
4	5 000
Mais de 4 filhos	3 000
Total	43 000

b. Tabela de distribuição de freqüência de dados agrupados em classe – os valores observados são agrupados em classes. Esse tipo de tabela é mais usado quando a variável for contínua ou quando ela for discreta, mas o número de dados é muito grande.

Exemplo

Distribuição dos alunos de uma série segundo a nota obtida em um teste de matemática

Classes	Freqüência
0 ├ 10	15
10 ├ 20	25
20 ├ 30	40
30 ├ 40	30
40 ├ 50	10
Total	120

Observações:

a) Se só havia notas inteiras (0, 10, 20 …) essa é uma tabela de dados agrupados em classes de uma variável discreta, mas se havia notas "picadas" (11; 22; 36; …) então a variável é contínua.

b) O sinal ├ significa que o valor a sua esquerda (limite inferior) pertence àquela classe e o valor a sua direita (limite superior) não pertence à classe. Assim, na classe 20 ├ 30 estão agrupados todos os alunos que obtiveram nota maior ou igual a 20 e menor que 30.

c. Como construir uma tabela de dados agrupados em classes

1º) Determina-se a amplitude total ou o intervalo total, que é a diferença entre o maior e o menor valor observado. No exemplo anterior, considerando que a variável é contínua, se a maior nota fosse 48 e a menor 2, teríamos que a amplitude total seria 48 – 2 = 46

2º) Determina-se o número de classes. Para isso há diversos métodos, entre eles a regra de Sturges, que estabelece que o número de classes (k) é dado por

k = 1 + 3,3 $\log n$, onde n é o número total de observações.

Assim, se n = 120, como no exemplo dado, o número de classes seria

k = 1 + 3,3 log 120. Como log 120 ~ 1,30103

k = 1 + 3,3 . 1,30103 = 1 + 4,29340 = 5,29340. Arredondando teremos 5 classes.

Para evitar um número muito pequeno ou muito grande de classes, Truman L. Kelley sugere a seguinte tabela

N	5	10	25	50	100	200	500	1 000
k	2	4	6	8	10	12	15	15

3º) Determina-se as classes e suas respectivas freqüências e constrói-se a tabela.

d. Limites reais de classe – são as médias aritméticas entre o limite superior de uma classe e o inferior da classe seguinte. Assim, no exemplo anterior, considerando as notas como variável contínua, o limite real da classe 10 ├ 20 é $\dfrac{19+20}{2} = 19{,}5$

e. Ponto médio da classe (x_j) – é o valor obtido quando se adiciona ao limite inferior a metade da amplitude da classe. Assim teremos no exemplo da distribuição de notas a seguinte tabela

Distribuição dos alunos de uma série segundo a nota obtida em um teste de matemática

Classes	Freqüência	Valores médios (x_i)
0 ├ 10	15	5
10 ├ 20	25	15
20 ├ 30	40	25
30 ├ 40	30	35
40 ├ 50	10	45

Os valores médios são importantes no cálculo de algumas medidas como, por exemplo, a média aritmética de valores agrupados em classes.

f. Tabela de dupla entrada - é uma tabela na qual se apresentam mais de uma série conjugadas.

Exemplo

Distribuição de nascimentos semanal de uma maternidade, segundo o sexo

Dias da Semana	Sexo	
	Feminino	Masculino
Domingo	5	8
Segunda	4	4
Terça	6	3
Quarta	2	7
Quinta	3	1
Sexta	1	8
Sábado	6	5
Total	27	36

Cada linha, encabeçada pela frase “Dias da semana” – domingo, segunda, terça, quarta, quinta, sexta, sábado – representa uma série temporal. Cada coluna encabeçada por Feminino, Masculino – representa uma série específica. Tem-se, portanto uma série específico-temporal.

6 - Exercício Resolvido

Em uma escola com 5 classes de 1ª série do ensino médio, cada uma com 45 alunos, foi feita uma pesquisa para traçar o perfil da 1ª série. Para tanto, foram selecionados 5 alunos de cada classe, que responderam a um questionário do qual foi elaborada a seguinte tabela:

Distribuição dos alunos da 1ª série, segundo características pessoais e desempenho em Matemática

Nome	Sexo	Idade (anos e meses)	Altura (cm)	Peso (kg)	Nº de irmãos	Cor do cabelo	Hobby	Nº do sapato	Manequim	Desempenho em Mat
Antônio	M	15a 4m	156	49	2	Cast	Esporte	36	38	Ótimo
Artur	M	14a 7m	166	48	0	Cast	Esporte	39	38	Bom
Áurea	F	15a 2m	165	66	1	Cast	Música	36	42	Insuficiente
Bruno	M	14a 8m	175	63	0	Cast	Patinação	40	42	Regular
Carla	F	14a 5m	165	57	2	Loiro	Música	36	40	Regular
Cláudia	F	15a 3m	164	50	2	Loiro	Dança	36	38	Bom
Domingos	M	14a 6m	163	51	1	Cast	Esporte	36	38	Bom
Edite	F	14a 7m	160	60	3	Cast	Música	36	40	Ótimo
Flávia	F	14a 7m	175	65	1	Cast	Esporte	37	42	Bom
Fúlvio	M	14a 5m	150	38	1	Ruivo	Esporte	34	36	Insuficiente
Geraldo	M	15a 11m	146	38	0	Cast	Aeromodelismo	34	36	Regular
José	M	14a 10m	165	52	1	Cast	Dança	38	38	Regular
Laura	F	14a 0m	165	53	2	Cast	Dança	36	38	Bom
Lúcia	F	14a 8m	167	65	2	Cast	Música	37	42	Bom
Mário	M	15a 4m	165	50	3	Loiro	Patinação	36	38	Insuficiente
Mauro	M	14a 11m	163	54	4	Cast	Esporte	38	40	Ótimo

Nome	Sexo	Idade (anos e meses)	Altura (cm)	Peso (kg)	Nº de irmãos	Cor do cabelo	Hobby	Nº do sapato	Mane quim	Desempenho em Mat
Nívea	F	15a2m	164	63	1	Loiro	Esporte	38	42	Bom
Orlando	M	14a 8m	159	64	2	Cast	Música	37	42	Regular
Patrícia	F	15a 1m	158	43	1	Loiro	Dança	36	36	Insuficiente
Paula	F	14a 11m	163	53	1	Cast	Dança	36	38	Bom
Renata	F	14a 3m	162	52	1	Cast	Dança	36	38	Ótimo
Roberto	M	14a 2m	167	53	0	Cast	Esporte	40	38	Ótimo
Sandra	F	14a 10m	167	58	1	Loiro	Dança	40	40	Ótimo
Teresa	F	15a 9m	155	49	0	Cast	Patina ção	35	36	Ótimo
Vânia	F	15a 2m	152	41	3	Cast	Música	34	36	Bom

Em relação à tabela anterior, responda:

a) Qual é o universo estatístico?
b) Qual o tamanho da amostra?
c) Cite uma variável qualitativa nominal
d) Cite uma variável quantitativa discreta.
e) Cite uma variável qualitativa ordinal.
f) Cite uma variável quantitativa contínua.
g) Que valor da variável Hobby tem freqüência absoluta igual a 7 e freqüência relativa igual a 28%?
h) Qual a freqüência absoluta e relativa (em %) do valor 38 da variável "manequim"?
i) Elabora a tabela de freqüência absoluta simples da variável "Desempenho em Matemática".
j) Construa a tabela de freqüência relativa simples (F_j) variável "idade" agrupada em classes.
k) Construa uma tabela de freqüência absoluta acumulada (fr_j) "acima de" da variável "altura"
l) Construa uma tabela de freqüência relativa acumulada "abaixo de"(Fr_j) da variável "Desempenho em Matemática".

Solução

a) O universo é o total de alunos da 1ª série, isto é, 225 alunos.
b) A amostra é a quantidade de pessoas pesquisadas, isto é, 25 alunos.

c) Cor de cabelo ou Hobby.
d) Altura, Peso, Nº de irmãos, Nº de sapato e Manequim.
e) Desempenho em Matemática
f) Idade

Construindo a tabela de freqüências simples e relativa da variável "Hobby" temos

Distribuição dos alunos do 1º ano, segundo o Hobby

Hobby	f_j	F_j (%)
Aeromodelismo	1	4
Dança	7	28
Esporte	8	32
Música	6	24
Patinação	3	12
Total	25	100

Logo, Dança é o hobby que tem freqüência simples absoluta 7 e relativa 28%

g) f_j = 10, pois esse hobby aparece 10 vezes na distribuição de valores.

F_j = 10 : 25 = 0,4 = 40%

h) Distribuição de alunos do 1º ano, segundo o desempenho em Matemática

Desempenho em Matemática	f_j
Insuficiente	4
Regular	5
Bom	9
Ótimo	7

i) Usando a tabela de Kelly, como se tem 24 dados (n = 25) é conveniente agrupá-los em 6 classes.

Amplitude total: maior idade: 15 anos e 11 meses – menor idade: 14 anos e 0 meses = 21 meses

Amplitude de cada classe: 21 : 6 = 3,5 ~ 4

Distribuição dos alunos do 1º ano, segundo a altura

Classes	F_j
14a ├ 14a e 4m	12
14a e 4m ├ 14a e 8m	24
14a e 8m ├ 15 a	28
15 a ├ 15 a e 4m	20
15 a e 4m ├ 15 a e 8m	8
15 a e 8m ├ 16 a	8
Total	100

j)

Distribuição dos alunos do 1º ano, segundo

Classes	f_j	fr_j
14a ├ 14a e 4m	3	25
14a e 4m ├ 14a e 8m	6	22
14a e 8m ├ 15 a	7	16
15 a ├ 15 a e 4m	5	9
15 a e 4m ├ 15 a e 8m	2	4
15 a e 8m ├ 16 a	2	2
Total	100	

k)

Distribuição dos alunos do 1º ano, segundo o peso

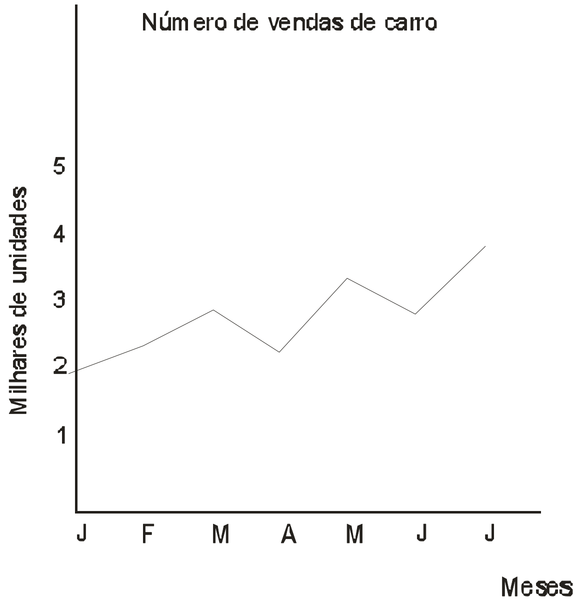
Peso	F_j	Fr_j
Insuficiente	16	16
Regular	20	36
Bom	36	72
Ótimo	28	100
Total	100	

7 – GRÁFICOS

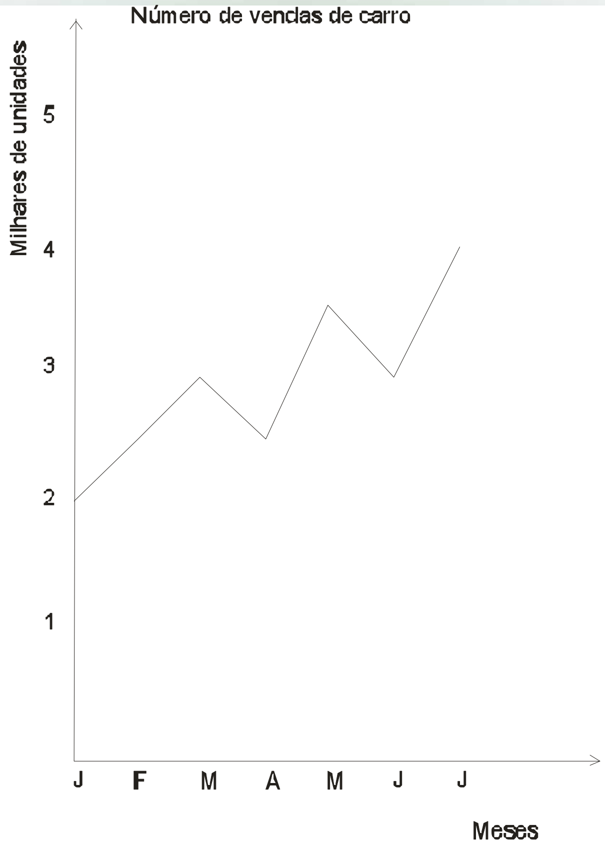
O gráfico estatístico é uma representação geométrica dos dados estatísticos que tem por objetivo fornecer ao investigador e ao público, uma visão clara e rápida do fenômeno. Entretanto, há que se ter cuidado ao construir um gráfico pois, caso contrário, ele pode confundir o leitor.

Exemplo: Observe os gráficos a seguir.

a)



b)



Embora os dois gráficos se refiram aos mesmos dados visualmente eles se diferem, isto porque foram tomadas unidades diferentes nos dois eixos. A impressão que se tem é que no gráfico b) a variação do número de carros vendidos foi maior.

7.1 - TIPOS DE GRÁFICOS

a. Diagramas – são gráficos geométricos bidimensionais. São os mais usados nas representações de séries estatísticas.

b. Cartogramas – são ilustrações relativas a cartas geográficas

c. Estereogramas – representam volumes e são tridimensionais. Ou são desenhados em perspectiva ou usando-se cartolina ou madeira.

7.2 - CLASSIFICASSÃO DOS GRÁFICOS SEGUNDO O OBJETIVO

Os gráficos podem ser classificados segundo o seu uso. Tem-se dois tipos de gráficos, considerando-se seu objetivo:

a. **Gráfico de informação** – seu objetivo é fornecer ao grande público uma visualização rápida e clara do maior número de informações possíveis sobre o fenômeno. É imprescindível o título, mas a legenda pode ser omitida desde que sejam visíveis as informações desejadas (ao longo dos eixos)

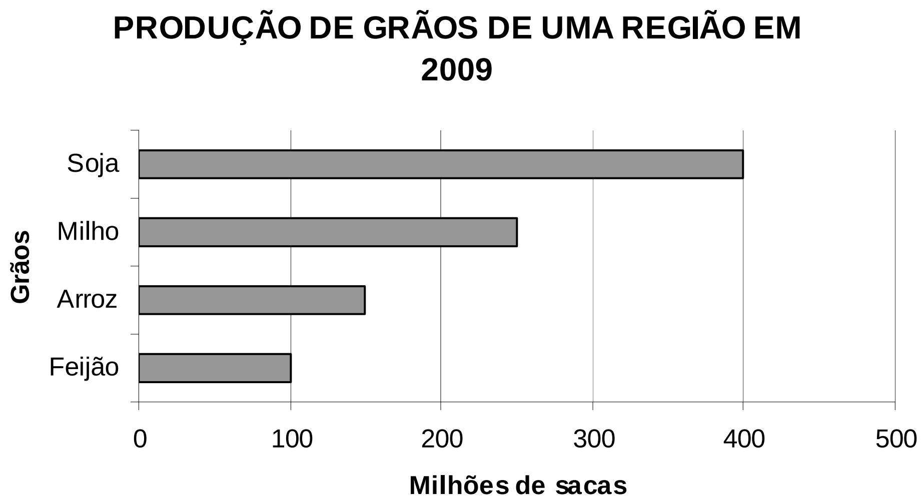
b. **Gráficos de análise** – são usados na fase de análise dos dados de um trabalho estatístico, embora também contenham informações. Normalmente vêm acompanhados de um texto explicativo e/ou de tabela

7. 3 - TIPOS DE GRÁFICOS DE INFORMAÇÃO

7.3.1 - GRÁFICOS DE BARRAS

Servem para comparar grandezas através da observação dos retângulos que o compõem, sendo que estes devem ter a mesma largura e alturas proporcionais às respectivas grandezas.

Exemplo



Orientações para se construir um gráfico de barras

- as barras têm a mesma largura e diferem no comprimento.

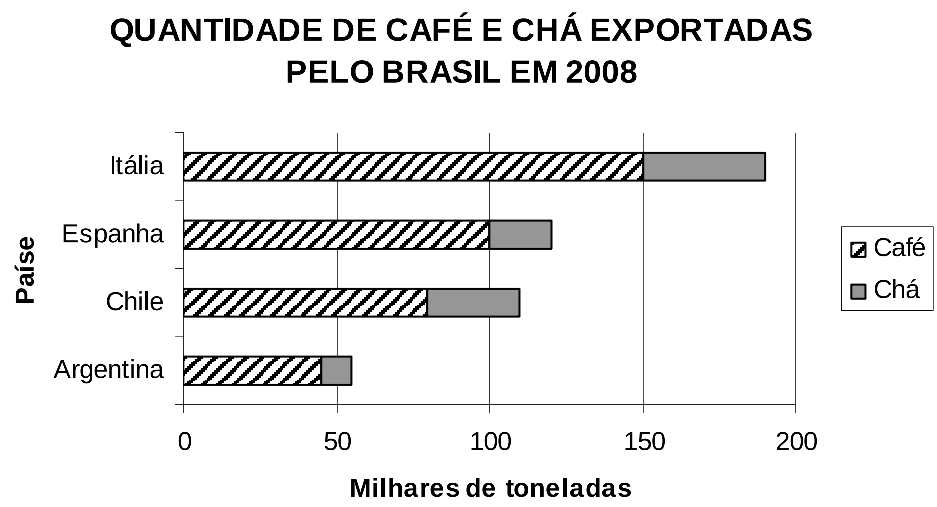
- o espaço entre as barras deve ser o mesmo e suficiente para as inscrições que as identificam não confundir o leitor. (Usualmente toma-se o espaço entre as barras como sendo aproximadamente igual à metade ou a dois terço de suas larguras)

- as barras devem ser desenhadas observando sua ordem de grandeza para facilitar a comparação entre os dados.

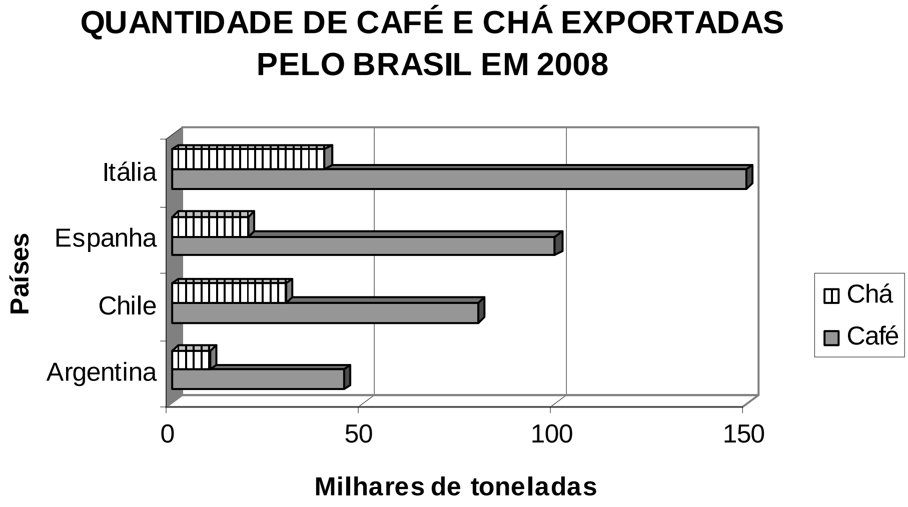
- o gráfico, construído para apresentar grandezas absolutas, deve apresentar a linha zero bem definida e uma escala de quantidades ininterrupta.

Outros tipos de gráfico de barras

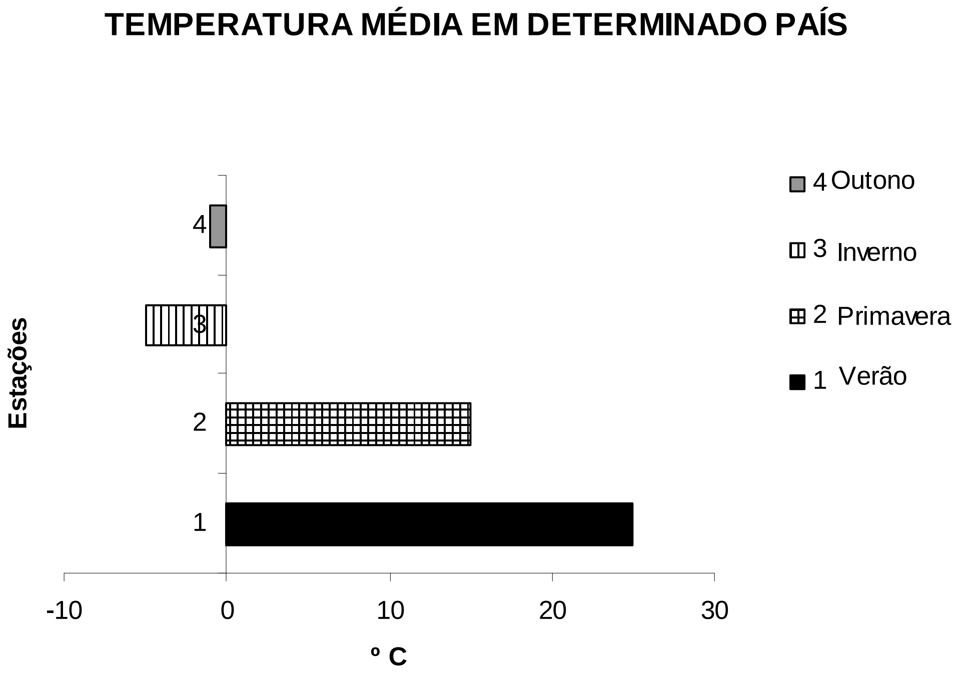
a. Gráfico de barras compostas



b. Gráfico de barras compostas



c. Gráfico de barras bidirecionais



7.3.2 - GRÁFICO DE COLUNAS

Embora tenham a mesma finalidade que os gráficos de barras, os gráficos de colunas são mais aconselháveis quando as legendas a serem escritas sob os retângulos são menores.

Exemplo

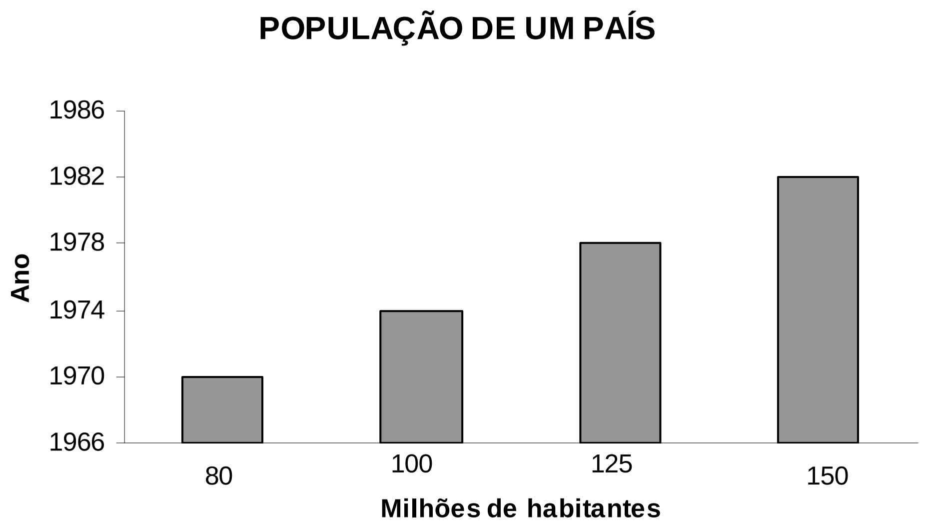
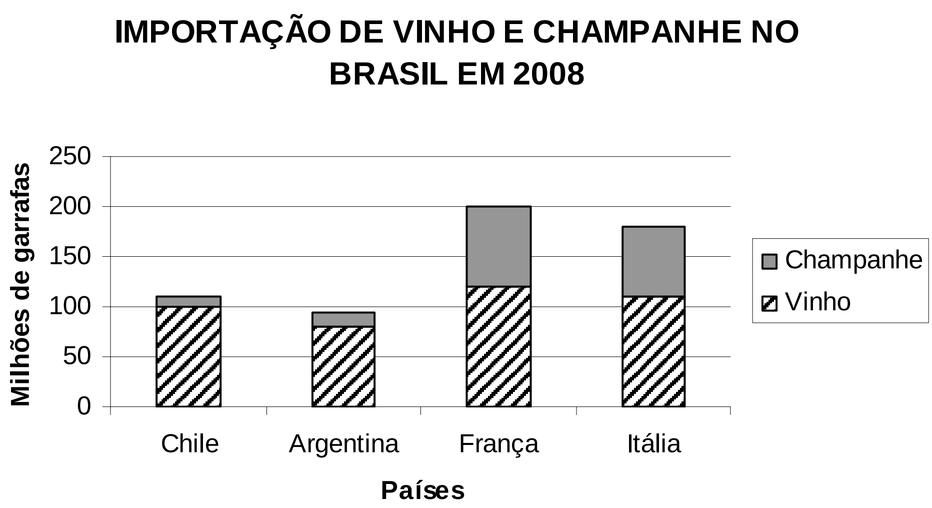


Gráfico de colunas superpostas – corresponde ao gráfico de barras compostas.

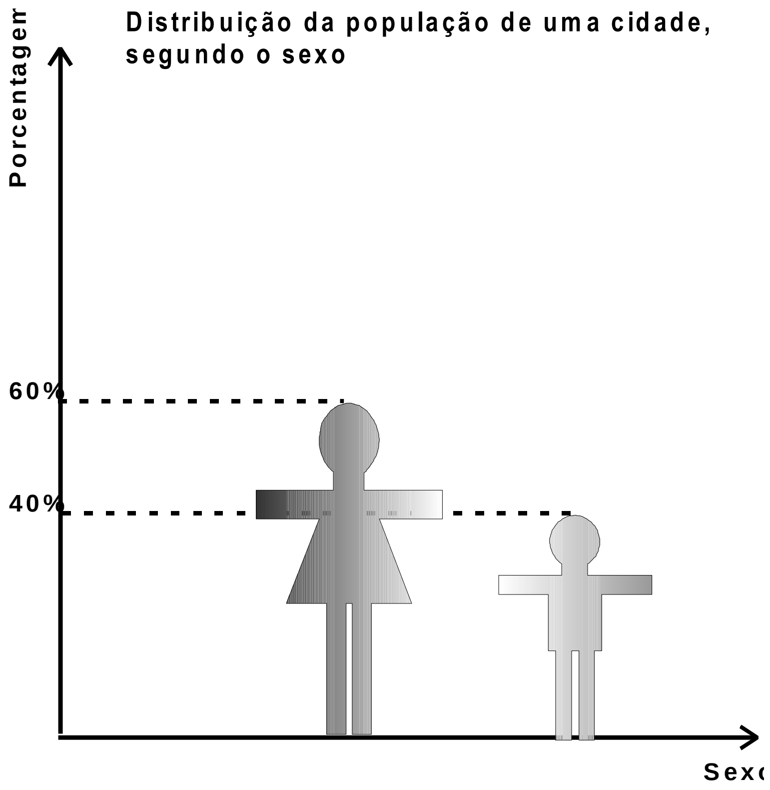
Exemplo



7.3.3 - PICTOGRAMAS

Os pictogramas são gráficos que usam conjuntos de figuras representativas da intensidade ou das modalidades do fenômeno.

Exemplos

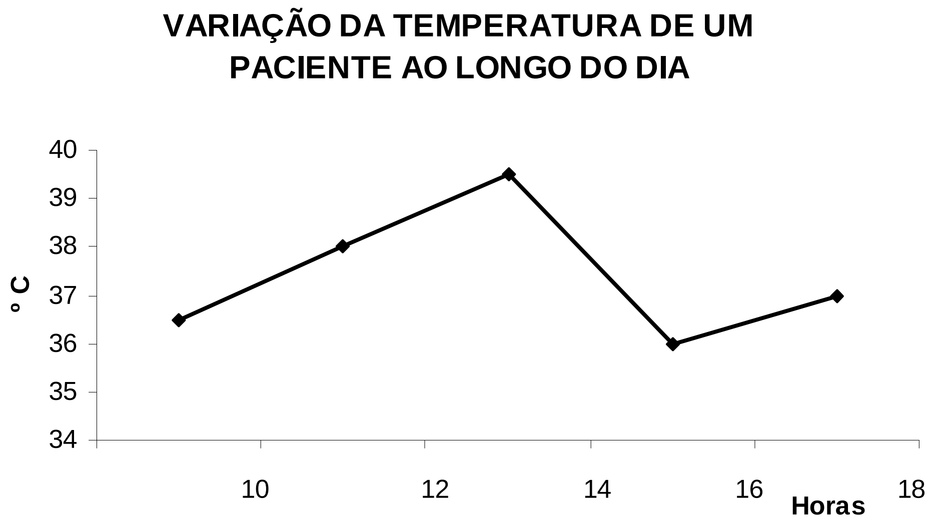

Distribuição da população de uma cidade, segundo o sexo
Porcentagem
60%
40%
Sexo

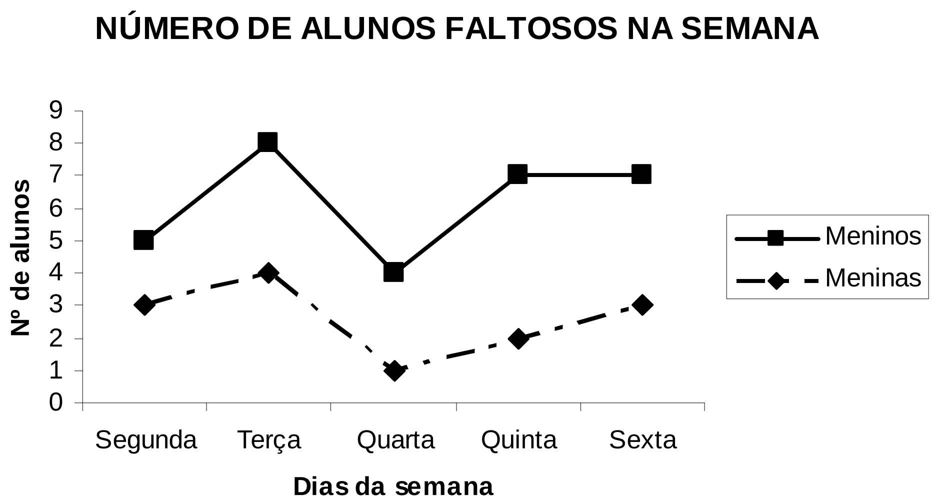

NÚMERO DE CARROS VENDIDOS POR DÉCADA
2000
1990
1980
1970
= 1 milhão de carros

7.3.4 - GRÁFICOS DE LINHAS OU GRÁFICOS LINEARES

São usados quando existem intensas flutuações nas séries ou quando se quer representar várias séries em um mesmo gráfico. Para construí-los basta marcar os pontos correspondentes aos valores observados e uni-los por uma linha contínua.

Exemplos:



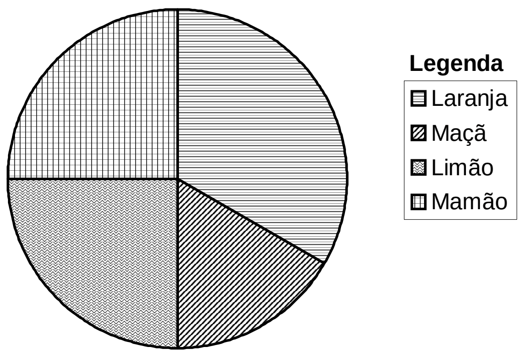


7.3.5 - GRÁFICOS EM SETORES OU CARTOGRAMAS EM SETORES

São usados para representar valores absolutos ou porcentagens.

Exemplo:

Produção de frutas de uma certa cidade



Fonte: Dados hipotéticos

Como construir o gráfico em setores.

Considerando-se a circunferência toda como um setor cujo o ângulo central é 360º, para determinar o ângulo central de um setor correspondente a um determinado valor basta fazer uma regra de três simples. Depois, usando um transferidor, marca-se esse ângulo central. Fazendo-se o mesmo para todos os valores, constrói-se o gráfico

Por exemplo, considere um problema simples: “Em uma Escola o estudo de uma língua estrangeira é obrigatório, mas os alunos podem optar entre Inglês, Francês ou Espanhol. Entretanto cada aluno só pode optar por uma delas. Em uma classe de 40 alunos, 20 optaram por estudar Inglês, 15 por Espanhol e o restante por Francês".

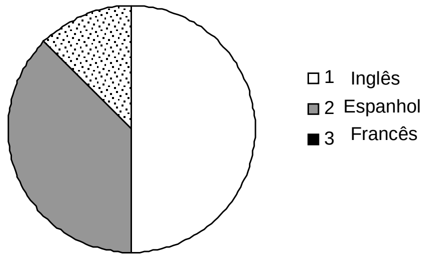
Construa um gráfico de setores que represente esta situação.

$40 \rightarrow 360º$ $\qquad$ $40 \rightarrow 360º$ $\qquad$ $40 \rightarrow 360º$

$20 \rightarrow x$ $\qquad$ $15 \rightarrow x$ $\qquad$ $5 \rightarrow x$

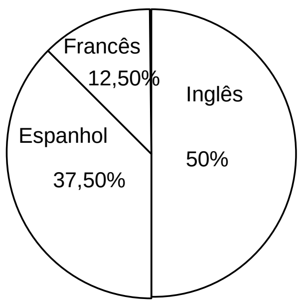
$$x = \frac{360x20}{40} = 180\,º \qquad x = \frac{360x15}{40} = 135^{o} \qquad x = \frac{360x5}{40} = 45^{o}$$

DISTRIBUIÇÃO DE ALUNOS DE ACORDO COM A LÍNGUA ESTRANGEIRA ESTUDADA



A legenda pode ser dispensada, quando se escreve o nome de cada variável e sua porcentagem no interior de cada setor

DISTRIBUIÇÃO DE ALUNOS DE ACORDO COM A LÍNGUA ESTRANGEIRA ESTUDADA

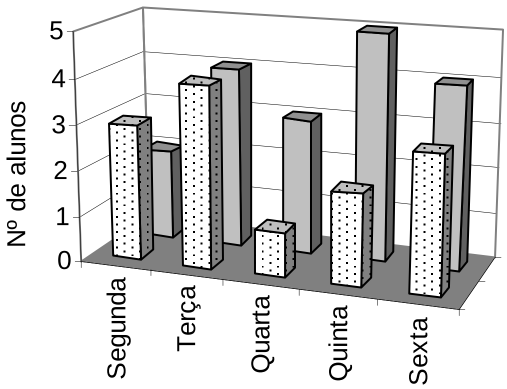


7.3.6 - ESTEREOGRAMAS

São usados para a representação gráfica de tabelas de dupla entrada. Esse tipo de gráfico dificulta ao leitor verificar facilmente as variações do fenômeno.

Exemplo:

NÙMERO DE ALUNOS FALTOSOS NA SEMANA



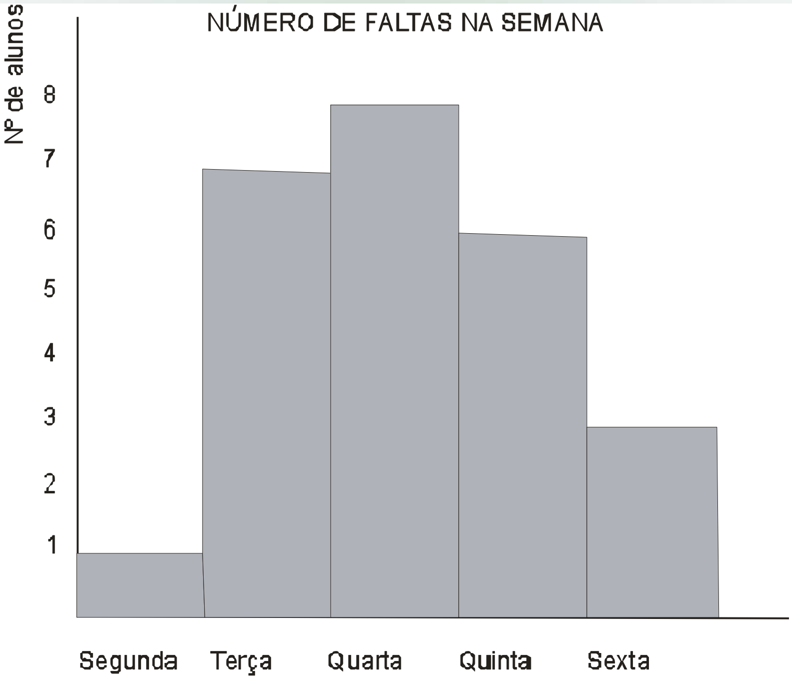
7.3.7 - REPRESENTAÇÃO GRÁFICA DE UMA DISTRIBUIÇÃO DE FREQUÊNCIAS

As distribuições de freqüências são representadas por gráficos de análise. Normalmente se usa o histograma ou polígono de freqüência para representar as freqüências simples e o polígono de frequências ou ogivas para representar as freqüências acumuladas.

Histogramas – são gráficos formados por um conjunto de retângulos agrupados, de forma que a área de cada retângulo seja proporcional à freqüência da classe que ele representa. A soma dos valores das áreas dos retângulos corresponde à freqüência total.

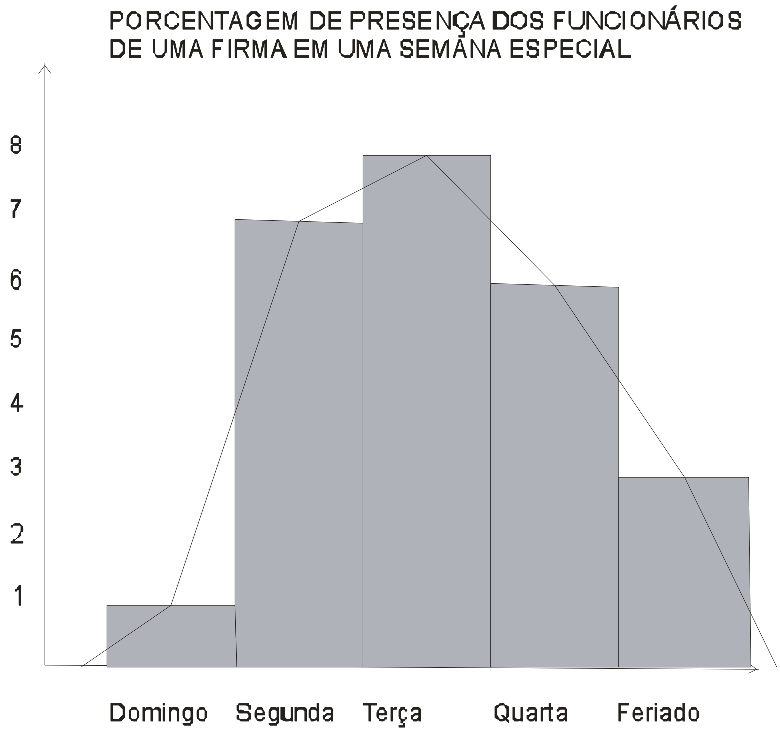
Na construção de um histograma, os valores individuais da variável ou os limites das classes são colocados no eixo horizontal. No eixo vertical coloca-se o número de observações ou freqüência da classe.

Exemplos



Polígono de frequências – é a representação dos dados obtida quando se une os pontos médios das bases superiores dos retângulos de um histograma. Assim como o próprio histograma, o polígono de frequências pode se referir a frequências absolutas ou relativas.

Exemplo



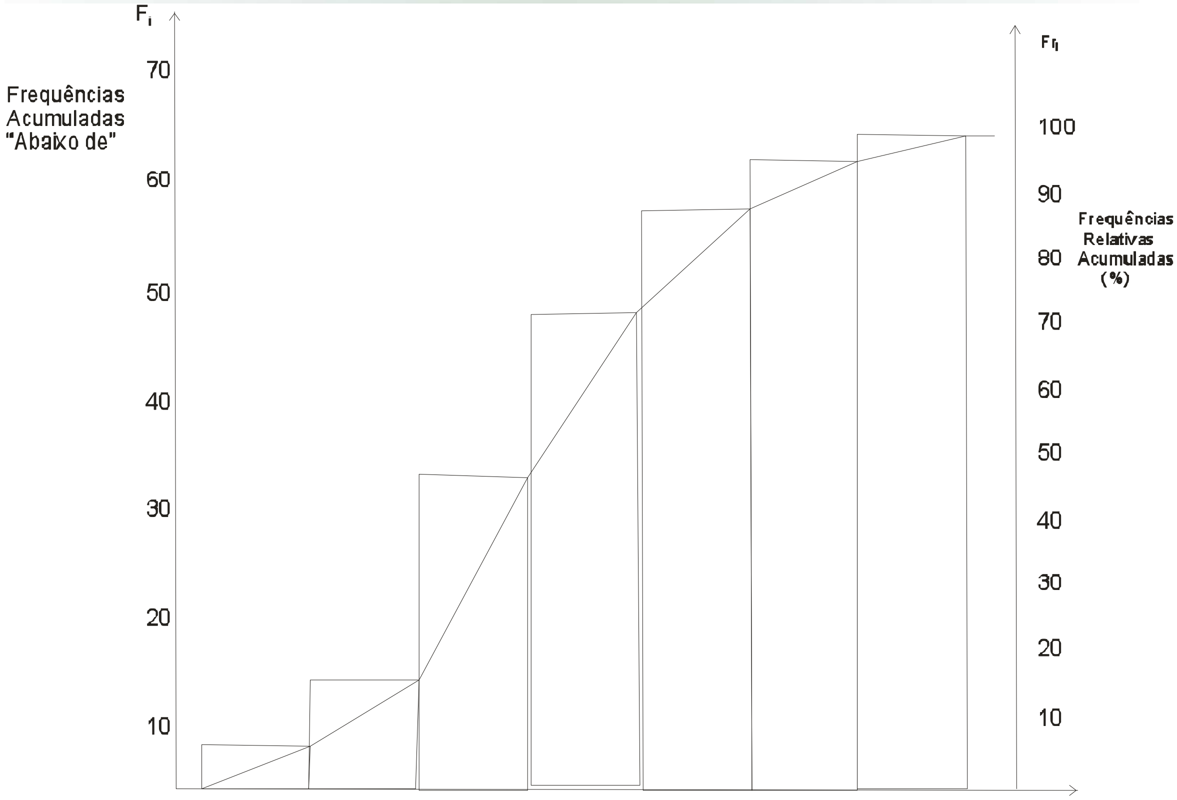
Polígono de frequências acumuladas – é a representação gráfica das frequências acumuladas. Quando o polígono de freqüência acumuladas se refere às frequências relativas ele é chamado de **ogiva percentual**.

Exemplos

Considere a tabela.

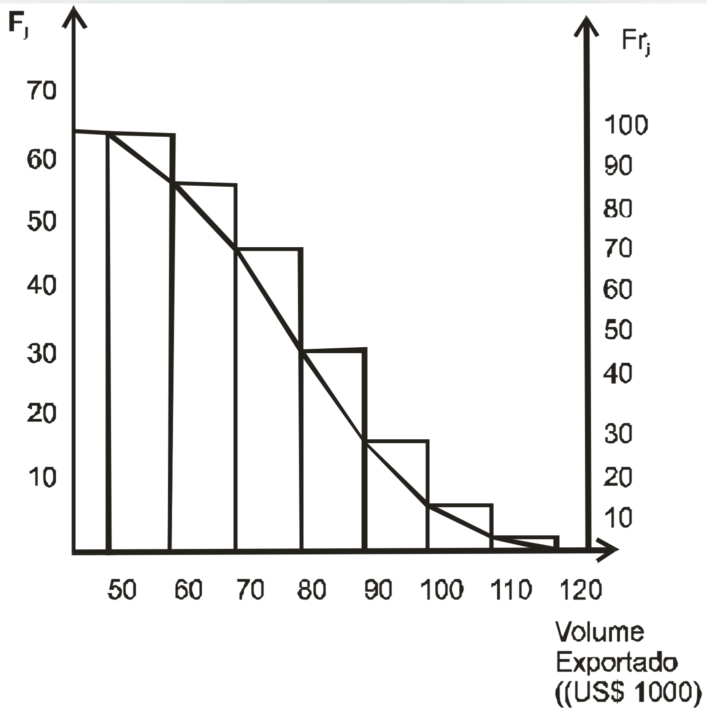
Volume exportado US$	Nº de empresas	F_j "Abaixo de"	F_j "Acima de"
50 000 ├ 60 000	8	8	65
60 000 ├ 70 000	10	18	57
70 000 ├ 80 000	16	34	47
80 000 ├ 90 000	14	48	31
90 000 ├ 100 000	10	58	17
100 000 ├ 110 000	5	63	7
110 000 ├ 120 000	2	65	2

Gráfico de Frequências Acumuladas "Abaixo de"



Suponhamos que se deseja saber qual o número de empresas com um volume de exportação até 80 000 dólares. Olhando na tabela, a freqüência acumulada é 34. Ou então, olhando-se no gráfico, considerando a tabela, também encontramos esse valor. Se se quiser a porcentagem de empresas com volume de exportação até 80 000 dólares, olha-se no eixo vertical à direita do gráfico e encontra-se um valor aproximado de 52%

Gráfico de Frequências Acumuladas "Acima de"



Para determinar o número de empresas com volume de exportação superior a 90 000 dólares faz-se a correspondência entre essa quantia e a freqüência acumulada (olhando no eixo vertical à esquerda). Temos que esse valor é aproximadamente é aproximadamente igual a 17. Quando se quer verificar a porcentagem, olha-se no eixo vertical à esquerda do gráfico, ou seja, há aproximadamente 26% das empresas exportando mais de 90 000 dólares.

8 – MEDIDAS DE POSIÇÃO

Existem medidas que sintetizam certas características importantes de uma distribuição de frequências. São elas: as medidas de posição ou de tendência central, as de dispersão ou de heterogeneidade, as de assimetria e as medidas de curtose. Entre elas, as medidas de posição e de dispersão são as mais importantes.

Entre as medidas de tendência central, as mais usadas são a média aritmética, a moda e a mediana. Existem outras: média geométrica, média harmônica, quadrática, cúbica e média biquadrática. Aqui tratar-se-á somente das médias aritmética, geométrica, harmônica, moda e mediana.

8.1 - MÉDIA ARITMÉTICA

Média aritmética simples: A média aritmética ($\overline{x}$, lê-se x barra) simples é o quociente entre o conjunto de valores e o número total de valores.

Exemplo: Os salários das funcionárias de uma micro empresa são: R$450,00; R$ 500,00; R$ 550,00; R$ 600,00 e R$ 650,00. Qual o salário médio dessa empresa?

$$\overline{x} = \frac{450+500+550+600+650}{5} = 550$$

De forma geral tem-se que $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, onde x_i são os valores

Média aritmética ponderada: è quando a freqüência absoluta de cada valor é diferente de zero ou quando os valores possuem pesos diferentes.

Por exemplo: Em uma escola são feitas, em cada etapa, três tipos de avaliação: exercícios em grupo, que tem peso 1, prova com consulta, que tem peso 2 e prova individual sem consulta, cujo o peso é 3. Se um aluno teve, em um total de 10 pontos em cada avaliação, 8 nos exercícios em grupo, 9 na prova com consulta e 7 na prova sem consulta, qual será sua média?

Tem-se: $\frac{1.8+2.9+3.7}{6} = \frac{47}{6} = 7{,}8$

Outro exemplo. Considere a tabela a seguir

Distribuição de alunos de uma série,

segundo a idade.

Idade (anos)	f_j
14	11
15	6
16	3
Total	20

$$\overline{x} = \frac{11.14 + 6.15 + 3.16}{20} = \frac{154 + 90 + 48}{20} = \frac{292}{20} = 14,6$$

Generalizando tem-se a fórmula

$$\overline{x} = \frac{\sum_{j=1}^{k} x_j f_j}{\sum_{j=1}^{k} f_j} = \frac{\sum xf}{\sum f} = \frac{\sum xf}{n},$$

onde x_j = valores da variável ou pontos médios das classes

$\sum_{j=1}^{k} f_j$ = n = número total de valores e k = número de classes ou de valores individuais diferentes da variável.

8.2 - MÉDIA GEOMÉTRICA

A média geométrica ($\overline{x_g}$) pode ser simples ou ponderada.

Média geométrica simples de uma sequência $x_1, x_2, x_3, \ldots, x_n$ é definida por

$$\overline{x_g} = \sqrt[n]{x_1.x_2.x_3\ldots x_n}$$

Assim, a média geométrica simples de um conjunto de dados {10, 30,90} é

$$\overline{x_g} = \sqrt[3]{10.30.90} = \sqrt[3]{27000} = 30$$

A média geométrica ponderada é dada por

$$\overline{x_g} = \sqrt[\sum f_n]{x_1^{f_1}.x_2^{f_2}.x_3^{f_3}\ldots.x_n^{f_n}},$$

onde cada $x_n^{f_n}$ é igual a valor n elevado à sua respectiva freqüência e $\sum f_n$ é a soma das frequências.

Exemplo: Considere a tabela

Distribuição de famílias, segundo o número de filhos

Nº de filhos	f_j
1	3
2	3
5	1

$$\overline{x_g} = \sqrt[7]{1^3.2^3.5^1} = \sqrt[7]{1.8.5} = \sqrt[7]{45} \cong 1{,}69$$

8.3 - MÉDIA HARMÔNICA ($\overline{x_h}$)

A média harmônica também pode ser simples (quando os valores têm freqüência absoluta igual a 1) ou ponderada (os valores têm frequências diferentes).

Média harmônica simples: É definida como:

$$\overline{x_h} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}} = \overline{x_h} = \frac{n}{\sum\limits_{j=1}^{n} x_j}$$

Exemplo:

Seja um conjunto de dados {2, 5, 6} Tem-se

$$\overline{x_h} = \frac{3}{\frac{1}{2} + \frac{1}{5} + \frac{1}{6}} = \frac{3}{\frac{26}{30}} = \frac{90}{26} \cong 4{,}46$$

Média harmônica ponderada: É definida por

$$\overline{x_h} = \frac{\sum\limits_{j=1}^{n} f_j}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \ldots + \frac{f_n}{x_n}}$$

Exemplo:

Distribuição de famílias segundo o nº de filhos

Nº de filhos	F_j
1	5
2	4
3	2
4	2

$$\overline{x_h} = \frac{13}{\frac{5}{1}+\frac{4}{2}+\frac{2}{3}+\frac{2}{4}} = \frac{13}{\frac{49}{6}} = \frac{78}{49} \cong 1{,}59$$

8.4 - MODA

Moda (M_o) é o valor de maior freqüência. Quando os valores são apresentados em uma tabela de frequências (valores brutos ou rol), basta verificar qual o elemento de maior freqüência. Por exemplo: São dados os valores 5, 7, 1, 3, 4, 5, 8, 6. O valor que aparece mais vezes é o 5, logo essa é uma sequência unimodal (tem uma única moda).

Quando os valores estão tabulados e a variável é discreta, basta olhar o valor de maior freqüência. Por exemplo

Distribuição de famílias segundo o nº de filhos

Nº de filhos	F_j
1	10
2	15
3	8
4	5

O valor que tem maior freqüência é

Quando os dados estão agrupados em classe há diversos métodos para se calcular a moda.

Cálculo da moda pela fórmula de Czuber

Sendo l_{mo} o limite inferior da classe modal, c a amplitude do intervalo de classe, Δ_1 a diferença entre as frequências simples das classes modal e a anterior à classe modal e Δ_2 a diferença entre as frequências simples da classe modal e a posterior a ela, temos

$$M_o = l_{mo} + c.\left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)$$

Exemplo

Seja a seguinte distribuição de frequências

Pesos (kg)	Frequências simples absolutas
$2 \mapsto 4$	9
$4 \mapsto 6$	12
$6 \mapsto 8$	6
$8 \mapsto 10$	2
$10 \mapsto 12$	1

Calcule a moda dessa distribuição

Classe modal: $4 \mapsto 6$

Freqüência simples da classe modal: 12

$l_{mo} = 4$

$\Delta_1 = 12 - 9 = 3$

$\Delta_2 = 12 – 6 = 6$

Amplitude do intervalo: 2

$$M_o = l_{mo} + c.\left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)$$

$M_o = 4 + 2.\left(\frac{3}{3+6}\right)$

$M_o = 4 + 2.\frac{1}{3}$

$M_o = 4{,}66$ kg

Cálculo da moda pelo método de King

Nesse caso a moda é dada pela fórmula

$$M_o = l_{mo} + \frac{f_{post}}{f_{ant} + f_{post}}.h \,,$$

onde l_{mo} = limite inferior da classe de maior freqüência absoluta (classe modal).

f_{post} = freqüência da classe posterior à classe modal

f_{ant} = freqüência da classe anterior à classe modal

h = amplitude do intervalo de classe.

Por exemplo, considere a tabela

Classe	f_j
0 ├ 10	8
10 ├ 20	12
20 ├ 30	25
30 ├ 40	15

A classe modal é 20 ├ 30. Seu limite inferior é 20 e a amplitude 10. Pela fórmula de King, tem-se

Portanto, 25,6 é o valor mais freqüente nessa distribuição.

8.5 - MEDIANA

A mediana 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.

Vejamos como encontrar o valor da Mediana nos seguintes casos:

Mediana em valores não tabulados

Inicialmente, no caso de termos um número ímpar de valores

Por exemplo: X = { 3, 5, 7, 9, **11**, 13, 12, 13, 14}, onde n = 9 (ímpar)

Determinamos o elemento central, *E*, da seguinte forma:

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

Buscamos, então, na seqüência ordenada o valor correspondente à posição E = 5. Nesse caso temos que a mediana será o valor $\mathbf{M_d = 11}$

No caso de termos um número par de valores

Por exemplo: X = { 1, 3, <u>5</u>, <u>7</u>, 9, 11}, onde n = 6 (par)

Determinamos os elementos centrais:

$$E = \frac{n}{2} = \frac{6}{2} = 3$$

Retomamos a seqüência ordenada e identificamos o elemento correspondente à posição 3 (valor de *E*) e a posição seguinte, ou seja *E* = 3 e *E* = 4, que são as posições dos valores centrais. Para determinar a Mediana calculamos a **média aritmética dos dois valores centrais.**

Temos:

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

Mediana em **valores tabulados**

Os valores tabulados podem se apresentar agrupados em classes ou não.

Trataremos inicialmente dos valores não agrupados em classes (dados discretos);

Veja:

x_i	f_{ji}
2	5
4	10
6	15
8	12
10	5
12	3
	$\sum = 50$

$\sum f_i = n = 50 \Rightarrow$ número **par**

Calculamos *E* para um número par de valores:

$E = \frac{n}{2} = \frac{50}{2} = 25 \therefore E = 25$

Calculamos então o valor das freqüências acumuladas:

x_i	f_{ji}	f_{ac}
2	5	5
4	10	15
6	15	30
8	12	42
10	5	47
12	3	50
	$\sum = 50$	

Buscamos a posição, E = 25 na relação das freqüências acumuladas (f_{ac}), verificamos que ela esta após o 15 e antes do 30. Sendo o somatório das freqüências par ($\sum = 50$), calculamos a mediana pela média aritmética entre as posições 25 e 26, que no caso são 6 e 6.

Assim: $M_d = \frac{6+6}{2}$ $\therefore$ $\mathbf{M_d = 6}$

Fique atento para o caso de **n** ser ímpar, nesse caso teremos E = $\frac{n+1}{2}$. Resolvemos de maneira análoga e obtemos a mediana, lembrando que neste caso teremos um único valor central, não sendo necessário o cálculo da média aritmética.

Veja, agora, como calculamos a mediana no caso de termos os dados agrupados em classes (dados contínuos).

A fórmula para o cálculo da mediana de dados agrupados será:

$$\boxed{M_d = l_i + h\frac{E - f_{fantac}}{f_{Md}}}$$

Onde:

l_i $\rightarrow$ limite inferior da classe mediana

h $\rightarrow$ amplitude o intervalo da classe (diferença entre os limites superior e inferior de uma classe)

E $\rightarrow$ será sempre determinado por $\frac{n}{2}$

f_{Md} $\rightarrow$ freqüência absoluta da classe mediana

f_{fantac} $\rightarrow$ freqüência acumulada anterior à classe mediana

Veja o exemplo:

Classe	f_j	f_{ac}

1 ├ 3	2	2
3 ├ 5	3	5
5 ├ 7	5	10
7 ├ 9	4	14
9 ├ 11	6	20
	$\sum = 20$	

E = $\frac{n}{2} = \frac{20}{2} = 10$

Classe mediana: 5 ├ 7

$M_d = l_i + h\frac{E - f_{fantac}}{f_{Md}} \Rightarrow M_d = 5 + 2\frac{10-5}{5}$

$M_d = 5 + \frac{5 \cdot 2}{5} \Rightarrow M_d = 5 + 2 \Rightarrow$ **M_d = 7**

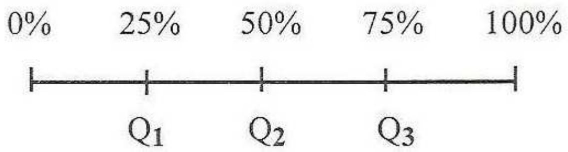
8.6. SEPARATRIZES

Das medidas chamadas separatrizes, a mediana, que acabamos de estudar, é uma delas que tem a característica de separar a série em dois grupos que com o mesmo número de valores.

Existem outras separatrizes tão importantes quando a mediatriz, que dividem a série de outras formas. São elas: os **quartis** (divide a série em quatro parte iguais), os **decis** (divide a série em dez parte iguais) e os **percentis** (divide a série em cem parte iguais).

Quartis:

Os quartis dividem um conjunto de valores de uma distribuição de freqüência em quatro partes iguais.



Primeiro Quartil (Q_1) $\rightarrow$ 25% da distribuição esta à sua esquerda e 75% á sua direita.

Segundo Quartil (Q_2) $\rightarrow$ 50% da distribuição esta à sua esquerda e 50% á sua direita.

O segundo quartil é a própria mediana.

Terceiro Quartil (Q_3) $\rightarrow$ 75% da distribuição esta à sua esquerda e 25% á sua direita.

Para calcular os quartis utilizamos, de forma análoga, a fórmula para o cálculo da mediana substituindo o valor de E (posição desejada), E = $\frac{n}{2}$ por

$E_{Qi} = \frac{in}{4}$, onde i = 1, 2, 3.

Temos:

$$Q_i = l_i + \frac{E_{Qi} - f_{fantac}}{f_{Qi}}$$

Decis:

Os decis dividem o conjunto de valores de uma distribuição de freqüência em 10 partes iguais.

Calculamos o elemento do decil:

$E_{Di} = \frac{in}{10}$, onde i = 1, 2, 3, ..., 9.

Calculamos o decil desejado:

$$\boxed{D_i = l_i + h\frac{E_{Di} - f_{fantac}}{f_{Di}}}$$

Centis:

De maneira análoga aos outros cálculos, calculamos o elemento centil:

$E_{Ci} = \frac{in}{100}$, onde i = 1, 2, 3, ..., 39,..., 99.

O centil desejado será calculado pela fórmula:

$$\boxed{C_i = l_i + h\frac{E_{Ci} - f_{fantac}}{f_{Ci}}}$$

Veja o exemplo:

Ao aplicar uma prova de Estatística a uma turma de 40 alunos, encontrou-se o resultado tabelado da seguinte forma:

Classes	freqüência= fi	Freqüência acumulada
50 ├ 54	4	4
54 ├ 58	9	13
58 ├ 62	11	24
62 ├ 66	8	32
66 ├ 70	5	37
70 ├ 74	3	40
Total	40	

Desejamos saber:

os quartis da distribuição dada:

Primeiro Quartil

$Q_i = l_i + h\dfrac{E_{Qi} - f_{fantac}}{f_{Qi}}$

$E_{Q1} = \dfrac{40}{4} \Rightarrow E_{Q1} = 10$, logo a classe do 1º Quartil será : 54 ├ 58

$L_i = 54$, $h = 4$, $f_{antac} = 4$, $f_{Qi} = 9$

$Q_i = 54 + 4\frac{10-4}{9}$

$Q_i = 56{,}66$

Segundo quartil

Calcular o segundo quartil é o mesmo que calcular a mediana.

$E_{Q2} = 2.\ \frac{40}{4} \Rightarrow E_{Q1} = 20$, logo a classe do 2º Quartil será : 58 ├ 62

$L_i = 58$, $h = 4$, $f_{antac} = 13$, $f_{Qi} = 11$

$Q_i = 58 + 4\frac{20-13}{11}$

$Q_i = 60{,}54$

Terceiro quartil

$E_{Q2} = 3.\ \frac{40}{4} \Rightarrow E_{Q1} = 30$, logo a classe do 2º Quartil será : 62 ├ 66

$L_i = 62$, $h = 4$, $f_{antac} = 24$, $f_{Qi} = 8$

$Q_i = 62 + 4\frac{30-24}{8}$

$Q_i = 65$

U

9 – MEDIDAS DE DISPERSÃO OU VARIABILIDADE

Estudando as medidas de posição vimos que um conjunto de dados pode ser resumido através dos valores da média aritmética, da moda e da mediana.

Entretanto, para representação de dados não basta a medida de posição, é necessário ter a noção de concentração (homogeneidade, heterogeneidade) existente entre os dados.

Para tal estudaremos a seguir as medidas de dispersão absoluta (amplitude total, desvio médio, variância e desvio padrão), e as medidas de dispersão relativa (coeficiente de variação).

Considere o exemplo a seguir:

X = 30, 30, 30, 30, 30

Y = 26, 28, 30, 32, 34

Z = 8, 22, 28, 42, 50

As médias aritméticas de X, Y, e Z são: $\overline{X}$ = 30, $\overline{Y}$ = 30 e $\overline{Z}$ = 30.

Os conjuntos X, Y e Z possuem a mesma média aritmética, entretanto se diferem em relação a variabilidade. O conjunto X tem dispersão ou variabilidade nula, enquanto o conjunto Y tem uma dispersão menor que o conjunto Z.

9.1 - AMPLITUDE TOTAL (A_t)

Chamamos de amplitude total (A_t) a diferença entre o maior e o menor valor dos dados observados.

$\mathbf{A_t = X_{(máx)} - X_{(min)}}$

Na sua determinação podemos encontrar as seguintes situações:

Dados não agrupados:

Exemplos:

X = 30, 30, 30, 30, 30

$A_t = 30 – 30 = 0$

$A_t = 0$ (dispersão nula)

Y = 26, 28, 30, 32, 34

$A_t = 34 – 26$

$A_t = 8$

Z = 8, 22, 28, 42, 50

$A_t = 50 – 8$

$A_t = 42$

Dados tabulados não agregados em classes (dados discretos):

Exemplo:

x_i	f_{ji}
1	10
3	20
5	40
7	20
9	10

$A_t = 9 – 1 = 8$

Dados tabulados agregados em classes (dados contínuos)

Classe	f_j
0 ├ 10	8
10 ├ 20	12
20 ├ 30	25
30 ├ 40	15

Neste caso a amplitude total será calculada pela diferença entre o limite superior da última classe e o limite inferior da primeira classe.

$A_t = L_{max} – L_{min}$

$A_t = 40 - 0$

$A_t = 40$

9.2 – DESVIO MÉDIO

Para determinamos o desvio médio(DM) de uma distribuição estabelecemos a relação entre a média aritmética dos módulos dos desvios e a média aritmética da série.

Vejamos o cálculo do desvio médio para **dados não agrupados**:

Seja a série: (4, 6, 8, 12)

a média dessa distribuição é:

$$\overline{X} = \frac{4+6+8+12}{4} = 7{,}5$$

Obtermos o desvio (d_i) calculando a diferença entre cada valor da distribuição e o $\overline{X}$

x_i	$d_i = x_i - \overline{x}$	$\lvert d_i \rvert$
4	4 – 7,5	3,5
6	6 - 7,5	1,5
8	8 - 7,5	0,5
12	12 - 7,5	4,5
		$\sum = 10$

$$DM = \frac{\Sigma\lvert d_i \rvert}{n} \Rightarrow DM = \frac{10}{4} \Rightarrow \mathbf{DM = 2{,}5}$$

onde:

DM = desvio médio

n = número de termos da série e

$\lvert d_i \rvert$ = módulo da diferença entre cada termo da série e sua média geométrica.

Cálculo do desvio médio para **valores agrupados** em uma distribuição de freqüência.

x_i	f_{ji}
1	2
2	3
3	4
4	5
5	8
	$\sum 22$

Calculando $x_i . f_{,i,}$ temos

x_i	f_{ji}	$x_i f,$
1	2	2
2	3	6
3	4	12
4	5	20
5	8	40
	$\sum = 22$	$\sum = 80$

A média dos valores será:

$$\bar{x} = \frac{\sum_{i=1}^{n} xifi}{\sum_{i=1}^{n} f_i} \Rightarrow \bar{x} = \frac{80}{22} \Rightarrow \bar{x} = 3{,}6364$$

Podemos, agora, calcular o desvio médio;

Veja como ficará nosso quadro de distribuições:

x_i	f_{ji}	$d_i = x_i - \bar{x}$	$\lvert d_i \rvert$	$\lvert d_i \rvert \cdot f_i$
1	2	1 – 3,6364	2,6364	5,2728
2	3	2 - 3,6364	1,6364	4,9092
3	4	3 - 3,6364	0,6364	2,5456
4	5	4 - 3,6364	0,3636	1,8180
5	8	5 - 3,6364	1,3636	10.9088
	$\sum = 22$			$\sum = 25{,}4544$

DM = $\dfrac{\sum \lvert d_i \rvert . f_i}{\sum f_i}$ $\Rightarrow$ DM = $\dfrac{25{,}4544}{22}$ $\Rightarrow$ **DM = 1,1570**

Dados agrupados de uma **distribuição em classes**

Classe	f_j	Ponto Médio (x_{pm})	$X_{pm}. f_i$
0 ├ 2	2	1	2
2 ├ 4	1	3	3
4 ├ 6	2	5	10

6 ├ 8 8 ├ 10	3 4	7 9	21 36
	$\sum = 12$		$\sum = 72$

Calculando a média aritmética:

$$\bar{x} = \frac{\Sigma\left(x_{pm}.f_i\right)}{\Sigma f_i} \Rightarrow \bar{x} = \frac{72}{12} \Rightarrow \bar{x} = 6$$

Calculando o desvio médio:

$$DM = \frac{\Sigma\left(\left|x_{pm} - \bar{x}\right|.f_i\right)}{\Sigma f_i}$$

$$DM = \frac{|1-6|.2+|3-6|.1+|5-6|.2+|7-6|.3+|9-6|.4}{12}$$

$$DM = \frac{5.2+3.1+1.2+1.3+3.4}{12}$$

DM = 2,5

9.3 - VARIÂNCIA

Para o cálculo da variância (S^2)utilizaremos os valores de d_i , f_i já calculados nos exemplos anteriores.

Para dados não tabulados a variância será calculada por:

$$S^2 = \frac{\Sigma d_i^2}{n}$$

Para dados tabulados a variância será:

$S^2 = \dfrac{\sum d_i^2 . f_i}{n}$

9.4 - DESVIO PADRÃO

É a medida de dispersão mais usada, pois leva em consideração a totalidade dos valores da variável em estudo. É um indicador de variabilidade bastante confiável. O desvio padrão baseia-se no desvios em torno da média aritmética e sua fórmula básica pode ser traduzida como : a raiz quadrada da média aritmética dos quadrados dos desvios e é representada por:

$S = \sqrt{\dfrac{\sum d_i^2}{n}}$ onde $d_i = x_i - \overline{x}$, quando temos dados não tabulados e

$S = \sqrt{\dfrac{\sum d_i^2 f_i}{n}}$, para dados tabulados.

Observe que se o cálculo do desvio padrão for aplicado a uma amostra, utiliza-se **n - 1** no lugar de **n** nas duas formulas dadas. (isto é feito pela utilização do "fator de correção de Bessl")

No caso de dados agrupados em uma distribuição de freqüência por classes, temos:

$$S = \sqrt{\dfrac{\sum \left(x_{pm} - \overline{x}\right) f_i}{\sum f_i}}$$

Veja o exemplo:

Classe	f_j	Ponto Médio (x_{pm})	$X_{pm} - \overline{x}$	$\left(x_{pm} - \overline{x}\right)^2$	$\left(x_{pm} - \overline{x}\right)^2 .f_i$
0 ├ 2	2	1	-4,92	24,21	48,42
2 ├ 4	3	3	-2,92	8,53	25,59
4 ├ 6	5	5	-0.92	0,85	4,25
6 ├ 8	10	7	1,08	1,17	11,7

8 ├ 10	4	9	3,08	9,49	37,96
	$\sum = 24$				$\sum = 72$

$$S = \sqrt{\frac{\Sigma\left(x_{pm} - \overline{x}\right) f_i}{\Sigma f_i}} \Rightarrow S = \sqrt{\frac{127{,}92}{24}} \Rightarrow S = 2{,}3087$$

10 – MEDIDAS DE ASSIMETRIA

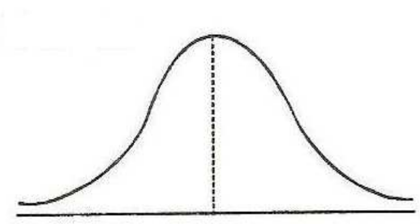
10.1 - ASSIMETRIA

Introdução

Podemos definir assimetria como o desvio ou afastamento da simetria de uma cura de freqüência de uma distribuição estatística. Assim, podemos caracterizar as distribuições de freqüência em : assimétrica à direita, assimétrica à esquerda e simétrica.

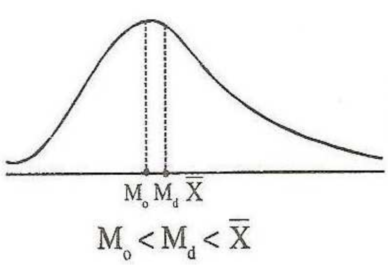
Veja os exemplos:

Simétrica



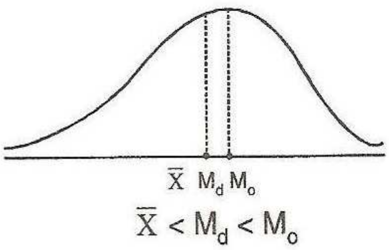
Média = Mediana = Moda

Assimétrica à direita



Moda menor que a mediana e mediana menor que a média.

Assimétrica à esquerda



Média menor que mediana e mediana menor que a moda.

10.2 - MEDIDAS DE ASSIMETRIA

A_1 : 1º coeficiente de assimetria de Pearson

$A_1 = \dfrac{\overline{x} - M_o}{S}$, onde $\overline{X}$ = Média

M_o = Moda

S = desvio padrão

Quando $\overline{X} = M_o \Rightarrow$ Simétrica

Quando $\overline{X} < M_o \Rightarrow$ Assimétrica negativa

Quando $\overline{X} > M_o \Rightarrow$ Assimetria positiva

A_2 = 2º coeficiente de assimetria de Pearson

$A_2 = \dfrac{3\left(\overline{x} - M_d\right)}{S}$, onde $\overline{X}$ = Média

M_d = Moda

S = desvio padrão

Quando $\overline{X} = M_d \Rightarrow$ Simétrica

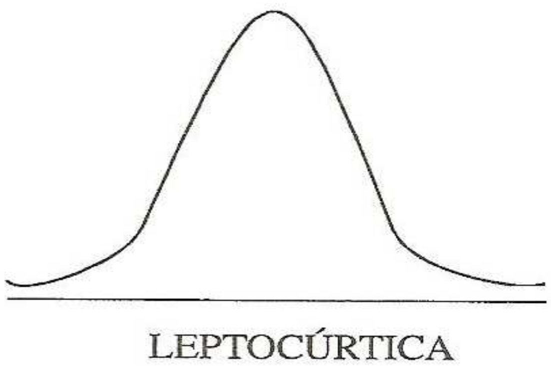
Quando $\overline{X} < M_d \Rightarrow$ Assimétrica negativa

Quando $\overline{X} > M_d \Rightarrow$ Assimetria positiva

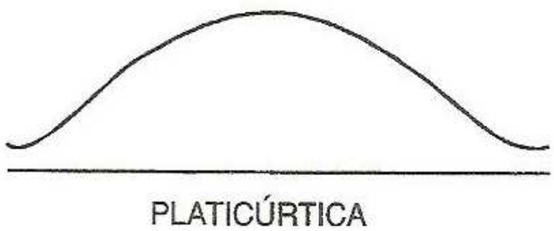
10.3 – CURTOSE

Denominamos **CURTOSE** o grau de achatamento de uma distribuição em relação a uma distribuição padrão, denominada curva normal (curva correspondente a uma distribuição teórica de probabilidade).

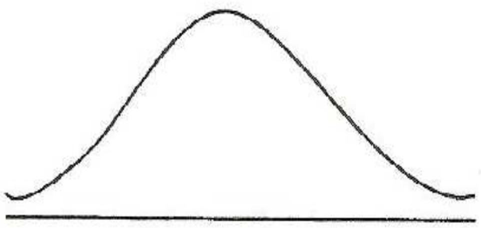
Quando a distribuição apresenta uma **curva de freqüência mais fechada que a normal** (ou mais aguda ou afilada em sua parte superior), ela recebe o nome de **leptocúrtica**.



Quando a distribuição apresenta uma **curva de freqüência mais aberta que a normal** (ou mais achatada em sua parte superior), ela recebe o nome de **platicúrtica**



A **curva normal**, que é a nossa base referencial, recebe o nome de **mesocúrtica**



MESOCÚRTICA

Coeficiente de curtose

$$C_1 = \frac{Q3 - Q1}{2(P90 - P10)}$$

Este coeficiente é conhecido como percentílico de curtose.

Relativamente a curva normal, temos:

C1 = 0,263 ➔ curva mesocúrtica

C1 < 0,263 ➔ curva leptocúrtica

C1 > 0,263 ➔ curva platicúrtica

O coeficiente (C2):

$$\frac{\dfrac{\sum (Xi - \overline{X})^4 . fi}{\sum fi}}{S^4}$$

onde S é desvio padrão

C2 = 3 ➔ curva mesocúrtica

C2 > 3 ➔ curva leptocúrtica

C2 < 3 ➔ curva platicúrtica

EXERCICIOS

1. População ou universo é:
   a) Um conjunto de pessoas;
   b) Um conjunto de elementos quaisquer
   c) Um conjunto de pessoas com uma característica comum;
   d) Um conjunto de elementos com pelo menos uma característica em comum;
   e) Um conjunto de indivíduo de um mesmo município, estado ou país.

2. Uma parte da população retirada para analisá-la denomina-se:
   a) Universo;
   b) Parte;
   c) Pedaço;
   d) Dados Brutos;
   e) Amostra.

3. A parte da estatística que se preocupa somente com a descrição de determinadas características de um grupo, sem tirar conclusões sobre um grupo maior denomina-se:
   a) Estatística de População;
   b) Estatística de Amostra;
   c) Estatística Inferencial
   d) Estatística Descritiva;
   e) Estatística Grupal.

4. Uma série estatística é denominada Temporal quando?

a) O elemento variável é o tempo;

b) O elemento variável é o local;

c) O elemento variável é a espécie;

d) É o resultado da combinação de séries estatísticas de tipos diferentes;

e) Os dados são agrupados em subintervalos do intervalo observado.

5. Assinale a afirmativa verdadeira:

a) Um gráfico de barras ou colunas é aquele em que os retângulos que o compõem estão dispostos horizontalmente.

b) Um gráfico de barras ou colunas é aquele em que os retângulos que o compõem estão dispostos verticalmente.

c) Um gráfico de barras é aquele em que os retângulos que o compõem estão dispostos verticalmente e um gráfico de colunas, horizontalmente.

d) Um gráfico de barras é aquele em que os retângulos que o compõem estão dispostos horizontalmente e um gráfico de colunas, verticalmente.

e) Todas as alternativa anteriores são falsas.

As questões de 6 a 12 se referem ao enunciado abaixo:

Um dado foi lançado 50 vezes e foram registrados os seguintes resultados

5	4	6	1	2	5	3	1	3	3
4	4	1	5	5	6	1	2	5	1
3	4	5	1	1	6	6	2	1	1
4	4	4	3	4	3	2	2	2	3
6	6	3	2	4	2	6	6	2	1

Construa uma distribuição de freqüência sem intervalo de classe e determine:

6. A amplitude Total (n)

a)5

b)6

c)7

d)10

e)50

7- A freqüência total

a) 5

b) 6

c) 7

d)10

e) 50

8- A freqüência simples absoluta do primeiro elemento:

a)10%

b)20%

c)1

d)10

e)20

9-A freqüência simples relativa do primeiro elemento:

a) 10%

b) 20%

c) 1

d) 10

e) 20

10- A freqüência acumulada do primeiro elemento:

a) 10%

b) 20%

c) 1

d) 10

e) 20

11- A freqüência acumulada relativa do primeiro elemento:

a)10%

b)20%

c) 1

d)10

e)84%

12 A freqüência simples absoluta do segundo elemento:

a)19

b) 9

c) 2

d) 38%

e)18%

13- A freqüência acumulada relativa do sexto elemento:

a) 50

b) 8

c) 6

a)100%

b)16%

14- Considere as afirmações abaixo:

I- A estatística aplicada é o ramo da estatística que se preocupa coam a coleta, organização, interpretação de dados.

II- A estatística indutiva ou inferencial esta voltada para a coleta, organização e interpretação de dados.

III- A estatística descritiva se preocupa com a análise e interpretação dos dados.

IV- Entendemos por população ou universo o conjunto de pessoas ou objetos que apresentam uma característica comum.

As afirmativas corretas são:

a) II e III
b) IV e I
c) II e I
d) III e IV
e) n.r.a.

15 – (Fundação Carlos Chagas) Constitui ramo da teoria estatística conhecido como inferência estatística.

a) organização de dados qualitativos e quantitativos de uma amostra.

b) técnicas que servem de instrumentos para a descrição de um conjunto de dados.

c) elaboração de gráficos com base em uma coleção de dados.

d) métodos que permitem analisar os dados de uma população, independente de se tirar quaisquer conclusões.

e) métodos que tornam possível a estimação de características de uma população baseados nos resultados amostrais.

16 - (TTN) Marque a opção correta:

a) um evento tem, no mínimo, dois elementos do espaço amostral de um experimento aleatório.

b) em um experimento aleatório uniforme, todos os elementos do espaço amostral são iguais.

c) Dois experimentos aleatórios distintos têm, necessariamente, espaços amostrais distintos.

d) Uma parte não nula do espaço amostral de um experimento aleatório define um evento.

e)Um experimento aleatório pode ser repetido indefinidamente, mantidas as condições iniciais.

17 – Assinale a alternativa incorreta:

a) A estatística é um ramo da matemática aplicada que se preocupa com a coleta, organização e descrição dos dados.

b) Entendemos por população um conjunto de indivíduos ou objetos que apresentam um atributo comum.

c) A quantidade de pagamentos é uma variável qualitativa.

d) Espaço amostral é o conjunto que reúne todos os elementos de uma amostra.

e) n.r.a.

18 – (AFC) A tabela abaixo representa a distribuição de um grupo de 200 estudantes Segundo o curso que fazem (matemática ou estatística) e o sexo (homem ou mulher)

	Homem	Mulher
estatística	40	20
matemática	80	60

A alternativa incorreta é:

a) 40% dos homens estudam matemática.

b) 75% das mulheres fazem o curso de matemática.

c) dois em cada três estudantes de estatística são homens.

d) um em cada três homens faz o curso de estatística.

e) 60% dos estudantes são homens.

19 -(TNT) Os intervalos de classes podem ser apresentados de várias maneiras. Dentre as situações abaixo, a correta é:

a) 2___6 compreende todos os valores entre 2 e 6, inclusive os extremos

b) 2|__| 6 compreende todos os valores entre 2 e 6 exclusive os extremos.

c) 2|__ 6 compreende todos os valores entre 2 e 6 exclusive o 2 e inclusive o 6

d) 2 __ | 6 compreende todos os valores entre 2 e 6 inclusive o 2 e exclusive o 6

e) 2___6 compreende todos os valores entre 2 e 6, exclusive os extremos

Responda às questões 20 a 24 baseando-se na tabela abaixo:

salários mínimos (R$)	f_i
$0 \mid\!- 2$	3
$2 \mid\!- 4$	2
$4 \mid\!- 6$	5
$6 \mid\!- 8$	3
$8 \mid\!- 10$	2
	$\sum = 15$

20 – O ponto médio da terceira classe é:

a) 3

b) 4

c) 5

d) 6

e) n.r.a.

21 – O limite inferior da última classe é:

a) 4

b) 6

c) 7

d) 8

e) n.r.a.

22 – O limite superior da quarta classe é :

a) 2

b) 7

c) 8

d) 9

e) n.r.a.

23 – A amplitude total da distribuição é:

a) 9

b) 10

c) 8

d) 11

e) n.r.a.

24 – A porcentagem das pessoas que ganham até quatro salários mínimos é:

a) 33,33%

b) 50,71%

c) 45,6%

d) 50,71%

e) n. r. a.

25 –(TRF) O coeficiente de correlação entre duas variáveis Y e X é igual a +0,8. Considere, agora, a variável Z definida como:Z = 0,2 - 0,5 XO coeficiente de correlação entre as variáveis Z e X, e o coeficiente de correlação entre as variáveis Z e Y serão iguais, respectivamente, a:

a)–1,0;–0,8

b)+1,0;+0,8

c)–0,5;–0,8

d)–0,5;+0,8

e) – 0,2; – 0,4

26 – (TRF) No gráfico abaixo, as colunas representam as freqüências relativas do número de aparelhos de rádio por domicílio em uma certa área da cidade:O exame da forma da distribuição das freqüências relativas permite concluir corretamente que, nesse caso,epara essa variável:

a) A moda é maior do que a mediana, e a mediana maior do que a média.

b) A média é maior do que a moda, e a moda maior do que a mediana.

c) A média é maior do que a mediana, e a mediana maior do que a moda.

d) A moda é maior do que a média, e a média maior do que a mediana.

e) A mediana é maior do que a moda, e a moda maior do que média.

27 – (TRF) Paulo e Helena jogam, cada um, uma moeda. Se do lançamento dessas duas moedas resultar duas caras, Paulo paga a Helena R$ 5,00. Dando qualquer outro resultado, Helena paga a Paulo R$ 2,00. Supondo que ambas as moedas sejam estatisticamente honestas, o valor esperado dos ganhos de Helena (considerando-se como ganhos negativos os valores que ela paga a Paulo) é igual a

a) – R$ 0,25

b) + R$ 0,25

c) + R$ 3,00

d) – R$ 1,50

e) + R$ 1,25

28 – (TRF) Sobre a moda de uma variável, é correto afirmar que

a) para toda variável existe uma e apenas uma moda.

b) a moda é uma medida de dispersão relativa.

c) a moda é uma medida não afetada por valores extremos.

d) em distribuições assimétricas, o valor da moda encontra-se entre o valor da média e o da mediana.

e) sendo o valor mais provável da distribuição, a moda, tal como a probabilidade, pode assumir valores somente no intervalo entre zero e a unidade.

29 –(TRF) Um motorista de táxi faz 10 viagens ida-e-volta do aeroporto Santos Dumont ao aeroporto do Galeão, no Rio de Janeiro. Ele calcula e anota a velocidade média, em quilômetros por hora, em cada uma dessas viagens. O motorista quer, agora, saber qual a velocidade média do táxi para aquele percurso, em quilômetros por hora, considerando todas as 10 viagens ida-e-volta. Para tanto, ele deve calcular a média

a) aritmética dos inversos das velocidades médias observadas.

b) geométrica das velocidades médias observadas.

c) aritmética das velocidades médias observadas.

d) harmônica das velocidades médias observadas.

e) harmônica dos inversos das velocidades médias observadas.20

30 –(TRF) Considere a seguinte distribuição das freqüências absolutas dos salários mensais, em R$, referentes a 200 trabalhadores de uma indústria (os intervalos são fechados à esquerda e abertos à direita).

Classes de Salários: Freqüências Absolutas de

R$ 400 até R$ 500 : 50

de R$ 500 até R$ 600: 70

de R$ 600 até R$ 700 :40

de R$ 700 até R$ 800 : 30

de R$ 800 até R$ 900: 10

Sobre essa distribuição de salários é correto afirmar que:

a) O salário modal encontra-se na classe de R$ 800 até R$ 900.
b) O salário mediano encontra-se na classe de R$ 600 até R$ 700.
c) O salário modal encontra-se na classe de R$ 600 até R$ 700.
d) O salário modal encontra-se na classe de R$ 700 até R$ 800.
e) O salário mediano encontra-se na classe de R$ 500 até R$ 600.

31 - Para dados agrupados representados por uma curva de freqüências, as diferenças entre os valores da média, da mediana e da moda são indicadores da assimetria da curva. Indique a relação entre essas medidas de posição para uma distribuição negativamente assimétrica.

a) A média apresenta o maior valor e a mediana se encontra abaixo da moda.
b) A moda apresenta o maior valor e a média se encontra abaixo da mediana.
c) A média apresenta o menor valor e a mediana se encontra abaixo da moda.
d) A média, a mediana e a moda são coincidentes em valor.
e) A moda apresenta o menor valor e a mediana se encontra abaixo da média.

32 - Assinale a opção que expresse a relação entre as médias aritmética (X), geométrica (G) e harmônica (H), para um conjunto de n valores positivos (X , X , ..., X):1 2 n

a) G <_ H <_ A , com G = H = -X somente se os n valores forem todos iguais.
b) G 5 X <_ H, com G = X = H somente se os n valores forem todos iguais.
c) X <_ G 5 H, com X = G = H somente se os n valores forem todos iguais.
d) H 5 G 5 X, com H = G = X somente se os n valores forem todos iguais.
e) X <_ H 5 G, com X = H = G somente se os n valores forem todos iguais

33 - A distribuição de frequência correspondente aos diferentes preços de um determinado produto em vinte lojas pesquisadas. O número de lojas que apresentaram o preço de R$ 52,00, é:

Preços	No. De lojas
50	2
51	5
52	6
53	6
54	1
Total	20

a) 6
b) 2
c) 5
d) 1
e) N.d.a.

O valor da mediana nas distribuições de freqüência das **questões 34, 53 e 36,** é:

34- 82, 86, 88, 84, 91, 93

a) 88
b) 86
c) 87
d) 91
e) N.d.a.

35-

Xi	73	75	77	79	81
Fi	2	10	12	5	2

a) 77

b) 77,5

c) 73

d) 75

e) N.d.a.

36-

Classes	1 ⊢→ 3	3 ⊢→ 5	5 ⊢→ 7	7 ⊢→ 9	9 ⊢→ 11	11 ⊢→ 13
f_i	3	5	8	6	4	3

a) 6,62

b) 6,63

c) 7,23

d) 7,24

e) N.d.a.

O valor da moda nas distribuições de freqüência das **questões 37, 38 e 39**, é:

37- 3, 4, 7, 7, 7, 8, 9, 10

a) 8

b) 2

c) 5

d) 1

e) n. r. a.

38 -

Xi	2,5	3,5	4,5	6,5
Fi	7	17	10	5

a) 17
b) 10
c) 3,5
d) 6,5
e) N.d.a.

39-

Classes	10 - 20	20 - 30	30 - 40	40 - 50
Fi	7	19	28	32

a) 30,31
b) 40,11
c) 41,11
d) 40,00
e) N.d.a.

40 - Desvio Médio, Variância e Coeficiente de variação são medidas de :

a) Assimetria

b) Dispersão

c) Posição

d) Curtose

41 - Desvio Médio para o conjunto de dados abaixo será:

xi	Fi
5	2
7	3
8	5
9	4
11	2

a) 1,28

b) 1,20

c) 1,00

d) 0,83

42- O Desvio Padrão de um conjunto de dados é 9. A variância é:

a) 3

b) 36

c) 81

d) 18

43- Na distribuição de valores iguais, o Desvio padrão é:

a) negativo

b) a unidade

c) zero

d) positivo

44- O calculo da variância supõe o conhecimento da:

a) Fac

b) média

c) mediana

d) moda

45- A variância do conjunto de dados tabelados abaixo será:

Classes	Fi
03 |- 08	5
08 |- 13	15
13 |- 18	20
18 |- 23	10

a) 1,36

b) 18,35

c) 4,54

d) 20,66

46- Numa empresa o salário médio dos homens é de R$ 4000,00 com um desvio padrão de R$1500,00, e o das mulheres é na média de R$3000,00 com desvio padrão de R$1200,00. Qual dos sexos apresenta maior dispersão. (Analise pelo C.V.)

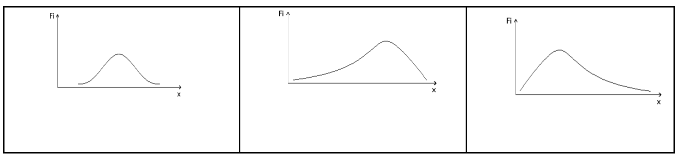
a) as mulheres

b) os homens

c) homens e mulheres

d) nenhuma das anteriores

47- Analisando as curvas abaixo marque a resposta correta.



(I) (II) (III)

a) a curva I é simétrica - $\overline{x} > \text{med} > \text{mo}$;

b) a curva II é assimétrica positiva - $mo > \sigma^2 > \bar{x}$;

c) a curva I é simétrica $\bar{x} = med = mo$;

d) a curva III é simétrica positiva $\bar{x} = med = mo$;

48- Para as distribuições abaixo foram calculados

Classes	Fi
02 |- 06	6
06 |- 10	12
10 |- 14	24
14 |- 18	12
18 |- 22	6

Distrib. A

Classes	Fi
02 |- 06	6
06 |- 10	12
10 |- 14	24
14 |- 18	30
18 |- 22	6

Distrib. B

Classes	Fi
02 |- 06	6
06 |- 10	30
10 |- 14	24
14 |- 18	12
18 |- 22	6

Distrib. C

Distrib. A: $\bar{x} = 12Kg$, Med = 12Kg, Mo = 12Kg, S = 4,42Kg

Distrib. B: $\bar{x} = 12,9Kg$, Med = 13,5Kg, Mo = 16Kg, S = 4,20Kg

Distrib. C: $\bar{x} = 11,1Kg$, Med = 10,5Kg, Mo = 8Kg, S = 4,20Kg

Marque a alternativa correta:

a) a distribuição I é assimétrica negativa;

b) a distribuição II é assimétrica positiva;

c) a distribuição III é assimétrica negativa moderada.

d) a distribuição I é simétrica;

49 - Considerando o conjunto de dados abaixo, o valor da média aritmética, da mediana e da moda, respectivamente são:

3, 5, 2, 6, 5, 9, 5, 2, 8, 6

a) $\overline{X} = 5{,}1$; $M_d = 5$; $M_o = 5$

b) $\overline{X} = 9{,}1$; $M_d = 5$; $M_o = 5$

c) $\overline{X} = 5{,}1$; $M_d = 6$; $M_o = 6$

d) $\overline{X} = 5{,}1$; $M_d = 9$; $M_o = 5$

e) n. d.a.

50 – Em uma classe de 50 alunos, as notas obtidas em Matemática formaram a seguinte distribuição:

Notas	2	3	4	5	6	7	8	9	10
N° de alunos	1	3	6	10	13	8	5	3	1

A nota modal será:

a) 6

b) 13

c) 5

d) 8

e) n.d.a.

51 - (ESAF/TTN) Dada a seguinte distribuição, onde f_i é a frequência simples absoluta da i-ésima classe, então

Classes	f_i
2 ⊢ 4	2
4 ⊢ 6	8
6 ⊢ 8	10
8 ⊢ 10	8
10 ⊢ 12	4

a) a distribuição é simétrica e o número de classes é 5

b) a distribuição é assimétrica e bimodal

c) a média aritmética é 6,4

d) por ser a maior frequência, a moda é 10

e) o ponto médio da 3ª classe e a moda são iguais

52 - (ESAF/TTN) De acordo com a distribuição transcrita a seguir, pode-se afirmar que

Diâmetro(cm)	Frequência simples absoluta
4 ⊢ 6	6
6 ⊢ 8	8
8 ⊢ 10	12
10 ⊢ 12	10
12 ⊢ 14	4

A moda da distribuição é aproximadamente

a) 9,5cm

b) 9,7cm

c) 9,3cm

d) 9,6cm

e) 9,4cm

53 - (IDR-DF/AFCE) Um órgão público divide suas despesas em doze rubricas diferentes. Os valores (em 1 000 reais) orçados por rubrica para o próximo ano, em ordem crescente, são: 20,22,28,43,43,61,61,61,64,72 e82. pode-se afirmar, então, que a mediana destes valores é:

a) 43

b) 50

c) 52

d) 61

54 - (ESAF/TTN) Considere as medianas dos grupos abaixo:

Grupo I: 10,6,30,2,5,8.

Grupo II: 7,4,2,10,7,15

Grupo III: 5,9,7,33,18,4

Grupo IV: 6,9,4,10,10,11

Os grupos que têm a mesma mediana são:

a) I e II

b) II e III

c) III e IV

d) I e III

e) II e IV

55 - (ESAF/TTN) De acordo com a distribuição transcrita a seguir, pode-se afirmar que

Pesos(kg)	Frequências simples absolutas
2 ⊢→ 4	9
4 ⊢→ 6	12
6 ⊢→ 8	6
8 ⊢→ 10	2
10 ⊢→ 12	1

A mediana da distribuição é igual a

a) 5,20kg

b) 5,30kg

c) 5,00kg

d) Um valor inferior a 5kg

e) 5,10kg

56 - (ESAF/TTN) de acordo com a distribuição de frequências transcrita a seguir, pode-se afirmar que

Diâmetro(cm)	Frequência simples absoluta
4 ⊢→ 6	6
6 ⊢→ 8	8
8 ⊢→ 10	12
10 ⊢→ 12	10
12 ⊢→ 14	4

A mediana da distribuição

a) é eqüidistante da média aritmética e da moda

b) é igual à média aritmética

c) é inferior à média aritmética

d) coincide com o ponto médio do intervalo de classe

e) pertence a um intervalo de classe distinto do que contém a média aritmética

GABARITO

1- D	2-E	3- D	4- A	5- D	6- E	7- E	8- D	9- B	10- D	11- B
12- B	13 - D	14– B	15– E	16– E	17– C	18– A	19– E	20– C	21– D	22– C
23– B	24– A		26- A	27- C	28- A	29- C	30- D	31- C	32–D	33– A
34– C	35– A	36– B	37– 7	38– C	39– C	40– B	41– B	42– C	43– C	44– B
45– D	46– A	47– C	48- D	49– A	50- A	51-E	52- C	53- C	54- A	55- C
56- D										

BIBLIOGRAFIA

ALVAREZ-LEITE, Elvira Noções de Estatística – Faculdades Milton Campos. 2005 – Apostila

BANCO DO BRASIL Apostila de Matemática para escriturário do Banco do Brasil. www.achei**concursos**.com.br

TOLEDO, Geraldo Luciano. OVALLE, Ivo Izidoro. Estatística Básica.São Paulo:Atlas, 1985