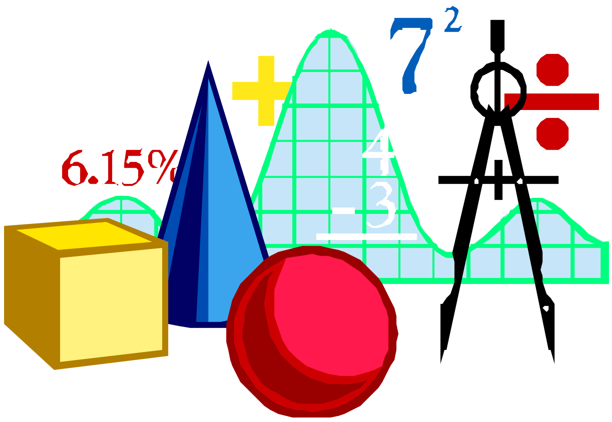
Estatística “I”



JAIR CROCE FILHO

MEMÓRIA

01. **Memória afiada. Não assobie e chupe cana:** bombardear a atenção com muitas informações juntas só atrapalha. Depois de alguns minutos, o cérebro acaba por dispersar. Não dá para ler e assistir à TV ao mesmo tempo. Você terá de decidir entre uma e outra ou perderá informações dos dois casos.

02. **Faça uma coisa de cada vez:** Evite estudar diversos conceitos novos em seqüência, tudo no mesmo dia. A mente precisa de tempo para absorver novas informações. Quando ainda não dominamos um assunto e tentamos passar para outro desconhecido, o cérebro não tem tempo para codificar tantas mensagens. Assim, corremos o risco de não memorizarmos nenhuma das duas.

03. **Evite tudo aquilo que lhe cause desconforto:** Estudar com fome, atrapalha a concentração. Ansiedade, depressão e tensão, também. Ao perceber que a atenção está pouca, tente descobrir o motivo e elimine-o, se o problema for difícil de resolver, peça ajuda médica, se for o caso.

04. **Estabeleça associações:** o modo mais simples de guardar conceitos recém-aprendidos é associá-los a algo que já sabemos. É mais fácil aprender uma novidade quando já temos conhecimento prévio do assunto, é como um quebra-cabeça: as primeiras peças são difíceis de encaixar mas, à medida que juntamos blocos, o processo de encaixe vai ficando rápido.

05. **Esquematize seus estudos:** Em matemática, estatística, português, etc. ao esquematizar o que aprendeu, você reorganiza o conhecimento com base em um conceito ou em um período da história. Isso significa trabalhar a mesma informação de forma diferente, o que ajuda a retê-la.

06. **Faça a retrospectiva do dia:** Antes de dormir, tente repassar seu dia de estudos. Que aulas você teve? Qual foi o conteúdo de cada uma delas? Parta do mais para o menos genérico, lembrando, por exemplo que parte ou capítulo da estatística foi estudado no dia, qual o conceito explicado, que tipo de exercício foi feito. Faça isso de forma leve, sem a preocupação de se recordar de tudo, porque as lembranças emergem como um novelo que se desenrola.

07. **Evite "decoreba":** repetir informações em bloco sem compreendê-las faz com que percamos todo o conjunto. A maneira ideal de reter um conceito é entendê-lo.

08. **Não esqueça o "corpore sano":** Exercícios físicos ajudam o corpo a liberar endorfina, o estimulante produzido pelo organismo que melhora a atenção. Atividade ao ar livre também colabora para diminuir a tensão, que é um dos fatores que mais atrapalham a concentração.

09. **Sistematize seu dia:** É o local onde guarda seus objetos. Guardá-los sempre no mesmo local não ajuda a memória, mas faz com que não precise acioná-la, por exemplo, para encontrar um lápis, reservando-a para coisas mais importantes.

10. **Elimine todas as drogas:** Incluindo álcool e medicamentos como tranqüilizantes e soníferos, a menos que sejam prescritos por médico. Tudo isso age no sistema nervoso central, podendo, a longo prazo, afetar sua memória; cocaína e maconha, por exemplo, deixam você mais esquecido.

Í N D I C E

S U M Á R I O

CAPÍTULO 1

Evolução Histórica

Todas as Ciências têm suas raízes na história do homem.

A Matemática, que é considerada "a ciência que une à clareza do raciocínio a síntese da linguagem", originou-se do convívio social, das trocas, da contagem, com caráter prático, utilitário, empírico.

A Estatística, ramo da Matemática Aplicada, teve origem semelhante.

1.1– Preparação dos Fatos

- Abrange a idade antiga, idade média e parte da idade contemporânea

- Característica principal - registro estatal - estatística administrativa

- Livro Sacro (**Chouking Vedas**) - escrito por Confúcio (**Sábio Chinês ± 2.230 a.C.**)
  - Censos de 2.275 e 2.238 a.C. (feito pelo Rei Yao)

Idade Antiga

- Vários povos já registravam o número de habitantes, de nascimentos, de óbitos, faziam estimativas das riquezas individual e social, distribuíam eqüitativamente terras ao povo, cobravam impostos e realizavam inquéritos qualitativos por processos que hoje, chamaríamos de "estatísticas".

- Período Romano

. *Census Romanus* realizado por Sérvio Túlio – 6º rei de Roma – 556 a.C.

. César Augusto convoca os Judeus para o recenseamento (**Bíblia - Nascimento de Jesus**)

Idade Média

- Colhiam-se informações, geralmente com finalidades tributárias ou bélicas.

- Havia coleta numérica de pessoas, cidades, fábricas e produtos alimentícios para controle das terras conquistadas

- Começam a surgir as primeiras análises sistemáticas de fatos sociais, como batizados, casamentos, funerais, originando as primeiras tábuas e tabelas e os primeiros números relativos.

- **Carlos Magno** - Rei dos Francos - Imperador do Ocidente - **800 d.C.**

. Com finalidades financeiras e administrativas estabeleceu o organismo do estado. Sua base foi a parte financeira.

- **Guilherme** - O Conquistador – **1083 a 1086 d.C.**

. *Doomsday Book* (Livro do Dia do Juízo) - organização de registros sistemáticos de informações e cadastros de interesse do Estado, com finalidade guerreira ou fiscal.

. Ordenou a elaboração de um cadastro dividindo o solo da Inglaterra entre várias classe sociais, para fins de arrecadação de impostos.

1.2– Preparação das Teorias

- No século XVI o estudo de fatos sociais foi adquirindo, aos poucos, feição verdadeiramente científica.

- As tabelas tornaram-se mais completas, surgiram as representações gráficas e o cálculo das probabilidades, e a Estatística deixou de ser simples catalogações de dados numéricos coletivos para se tornar o estudo de como chegar a conclusões sobre o todo (populações), partindo da observação de partes desse todo (amostra).

- Caracteriza-se pelo aparecimento de críticas polêmicas em torno da estatística como disciplina independente do estado.

<u>Hermann Conring</u> – (1600–1681) d.C. (Alemão)
- Primeiro a fazer distinção entre estatística e Estado
  - ✓ Estado = Governo = Território
  - ✓ Estatística = parte da matemática necessária na administração

<u>John Graunt</u> – (1620–1674) d.C. (Inglês) – Vendedor de tecidos de Londres.
- Inicia o trabalho de estatística demográfica. Publicou um estudo analítico sobre o registro de batismos, casamentos e enterros que há um século eram feitos, descobrindo certas regularidades nesses fatos.

<u>Willian Petty</u> – (1623–1687) d.C. (Inglês)
- Criador do termo “aritmética política” (fixar os fenômenos sociais por meio dos números e pelas relações numéricas entre os mesmos, dando-lhes características de precisão matemática com fundamentação teórica.
- Foi o primeiro a fazer conjeturas baseadas em informações estatísticas, utilizando tabelas e números relativos.

<u>Blaise Pascal</u> – (1623-1662) d.C. (Francês)
- Surgem os primeiros cálculos de probabilidade. (Chevalier de Méré). Dúvidas Pacioli.

<u>Edmund Halley</u> – (1656–1742) d.C. (Inglês) – Astrônomo e geômetra – Cont. Newton.
- Inicia o trabalho estatístico no campo social com o cálculo da mortalidade média de uma região. Ele notou que a morte, muito irregular e imprevisível para os casos considerados individualmente, seguia uma lei razoavelmente fixa se fosse computado um grande número de pessoas - daí se originando a primeira tábua de mortalidade. (Usada em seguros).

<u>John Peter Sussmilch</u> – (1707–1767) d.C. (Alemão) - Pastor
- Aritmético político, em sua obra “A ordem divina nas mutações do gênero humano” (1741), deu um aspecto verdadeiramente científico à estatística, procurando demonstrar relações entre causa e efeito no seu estudo. Demonstrou que o número de nascimentos e óbitos obedecia a certa “ordem divina” regular, que regia as manifestações naturais.

<u>Gottfried Achenwall</u> – (1719–1772) d.C. (Alemão)
- Batizou a nova ciência (ou método) com o nome de Estatística, determinando o seu objetivo e suas relações com as ciências. Portanto, o primeiro a utilizar a palavra Estatística.

<u>Lamber Adolph Jacques Quetelet</u> – (1796-1874) d.C. (Belga)
- Com sua obra “Física Social”, acabou por confirmar a regularidade com que os fatos demográficos, sociais e até morais se manifestam, aproximando a estatística do terreno matemático e como se desenvolve atualmente.

1.3– Aperfeiçoamento Técnico e Científico
- Inicia-se em 1853 com a reunião do “Primeiro Congresso de Estatística”. Dessa data até hoje o método estatístico vem sendo cada vez mais inteligentemente aplicado e em campos os mais diversos, e os processos de elaboração estatística mais aperfeiçoados.
- Um grande intercâmbio de informações e idéias, reuniões de congressos, unificação de pontos de vista, concepção da estatística como um método destinado a pesquisar as relações de causa e efeito dos fenômenos são características desse período, abrangendo até a idade contemporânea.
- Passou a ser usada em todos os campos de atividade humana (**Francis Galton**).

CAPÍTULO 2

Introdução

2.1– A Natureza da Estatística

A utilização da Estatística é cada vez mais acentuada em qualquer atividade profissional da vida moderna. Nos seus diversificados ramos de atuação, as pessoas estão freqüentemente expostas à Estatística, utilizando-a com maior ou menor intensidade. Isto se deve às múltiplas aplicações que o método proporciona àqueles que dele necessitam.

Geralmente, quando apresentamos uma citação estatística, somos levados, de pronto, a desacreditar em qualquer argumentação em contrário destituída de base numérica.

Devemos ser frontalmente contra os que afirmam que a ***"Estatística é capaz de provar qualquer coisa",*** o que implica, consequentemente, em querer dizer que a Estatística não prova coisa alguma.

Aquele que aceita dados estatísticos indiscriminadamente muitas vezes se deixará enganar, sem necessidade; também aquele que rejeita qualquer informe estatístico de pronto, estará dando prova de ignorância.

Atualmente, o público leigo (leitor de jornais e revistas) posiciona-se em dois extremos divergentes e igualmente errôneos quanto à validade das conclusões estatísticas: ou crê em sua infalibilidade ou afirma que elas nada provam. Os que assim pensam ignoram os objetivos, o campo e o rigor do método estatístico; ignoram a Estatística quer teórica quer prática, ou a conhecem muito superficialmente.

Há, evidentemente, a necessidade de especiais cuidados no manejo e na interpretação da Estatística; a interpretação não é monopólio dos estatísticos, sendo natural que, possuindo um maior conhecimento das técnicas estatísticas, levem vantagens no tocante à apreciação, análise e interpretação dos dados estatísticos. O raciocínio claro é indispensável para interpretar estatísticas, requerendo uma disposição mental receptiva e crítica.

Raramente, ou nunca, os dados estatísticos falam por si mesmos. A coisa mais importante acerca da interpretação dos dados estatísticos é saber que, se forem habilmente coletados e criticamente analisados podem ser extremamente úteis.

Infelizmente os maus empregos são tão numerosos quanto os usos válidos da Estatística. Ninguém - administrador, executivo, cientista ou pesquisador social deve deixar-se enganar pelas más Estatísticas, embora os casos de emprego indevido da Estatística sejam tantos que possam gerar a falsa impressão de que a Estatística é, raras vezes ou nunca, digna de confiança.

Como já dito, existem muitas concepções errôneas acerca da natureza desta disciplina. A idéia que um leigo possa fazer da Estatística difere em muito da de um profissional. É comum, por exemplo, as pessoas formarem conceitos distorcidos a respeito de um estatístico profissional. Para alguns, trata-se de um indivíduo que tem a capacidade de manipular números para demonstrar seus pontos de vista. Alguns estudantes, por outro lado, tendem a admiti-lo como alguém que, auxiliado por sua calculadora, tem a faculdade de converter qualquer assunto em um estudo "científico". Toda essa aura criada em torno da disciplina tem provocado, em estudantes e profissionais, uma dupla atitude: de apreensão, quanto à dificuldade de absorção de seu conteúdo, e de expectativa, quanto à sua potencialidade como instrumento auxiliar de resolução de problemas.

Na era da energia nuclear, os estudos estatísticos têm avançado rapidamente e, com seus processos e técnicas, têm contribuído para a organização dos negócios e recursos do mundo moderno.

Por essa razão, é extremamente difícil apresentar uma definição de Estatística, além do que muitos de seus conceitos fundamentais não apresentam uma definição explícita, ou, se a apresentam, esta não se revela suficientemente clara para dar uma idéia definitiva de seu significado.

Algumas definições:

a) é uma coleção de métodos para planejar experimentos, obter dados e organizá-los, resumi-los, analisá-los, interpretá-los e deles extrair conclusões;

b) é a <u>ciência</u> dos dados. Envolve a coleta, a classificação, o resumo, a organização, a análise e a interpretação da informação numérica;

c) é um conjunto de <u>métodos</u> e processos quantitativos que serve para estudar e medir os fenômenos coletivos;

d) é o <u>método</u> que estuda os fenômenos de massa;

e) é um <u>método</u> de trabalho que auxilia todas as outras <u>ciências</u>, no seu campo de ação;

f) porém, de um ponto de vista muito geral, <u>métodos</u> estatísticos são métodos de se tratar dados numéricos.

OBS.: Ciência - Conjunto de conhecimentos exatos racionais relativos as causas das realizações e suas deduções, portanto, ciência é tudo que tem objetivo fixo.

Ex: Matemática, Física, Química, etc.

Método - Conjunto de meios, processos e instrumentos usados pelos cientistas e estudiosos, para formularem seus princípios, teorias e normas.
Muitos dos conhecimentos que temos foram obtidos na antigüidade por acaso e, outros por necessidades práticas, sem aplicação de um método.
Atualmente, quase todo acréscimo de conhecimento resulta da observação e do estudo. Se bem que muito desse conhecimento possa ter sido observado inicialmente por acaso, a verdade é que desenvolvemos processos científicos para seu estudo e para adquirirmos tais conhecimentos.
- Método científico → observa
- Método experimental → faz experiência
- Método estatístico → relaciona fatos

2.2 – Origem da Palavra

Vem do Latim “Status” ou do Grego “Statizen” (ESTADO)

ABUSOS DA ESTATÍSTICA

Não é de hoje que ocorrem abusos com a estatística. Assim é que, a cerca de um século, o estadista *Benjamin Disraeli* disse:
“há três tipos de mentira: as mentiras, as mentiras sérias e a estatística.” Já se disse também que “os números não mentem; mas os mentirosos forjam números” e que “se torturarmos os dados por bastante tempo, eles acabarão por admitir qualquer coisa”.

O historiador *Andrew Lang* disse que algumas pessoas usam a estatística “como um bêbado utiliza um poste de iluminação – para servir de apoio e não para iluminar”.

Todas essas informações se referem aos abusos da estatística, quando os dados são apresentados de forma enganosa. Alguns dos que abusam da estatística o fazem simplesmente por descuido ou ignorância; outros porém, têm objetivos pessoais, pretendendo suprimir dados desfavoráveis enquanto dão ênfase aos dados que lhes são favoráveis.

OBSERVAÇÕES:

1ª) A ESTATÍSTICA NAS EMPRESAS

No mundo atual, a empresa e uma das vigas-mestras da Economia dos povos.

A direção de uma empresa, de qualquer tipo, incluindo as estatais e governamentais, exige de seu administrador a importante tarefa de tomar decisões, e o conhecimento e o uso da Estatística facilitarão seu tríplice trabalho de organizar, dirigir e controlar a empresa.

Por meio da sondagem, de coletas de dados e de recenseamento de opiniões, podemos conhecer a realidade geográfica e social, os recursos naturais, humanos e financeiros disponíveis, as expectativas da comunidade sobre a empresa, e estabelecer suas metas, seus objetivos com maior possibilidade de serem alcançados a curto, médio ou longo prazos.

A Estatística ajudará em tal trabalho, como também na seleção e organização da estratégica a ser adotada no empreendimento e, ainda, na escolha das técnicas de verificação e avaliação da quantidade e da qualidade do produto e mesmo dos possíveis lucros e/ou perdas.

Tudo isso que se pensou, que se planejou, precisa ficar registrado, documentado para evitar esquecimentos, a fim de garantir o bom uso do tempo, da energia e do material e, ainda, para um controle eficiente do trabalho.

O esquema do planejamento é o plano, que pode ser resumido, com auxílio da Estatística, em tabelas e gráficos, que facilitarão a compreensão visual dos cálculos matemático-estatístico que lhes deram origem.

O homem de hoje, em suas múltiplas atividades, lança mão de processos e técnicas estatísticas, e só estudando-os evitaremos o erro das generalizações apressadas a respeito de tabelas e gráficos apresentados em jornais, revistas e televisão, freqüentemente cometido quando se conhece apenas “por cima” um pouco de Estatística.

2ª) CAMPOS DE APLICAÇÃO

A Estatística encontra-se em quase todos os campos da atividade humana.
O Estado e a Sociologia têm necessidade de conhecer as populações por seus efetivos, por sexo, idade, estado civil, profissão, nacionalidade, etc.

Os serviços de meteorologia tão importantes para a navegação aérea e marítima, são essencialmente estatísticos, com seus estudos de temperaturas, pressões, quedas de chuvas, umidades, ventos, etc.

Na agricultura, a estatística serve como orientador seguro fornecendo informações sobre colheitas, rendimento das terras, valores da produção e outros.

Na indústria e no comércio podem-se comparar produções e volumes de vendas em relação ao total por região, estudar a situação dos mercados e suas tendências.

Grandes serviços a Estatística presta à Biologia desde o “homem médio” de Quetelet passando pela teoria da hereditariedade de Mendel, até as infinitas aplicações de hoje.

A Geografia conclui através de estudos estatísticos as densidades demográficas, correntes migratórias, clima, etc.

Na Informática também encontramos importantes aplicações, entre elas: avaliação de desempenho de redes de computadores, etc..

Na Inteligência Artificial, usam aplicações em redes neurais, artificiais e mineração de dados.

E ainda na História e Literatura, onde trabalhos estatísticos estudam a extensão dos períodos, coincidências, pontuações e estilos e, muitos outros.

3ª) RESUMO DA PROFISSÃO

O Estatístico promove levantamento de pesquisas estatísticas em suas aplicações técnicas e científicas, investigando, elaborando e testando métodos matemáticos e sistemas em amostragem, bem como coletando, analisando e interpretando os dados relacionados com os fenômenos estatísticos, e ainda estuda e renova a metodologia estatística a fim de estabelecer a sua evolução e desenvolvimento.

É fácil mentir com estatísticas

Antes de aceitar a verdade de um número, pergunte de onde veio e se não existe outro mais significativo

JORNAL DO BRASIL
12/03/97

Nunca fomos um povo com vocação científica. Não temos base, nem paciência, nem muito interesse pelos números. Num exame vestibular, feito a algum tempo, por exemplo, a média das notas das provas de matemática dos candidatos ao curso de administração foi 1,4. Como se trata de média, pode-se imaginar o número de zeros.

Por culpa do ensino, dos maus professores ou da nossa formação cultural, o fato é que a maioria das pessoas, no Brasil, tem dificuldades permanentes com a matemática, desde o curso primário até a universidade.

No entanto, somos diariamente bombardeados pela mídia com informações estatísticas (que são um ramo da matemática): o sabão em pó X lava mais branco (do que o quê?). O candidato B lidera as pesquisas de opinião (qual é a amostra? como foram feitas as perguntas?). Cai o poder aquisitivo da classe média (de quanto? onde? por quê? classe “média” na classificação de quem?). Sobe o poder aquisitivo da classe média. Os juros são de 500% ao ano (mas a taxa usada foi ao mês). Aumenta o rombo da previdência. As contas da previdência foram saneadas. Qual é a maior ameaça à saúde pública: a AIDS (mil casos verificados) ou a doença de Chagas (vários milhões de doentes?).

Foram provas dessa incompetência matemática as grandes decepções com os muitos planos de estabilidade da economia brasileira, porque aceitávamos, sem questionamento, planos de estabilidade econômica que simplesmente congelavam o resultado de uma complexa equação, sem que fossem tomadas as medidas necessárias para controlar cada um dos termos que a compunha, como os gastos governamentais e as emissões de moeda, por exemplo.

A informação de que há 10 milhões de menores abandonados no Brasil é freqüentemente mencionada como um dado real. Mas a composição etária da população, cruzada com a estratificação sócio-econômica da ABA/ABIPEME indica que o total de crianças com menos de 15 anos pertencente às classes D e E é de cerca de 17 milhões, o que indica que cerca de 60% deles são abandonados, ou então que há menores abandonados na classe média (o que é, no mínimo, inverossímil). Além disso, o que é um menor abandonado? Certamente um menino ou menina de 9 a 10 anos que perambula pelas ruas da cidade sem ter para onde ir: mas o mesmo não se pode afirmar de um jovem de 15 anos, que pode ter saído de casa por vontade própria. Existe a informação que mais de 2/3 dos jovens que perambulam pelas ruas da cidade do Rio durante o dia voltam para suas casas à noite... o que, aliás, não significa que o problema dos menores não seja motivo de sérias preocupações de todos nós.

Outro exemplo marcante é a renda per capita, ou salário médio. São mais de 2000 dólares, no Brasil. Existem, contudo, três taxas de câmbio diferentes: o comercial, o de turismo e o paralelo. Qual o que vale? E deve-se dividir o total anual por 12 meses? Ou por 13 salários?

Para complicar, existem, em estatística, três tipos de média: a média aritmética, que é comumente usada, soma todos os salários e divide pelo número de pessoas. Assim, a média de ganho de duas pessoas: um banqueiro que ganha 20 mil reais por mês e de um servente de pedreiro com salário de 200 reais é de 10 mil e 100 reais para cada um, divididos na tabela estatística, mas não na vida real. Mas existe uma outra média, denominada mediana, que estabelece o nível salarial mediano, isto é, um certo salário, onde há tantas pessoas que ganham mais do que aquele número, quantas pessoas que ganham menos. Tal cálculo (bastante democrático) não é feito no Brasil. Finalmente, existe a moda ou modal, que é o salário de maior incidência (ou freqüência) na população e, que no Brasil, é indiscutivelmente o salário mínimo, hoje de pouco mais de R$100... Qual dos três cálculos de conversão e qual das três médias devem ser utilizadas para os estudos sociais e os planos econômicos?

Resta o consolo de que a confusão matemática e estatística não é privilégio do Brasil. Faz pouco tempo, uma revista americana publicou um artigo sobre acidentes de trânsito nos Estados Unidos, comentando que a maioria dos acidentes com vítimas era causada por motoristas com menos de 30 anos e que, portanto, as pessoas mais jovens eram mais imprudentes ao dirigir. Um professor de Estatística escreveu à revista para observar que o simples registro numérico dos acidentes não era suficiente para apoiar a conclusão do editor. Para fazer a inferência correta (explicava) é necessário conhecer o número de motoristas habilitados com menos de 30 anos, em relação ao total de motoristas de todas as idades: determinar o número de horas por ano que cada faixa etária dedica a dirigir e os tipos de acidentes, para estabelecer suas reais causas. Afinal, ponderava, além do provável fato de que as pessoas mais jovens dirigem com mais freqüência, um motorista idoso conduzindo um carro a cinqüenta por hora numa auto-estrada pode ser o verdadeiro causador de um acidente, ainda que, tecnicamente, o culpado seja quem bate...

É claro que esse exemplo não significa que as pessoas mais jovens não possam ser mais imprudentes e provocar, proporcionalmente, mais acidentes que os mais velhos. Apenas serve para ilustrar o fato de que as afirmações feitas em cima de estatísticas frágeis tendem a merecer mais crédito quando estão de acordo com as opiniões e preceitos de quem as ouve.

Para se conhecer a verdade através dos números, não basta que eles existam e não sejam falsos. É preciso saber de onde vieram, quais os disfarces que podem estar usando e se não existem outros números mais importantes que estejam sendo (talvez deliberadamente) escondidos ou omitidos por quem os apresenta. Especialmente se for para "provar" alguma coisa ou tentar convencer alguém.

J. ROBERTO WHITAKER PENTEADO - Vice-presidente da mantenedora e diretor-geral da Escola Superior de Propaganda e Marketing do Rio de Janeiro.

2.3– Estatística Descritiva e Estatística Indutiva

2.3.1 – Estatística Descritiva ou Dedutiva

É a parte da estatística referente a coleta e tabulação de dados.

Utiliza métodos numéricos e gráficos para mostrar os padrões de comportamento dos dados, para resumir a informação contida nesses dados e para apresentar a informação de forma conveniente.

É aquela que tem por objetivo descrever e analisar determinada população, sem pretender tirar conclusões de caráter mais genérico.

Dado um conjunto de elementos, podemos em relação a um certo fenômeno estudar todos os seus elementos, classificando-os, fornecendo números indicativos que sumariem, certas características dos dados; são números sumariantes, que fornecem descrições de todo o conjunto sem a apresentação total dos elementos, ou mesmo medidas e relações do conjunto, não perceptíveis com a pura apresentação do rol de dados. A este setor da Estatística denomina-se ***Estatística Descritiva.***

Principalmente em pesquisa social, o analista defronta-se amiúde com a situação de dispor de tantos dados que se toma difícil absorver completamente a informação que está procurando investigar. É extremamente difícil captar intuitivamente todas as informações que os dados contêm. É necessário, portanto, que as informações sejam reduzidas até o ponto em que se possa interpretá-las mais claramente. Em outras palavras, é indispensável resumi-las, através do uso de certas ***medidas-sínteses***, mais comumente conhecidas como **estatísticas descritivas** ou simplesmente estatísticas. Por conseguinte, a estatística descritiva é um número que sozinho descreve uma característica de um conjunto de dados. Trata-se, portanto, de um ***número-resumo*** que possibilita reduzir os dados a proporções mais facilmente interpretáveis. Evidentemente, ao resumir os dados através do uso de estatísticas descritivas, muita informação irá necessariamente se perder, além de ser provável a obtenção de resultados distorcidos, a menos que eles sejam interpretados com muita precaução.

Em um sentido mais amplo, a Estatística Descritiva pode ser interpretada como uma função cujo objetivo é a observação de fenômenos de mesma natureza, a coleta de dados numéricos referentes a esses fenômenos, a organização e a classificação desses dados observados e a sua apresentação através de gráficos e tabelas, além do cálculo de coeficientes (estatísticas) que permitem descrever resumidamente os fenômenos.

2.3.2– Estatística Indutiva ou Inferência Estatística (indução, conseqüência, conclusão)

É a parte da estatística que, baseando-se em resultados obtidos da análise de uma amostra da população, procura inferir, induzir ou estimar as leis de comportamento da população da qual a amostra foi retirada.

Portanto, a **estatística indutiva** refere-se a um processo de generalização, a partir de resultados particulares. Consiste em obter e generalizar conclusões, ou seja, inferir propriedades para o todo com base na parte, no particular. A inferência estatística implica, pois, um raciocínio muito mais complexo do que o que preside a Estatística Descritiva. Entretanto, bem compreendida e utilizada, pode converter-se em um instrumento muito importante para o desenvolvimento de uma disciplina científica.

O processo de generalização, que é característico do ***método indutivo***, está associado a uma margem de incerteza. A existência da incerteza deve-se ao fato de que a conclusão, que se pretende obter para o conjunto de todos os indivíduos analisados quanto a determinadas características comuns, baseia-se em uma parcela do total de observações. A medida da incerteza é tratada mediante técnicas e métodos que se fundamentam na Teoria da Probabilidade.

2.4– População, Amostra e Amostragem

2.4.1– População ou Universo Estatístico

O conjunto da totalidade dos indivíduos sobre o qual se faz uma inferência recebe o nome de população ou universo. A população congrega todas as observações que sejam relevantes para o estudo de uma ou mais características dos indivíduos, os quais podem ser concebidos tanto como seres animados ou inanimados. Em linguagem mais formal, a população é o conjunto constituído por todos os indivíduos que apresentem pelo menos uma característica comum, cujo comportamento interessa analisar (inferir).

Pode-se classificar pelo tamanho, sendo finita, quando a população possui um número determinado de elementos e infinita, quando possui um número infinito de indivíduos, mas tal definição só existe na teoria, pois na prática, nunca encontraremos populações com infinitos elementos, mas sim, populações com um grande número de componentes, por isso, tais populações são tratadas como se fossem infinitas. Quanto maior a população, mais difícil a observação dos aspectos que se deseja estudar, devido ao alto custo, ao intenso trabalho e ao tempo gasto para tal.

Assim sendo, o objetivo das generalizações estatísticas (indução estatística) está em dizer-se algo acerca de diversas características da população estudada, com base em fatos conhecidos. Essas características da população são comumente chamadas de parâmetros (valores singulares), os quais são valores fixos e ordinariamente desconhecidos.

2.4.2– Amostra

A amostra pode ser definida como um subconjunto, uma parte selecionada da totalidade de observações abrangidas pela população, através da qual se faz um juízo ou inferência sobre as características da população. As características da amostra são chamadas de estatísticas (descritivas), sendo simbolizadas por caracteres latinos, enquanto que os parâmetros da população terão como símbolos, via de regra, os caracteres gregos.

“A amostra é um subconjunto, representativo ou não, da população em estudo. Essa represernta-tividade da amostra, ocorre quando ela apresenta as mesmas características gerais da população da qual foi extraída” (Milone, Angelini, 1993:16)

Ex.: Suponha-se, que se pretenda conhecer o conteúdo de ferro natural a ser exportado por um navio. O agregado ou população consiste em todo o minério de ferro a ser exportado por esse navio. Parte do minério é examinada, a fim de determinar seu teor de ferro, com o objetivo de tirar uma conclusão a respeito do teor de ferro natural do embarque completo. A parte de mineral selecionado constitui a amostra do embarque. Uma vez que se fará inferência sobre todo o minério embarcado a partir de apenas uma porção dele, a base do processo é a informação incompleta ou de amostra.

IMPORTANTE

É possível mentir usando estatísticas, mas se mente mais, e melhor, sem estatísticas. É preciso entender que as amostras podem levar a conclusões erradas. Contudo, as opiniões pessoais, sem base em dados, levam, em geral, a conclusões muito mais erradas.

Frederick Mosteller
Professor em Harvard

2.4.3– Amostragem

É um artifício ou uma técnica estatística que possibilita realizar a pesquisa em universos infinitos, quanto aos aspectos de custo e de tempo. Desta forma, a Estatística pode ser estendida ao estudo das populações chamadas “infinitas” nas quais não temos a possibilidade de observar todos os elementos do universo.

Mesmo no caso das populações finitas passou-se a empregar o estudo por amostragem, pela economia e rapidez dos resultados. Assim o estudo da qualidade dos produtos de uma partida industrial passou a ser feito a partir dos resultados obtidos pela inspeção dos elementos de uma amostra.

A teoria da amostragem é útil para determinar se as diferenças observadas entre duas amostras são realmente devidas a uma variação casual ou se são verdadeiras.

No geral, ao estudo da inferência de uma pesquisa a respeito de uma população mediante a utilização de amostras delas extraídas, junto com a precisão das inferências usando a teoria da probabilidade, denominamos *inferência estatística*.

Enfim, amostragem é o estudo das relações existentes entre a população e as amostras dela extraídas. É o conjunto de técnicas utilizadas para a seleção de uma amostra.

Este conjunto de técnicas pode ser subdividido em dois grupos básicos: a amostragem não aleatória e a amostragem aleatória.

A amostragem não aleatória incluí técnicas como:

a) Amostragem intencional

Ocorre quando o pesquisador seleciona intencionalmente os componentes da amostra.

b) Amostragem voluntária

Ocorre quando o componente da população se oferece voluntariamente para participar da amostra independentemente do julgamento do pesquisador.

Estas amostras não permitem o controle da variabilidade amostral, o que inviabiliza o controle da qualidade da estimação.

A amostragem aleatória incluí técnicas como:

1) Amostragem casual ou aleatória simples

Este tipo de amostragem é equivalente a um sorteio lotérico. Ela pode ser realizada numerando-se a população de 1 a n e sorteando por meio de um dispositivo aleatório qualquer, **K** números dessa população que corresponderá aos elementos pertencentes à amostra.

Ex.: obter uma amostra representativa de 10% para a pesquisa de estatura de 70 alunos:

- numeramos os alunos de 01 a 70;
- escrevem-se os números de 01 a 70 em pedaços iguais de papéis, coloque-os dentro de uma caixa e proceda ao sorteio, tirando um a um, sete números que formarão a amostra.

Quando o número de elementos da amostra é muito grande, nestes casos utiliza-se uma Tabela de Números Aleatórios, constituída de modo que os algarismos são distribuídos ao acaso nas linhas e colunas. Para obtermos os elementos da amostra usando a tabela, sorteamos um algarismo, a partir do qual iremos considerar números de dois, três ou mais algarismos, conforme a necessidade.

Medindo as alturas dos alunos correspondentes aos números sorteados, obteremos uma amostra das estaturas dos 70 alunos.

2) Amostragem proporcional ou estratificada

É quando a população se subdivide em sub-populações (estratos)

Quando a população se divide em estratos, convém que o sorteio dos elementos da amostra leve em consideração tais estratos, daí obtém-se os elementos da amostra proporcional ao número de elementos desse estrato.

Ex.: Se considerarmos o exemplo anterior, que, dos 70 alunos, 40 sejam meninas e 30 meninos, vamos obter amostra proporcional estratificada.

Temos dois estratos (sexo masculino e feminino) se queremos uma amostra de 10%, teremos 3 homens (10% de 30) e 4 mulheres (10% de 40).

3) Amostragem sistemática

Quando os elementos da população já se encontram ordenados, não há necessidade construir um sistema de referência.

Esta amostragem é semelhante à aleatória simples, mas a listagem é ordenada. Devemos seguir os seguintes passos:

1º) divide-se o tamanho da população (N) pelo tamanho da amostra (n), obtendo um intervalo de retirada (k).
2º) sorteia-se o ponto de partida.
3º) a cada k elementos retira-se uma para amostra.

Ex.: - no caso de uma linha de produção, podemos a cada 10 itens produzidos, retirar 01 para amostra de produção diária. Neste caso estaríamos fixando o tamanho da amostra em 10% da população.

- uma avenida contendo 800 prédios, dos quais desejamos obter uma amostra formada de quarenta prédios. Pode-se usar o seguinte procedimento: 800 : 40 = 20, escolhemos por um sorteio casual um número de 1 a 20, o qual indicaria o primeiro elemento sorteado para a amostra; os demais elementos serão periodicamente considerados de 20 em 20. Se o número sorteado fosse 06, por exemplo, tomaríamos pelo lado direito da avenida, o 6º prédio, o 26º, o 46º, o 66º, etc., até voltarmos ao início da rua, pelo lado esquerdo.

4) Amostragem por conglomerados

Em algumas situações, podemos identificar um grupo de elementos que tenha aproximadamente a mesma composição de população. Neste caso, pode ser interessante realizar a amostragem usando somente os elementos desse grupo.

Algumas empresas, quando pretendem avaliar a aceitação de um produto no eixo Rio-São Paulo, lançam o produto em Curitiba, cuja população se comporta com uma miniatura desse mercado.

Ex.: Dependendo da situação poderemos consultar todos os moradores de um único prédio para conhecer o pensamento de todo o bairro.

OBSERVAÇÕES:

1ª) PARÂMETROS: são valores singulares que existem na população e que servem para caracterizá-la. Para definirmos um parâmetro devemos examinar toda a população.

Ex.: Os alunos do 2º ano de uma Faculdade têm em média 1,68 metros de estatura.

2ª) ESTIMATIVA: é um valor aproximado do parâmetro e é calculado com o uso da amostra.

3ª) ATRIBUTO: quando os dados estatísticos apresentam um caráter qualitativo, o levantamento e os estudos necessários ao tratamento desses dados são designados genericamente de estatística de atributo.

Ex.: de classificação dicotômica do atributo:
a classificação dos alunos da Faculdade quanto ao sexo.
Atributo: sexo
Classe: alunos da faculdade
Dicotomia: duas subclasses (Masculino e Feminino)

de classificação policotômica do atributo:
alunos da Faculdade quanto ao estado civil.
Atributo: estado civil
Classe: alunos da faculdade
Dicotomia: mais de duas subclasses (solteiro, casado, divorciado, viúvo, etc.)

4ª) VARIÁVEL: é convencionalmente, o conjunto de resultados possível de um fenômeno.

2.5– Fenômenos Estatísticos

O fenômeno em estatística relaciona-se com qualquer evento que se pretenda analisar, cujo estudo seja passível da aplicação da técnica estatística. A Estatística dedica-se ao estudo dos fenômenos de massa, que são resultantes do concurso de um grande número de causas, total ou parcialmente desconhecidas, que serão chamadas de **"fenômenos estatísticos"**.

É possível não se conhecerem exatamente as causas subjacentes aos fenômenos, pois pode-se estudá-los através de suas manifestações, descobrindo-se neles alguns aspectos globais, sem remontar a essas causas. O que caracteriza tais fenômenos (sociais, biológicos, etc.) é o fato de serem eles provenientes de um concurso de causas nem sempre totalmente conhecidas pelo analista. Os fenômenos classificam-se em três tipos:

2.5.1– Fenômenos Coletivos ou Fenômenos de Massa

Os fenômenos coletivos são aqueles que não podem ser definidos por uma simples observação. A natalidade, a mortalidade, a nupcialidade, o preço médio de veículos usados, vendidos diariamente em uma grande cidade, são fenômenos coletivos.

2.5.2– Fenômenos individuais ou Particulares

Os fenômenos individuais são aqueles que irão compor os fenômenos coletivos. Cada nascimento, cada indivíduo que morre, cada casamento que ocorre, cada veículo usado que se vende diariamente em uma grande cidade, são fenômenos individuais.

2.5.3– Fenômenos de Multidão

Os fenômenos de multidão distinguem-se dos fenômenos coletivos pelo fato de as características observadas para a massa não se verificarem para o particular, para o indivíduo isoladamente.

De acordo com a forma como se manifestam, os fenômenos podem ser classificados sob dois aspectos:

a) *Fenômenos Típicos*
Os fenômenos típicos são aqueles que se manifestam de forma regular, revelando um comportamento definido.

Ex.: eleição, dia e noite, marés, censo, balanço, balancete, etc..

b) *Fenômenos Atípicos*
Os fenômenos atípicos referem-se àqueles fenômenos cuja manifestação se dá através de um comportamento irregular, não revelando uma tendência definida.

Ex.: epidemias, avalanches, chuva, etc..

2.6– Aspecto Qualitativo e Aspecto Quantitativo (Variável Discreta e Contínua)

2.6.1– Aspecto Qualitativo (Variável Categórica)

É o que representa qualidade, atributo, característica, etc.. Considera-se um caráter como qualitativo quando as modalidades que o compõem formam um conjunto amorfo (sem forma definida), não estruturado numericamente, ou seja, quando não há ligação entre essas modalidades, independentemente do fato de constituírem um conjunto completo. Se subdividem em:

A) Nominais - Ex.: sexo, cor, raça, etc.

B) Por Postos - Ex.: lista de concursos, tabela de campeonato, etc.

2.6.2– Aspecto Quantitativo (Variável Numérica)

Os resultados das observações serão expressos sempre através de *valores numéricos*. Os dados são de caráter nitidamente quantitativo, e o conjunto dos resultados possui uma estrutura numérica. Dir-se-á, então, que se trata de estatística quantitativa ou estatística de variável. Se subdividem em:

A)Variável Contínua: é a que permite subdivisões intermediárias entre dois pontos, ou seja, pode assumir qualquer valor num certo intervalo de medida. (Conjunto R dos números naturais).

Ex.: um pacote de arroz, diâmetro de um rolamento, mercúrio no termômetro, etc.

B) Variável Descontínua ou Discreta: é a que não permite valores intermediários entre dois pontos, ou seja, só pode assumir determinado valor num certo intervalo de medida. (Valores inteiros, inclusive zero).

x.: uma pessoas, um carro, uma consulta, um computador, etc.

CAPÍTULO 3

Fases do Método Estatístico (Estatística Descritiva)

Quando se pretende empreender um estudo estatístico completo, existem diversas fases do trabalho que devem ser desenvolvidas para se chegar aos resultados finais do estudo. Essas etapas ou operações são chamadas fases do trabalho estatístico.

3.1– Definição do Problema

A primeira fase do trabalho estatístico consiste em uma definição ou formulação correta do problema a ser estudado. Além de considerar detidamente o problema objeto do estudo, o analista deverá examinar outros levantamentos realizados no mesmo campo e análogos, uma vez que parte da informação de que necessita pode, muitas vezes, ser encontrada nesses últimos.

Ex.: Um fabricante de sabonete, que deseja lançar um produto novo no mercado, poderia estar interessado em um estudo sobre as características dos consumidores atuais. Não havendo estudos semelhantes, ele deverá formular o problema com base em sua própria experiência. Uma lista de fatores relevantes deverá resultar dessa investigação preliminar:

- número de unidades consumidas por família em cada ano;
- número médio de pessoas que compõe cada família;
- número de membros adultos da família, as marcas preferidas e assim por diante.

Saber exatamente aquilo que se pretende pesquisar é o mesmo que definir corretamente o problema.

3.2– Planejamento

O passo seguinte, após a definição do problema, compreende a fase do planejamento, que consiste em se determinar o procedimento necessário para resolver o problema e, em especial, como levantar informações sobre o assunto objeto do estudo.

- Que dados deverão ser obtidos?
- Como obtê-los?
- O que será pesquisado?
- Quem participará da pesquisa? (Critérios de inclusão e exclusão)
- Em que setores geográficos será feita a pesquisa?
- Qual o grau de precisão exigido na pesquisa?
- Qual o tipo de amostragem?
- Qual o tamanho da amostra?
- Quais materiais serão necessários para realizar a pesquisa?
- Qual o tempo disponível para fazer a pesquisa?
- Qual o custo previsto?
- Qual a verba destinada ao projeto? Etc.

É preciso *planejar* o trabalho a ser realizado, tendo em vista o objetivo que se pretende atingir. Mais especificamente, na fase do planejamento a preocupação maior reside na escolha das perguntas, bem como sua correta formulação, qualquer que seja a modalidade de coleta dos dados.

O planejamento pode ser dividido em: **censitário** que é utilizado quando a contagem for completa, ou por **amostragem** quando for parcial.

Ex.: Censitário → levantamento do IBGE
Amostragem → opinião dos eleitores sobre o presidente.

ATENÇÃO:
um mal planejamento pode comprometer, não só as diversas fases da análise mas também a própria análise, levando a conclusões ou a decisões erradas, podendo vir a prejudicar os resultados do negócio como um todo (processo, empresa, etc.).

3.3– Coleta dos Dados

O terceiro passo é essencialmente operacional, compreendendo a coleta das informações propriamente ditas. Formalmente, a coleta de dados se refere à obtenção, reunião e registro sistemático de dados, com um objetivo determinado.
Antes de se tecer qualquer outra consideração sobre esta fase do método estatístico, convém estabelecer uma distinção entre os dados estatísticos.

3.3.1– Origem dos Dados

A) Dados Primários
Os dados são primários quando são publicados ou comunicados pela própria pessoa ou organização que os haja recolhido.

B) Dados Secundários
Os dados são secundários quando são publicados ou comunicados por outra organização.

OBS.: Um conjunto de dados é, pois, primário ou secundário em relação a alguém. As tabelas do Censo Demográfico são fontes primárias. Quando determinado jornal publica estatísticas extraídas de várias fontes e relacionadas com diversos setores industriais, os dados são secundários para quem desejar utilizar-se deles em alguma pesquisa que esteja desenvolvendo. Embora muitas vezes possa ser conveniente recorrer a fontes secundárias, é mais seguro trabalhar com fontes primárias, por várias razões:

1. uma fonte primária oferece, em geral, informação mais detalhada do que uma fonte secundária;
2. é mais provável que as definições de termos e de unidades figurem somente nas fontes primárias;
3. o uso da fonte secundária traz o risco adicional de erros de transcrição;
4. uma fonte primária poderá vir acompanhada de cópias dos impressos utilizados para coletar as informações, juntamente com o procedimento adotado na pesquisa, a metodologia seguida e o tipo e tamanho da amostra.

Essas informações proporcionam ao usuário uma idéia do grau de garantia que os dados oferecem.

3.3.2– Como os Dados são Encontrados na Natureza (em quantidade)

A) Enumerados (Enumeração)
São aqueles que podem ser contados ou contabilizados.
Ex.: contagem física de objetos, levantamento de estoque, etc.

B) Mensurados (Mensuração)
São aqueles que podem ser obtidos através do uso de instrumentos de medida.
Ex.: balança, metro, litro, etc.

C) Avaliados (Avaliação)

São aqueles que não podem ser contados ou mensurados. É uma maneira empírica de se determinar quantidade.

Ex.: uma sala com $\pm$ 100 pessoas.

3.3.3– Tipos de Coleta de Dados

A) Coleta Direta

A coleta é direta quando é obtida diretamente da fonte, como no caso da empresa que realiza uma pesquisa para saber a preferência dos consumidores pela sua marca.

A coleta direta de dados pode ser classificada relativamente ao fator tempo em contínua, periódica e ocasional.

1ª - Coleta Contínua ou Automática

A coleta de dados é contínua quando estes são obtidos ininterruptamente, automaticamente e na vigência de um determinado período: um ano, por exemplo. É aquela em que é feito o registro tão logo se verifique o fato.

Ex.: Registros de nascimento, de casamento, de óbito, baixa automática de estoque, etc..

2ª - Coleta Periódica

A coleta de dados é periódica quando é realizada em períodos determinados, de tempos em tempos, com repetições cíclicas.

Ex.: recenseamento a cada dez anos, o censo industrial, anualmente, balanço, etc..

3ª - Coleta Ocasional

A coleta de dados é ocasional quando os dados forem colhidos esporadicamente, ocasionalmente, atendendo a uma conjuntura qualquer ou a uma emergência.

Ex.: coleta de casos fatais em um surto epidêmico, registro de pedidos de um determinado artigo que uma grande empresa recebe em um dia de greve, etc..

B) Coleta Indireta

A coleta dos dados é indireta quando é inferida **a partir dos elementos** conseguidos pela **coleta direta**, ou através do conhecimento de outros fenômenos que, de algum modo, estejam relacionados com o fenômeno em questão. É feita, portanto, por deduções e conjeturas (sem fundamento preciso, suposição), podendo ser realizada:

1ª - Coleta por Analogia

A coleta de dados é feita por analogia quando o conhecimento de um fenômeno é induzido a partir de outro que com ele guarda relações de casualidade.

Ex.: No carnaval do ano passado 10.000 pessoas visitaram a cidade, espera-se a mesma quantidade este ano.

2ª - Coleta por Proporcionalização

A coleta de dados é feita por proporcionalização, quando o conhecimento de um fato se induz das condições quantitativas de uma parte dele. É feito através de uma regra de três, em que se mede um elemento básico. Nada mais é que uma porcentagem.

Ex.: uma peça ocupa um espaço de $3m^2$, tenho um espaço de $30m^2$, portanto coloco 10 peças neste espaço.

3ª - Coleta por Indícios

A coleta por indícios se dá quando são escolhidos fenômenos sintomáticos para discutir um aspecto geral da vida social.

Ex.: reunião de elementos como prova de um crime para a descoberta dos culpados

4ª - Coleta por Avaliação

A coleta é feita por avaliação quando, através de informações fidedignas ou estimativas cadastrais, se presume o estado quantitativo de um fenômeno.

Ex.: supor que existam 150 pessoas numa sala.

<u>Resumo da Coleta de Dados</u>

- Coleta Direta	- Coleta Contínua ou Automática	→	Ex.: Registro de óbitos
	- Coleta Periódica	→	Ex.: Censo
	- Coleta Ocasional	→	Ex.: Epidemias
- Coleta Indireta	- Coleta por Analogia	→	Ex.: Pessoas no carnaval
	- Coleta por Proporcionalização	→	Ex.: Regra de três
	- Coleta por Indícios	→	Ex.: Elementos de um crime
	- Coleta por Avaliação	→	Ex.: Produção de um período

3.4– Apuração dos Dados

Antes de começar a analisar os dados, é conveniente que lhes seja dado algum tratamento prévio, a fim de torná-los mais expressivos.

A quarta etapa do processo é, então, a da apuração ou sumarização, que consiste em resumir os dados, através de sua contagem e agrupamento. É um trabalho de condensação e de tabulação dos dados, que chegam ao analista de forma desorganizada, tornando impossível a tarefa de apreender todo o seu significado pela simples leitura.

Há várias formas de se fazer a apuração, dependendo das necessidades e dos recursos disponíveis do interessado: manual (sem máquinas), mecânica (máquinas de somar manual), eletromecânica (máquinas de somar elétricas) ou eletrônica (computador).

Por conseguinte, através da apuração, tem-se a oportunidade de condensar os dados, de modo a obter um conjunto compacto de números, o qual possibilita distinguir melhor o comportamento do fenômeno na sua totalidade.

Entretanto, a contrapartida da melhor apreciação dos dados em seu conjunto é a perda correspondente de detalhes, uma vez que se trata de um processo de sintetização.

3.5– Crítica dos Dados

Obtidos os dados, eles devem ser cuidadosamente criticados, à procura de possíveis falhas e imperfeições, a fim de não incorrermos em erros grosseiros ou de certo vulto, que possam influir sensivelmente nos resultados. As críticas podem ser:

3.5.1– Crítica Interna

É a crítica feita sobre os dados originais da coleta.
Ex.: soma de números.

3.5.2– Crítica Externa

É aquela que visa a causa dos erros por parte do informante, por distração ou má interpretação das perguntas que foram feitas.
Ex.: perguntas mal formuladas ou indiscretas, respostas com duplo sentido, etc.

3.6– Apresentação dos Dados

Por mais diversa que seja a finalidade que se tenha em vista, os dados devem ser apresentados sob forma adequada, tornando mais fácil o exame daquilo que está sendo objeto de tratamento estatístico e ulterior obtenção de medidas típicas.

Há duas formas de apresentação dos dados, que não se excluem mutuamente:

3.6.1– Apresentação Tabular

A apresentação tabular é uma apresentação numérica dos dados. Consiste em dispor os dados em linhas e colunas distribuídas de modo ordenado, segundo algumas regras práticas adotadas pelos diversos sistemas estatísticos.

De maneira mais formal, define-se como tabela, a disposição escrita que se obtém, fazendo-se referir uma coleção de dados numéricos a uma determinada ordem de classificação.

3.6.2– Apresentação Gráfica

A apresentação gráfica dos dados numéricos constitui uma apresentação geométrica. Embora a apresentação **tabular** seja de extrema importância, no sentido de facilitar a análise numérica dos dados, não permite ao analista obter uma visão tão rápida, fácil e clara do fenômeno e sua variação como a conseguida através de um gráfico.

3.7– Análise e Interpretação dos Dados

A última fase do trabalho estatístico é a mais importante e também a mais delicada. Nesta etapa, o interesse maior reside em tirar conclusões que auxiliem o pesquisador a resolver seu problema. A análise dos dados estatísticos está ligada essencialmente ao cálculo de medidas, cuja finalidade principal é descrever o fenômeno. Assim, o conjunto de dados a ser analisado pode ser expresso por números-resumos, as estatísticas, que evidenciam características particulares desse conjunto. O significado exato de cada um dos valores obtidos através do cálculo das várias medidas estatísticas disponíveis deve ser bem interpretado. É possível mesmo, nesta fase, arriscar algumas generalizações, as quais envolverão, naturalmente, algum grau de incerteza, porque não se pode estar seguro de que o que foi constatado para aquele conjunto de dados (a amostra) se verificará igualmente para a população.

CAPÍTULO 4

Séries e Tabelas Estatísticas

4.1– Série Estatística

Uma vez coletados os dados, não é conveniente apresentá-los para análise, sob a forma a que se chegou pela simples apuração. Muitas vezes o conjunto de valores é extenso e desorganizado, e seu exame requer maior atenção. Além disso, como já foi salientado, há o perigo de se perder a visão global do fenômeno analisado, quando a lista de dados for extensa e desordenada.

Por outro lado, se a lista original de valores puder ser apresentada de uma forma mais simples e compacta, haverá menor dificuldade em interpretar os dados e trabalhar com eles. Reunindo, pois, os valores em tabelas compactas, consegue-se apresentá-los e descrever-lhes a variação mais eficientemente. Essa condensação dos valores permite ainda a utilização de representação gráfica, que normalmente representa uma forma mais útil e elegante de apresentação da característica analisada.

Enfim, qualquer processo de representação que contribua para proporcionar uma visão mais sintética do fenômeno estudado, sem tirar-lhe a precisão primitiva, contribuirá igualmente para facilitar e encaminhar qualquer desses estudos, quer seja o de caracterização de um conjunto, o de comparação com outros semelhantes ou ainda o de previsão de valores possíveis. É o caso, por exemplo, da série estatística.

Uma série estatística define-se como toda e qualquer coleção de dados estatísticos referidos a uma mesma ordem de classificação: quantitativa. No sentido mais amplo, série é uma sucessão de números referidos a qualquer variável. Se os números expressarem dados estatísticos, a série será chamada **de série estatística**. Em sentido mais estreito, pode-se dizer que uma série estatística é uma sucessão de dados estatísticos, referido a caracteres quantitativos, ao passo que sucessão de dados estatísticos configurará uma seriação. Em outros termos, a série é usada normalmente para designar um conjunto de dados dispostos de acordo com caracter variável, residindo a qualidade serial na disposição temporal ou espacial de indivíduos.

Para diferenciar uma série estatística de outra, há que se levar em conta, então os três caracteres presentes na tabela que apresenta:
1. a **época** (fator temporal ou cronológico) a que se refere o fenômeno analisado;
2. o **local** (fator espacial ou geográfico) onde o fenômeno acontece;
3. o **fenômeno** (espécie do fato ou fator especificativo) que é descrito.

As séries estatísticas podem ser de quatro tipos, conforme varie seus caracteres ou fatores. Embora seja a variação desses elementos a característica diferenciadora das séries, costuma-se dividi-las em dois grupos:

1ª - Série Homógrada

Série homógrada é aquela em que a variável descrita apresenta variação discreta ou descontínua. São séries homógradas a série temporal, a série geográfica e a série específica.

2ª - Série Heterógrada

A série heterógrada é aquela na qual o fenômeno ou o fato apresenta gradações ou subdivisões. Embora fixo, o fenômeno varia em intensidade. A distribuição de freqüências ou seriação é uma série heterógrada.

4.1.1– Tipos de Séries Estatística

A)- Cronológica - ou Temporal, ou Marcha, ou Histórica, ou de Andamento.
– Elemento variável: época
– Elementos fixos: local e fenômeno

São constituídas por dados produzidos e monitorados ao longo do tempo (anos, meses, dias, semanas, horas, minutos, semestres, bimestres, etc.).

Ex.: O Diretor de Marketing de uma empresa fabricante de componentes eletrônicos, deseja examinar a evolução de suas vendas durante o último ano, mês a mês. Para tanto solicita ao Departamento de Análise de Mercado uma tabela na qual conste os valores das vendas no período desejado.

B)- Geográfica - ou Espacial, ou Territorial, ou de Localização
– Elemento variável: local
– Elementos fixos: época e fenômeno

São constituídas por dados provenientes de diferentes regiões geográficas:
Continentes: Europa, África, Ásia, etc..
Localidades: Juiz de Fora, Matias Barbosa, Belo Horizonte, Bicas, etc..

Ex.: O Diretor de Marketing dessa empresa desejar saber agora, o comportamento das vendas efetuadas nos vários Estados do Brasil..

C)- Especificativa - ou Categórica, ou por Categoria.
– Elemento variável: fenômeno
– Elementos fixos: época e local

São constituídas por dados obtidos nas diferentes categorias de uma mesma variável. Varia o fenômeno:
Cores: Vermelha, Amarela, Branca, etc..
Produtos: Café, Arroz, Feijão, etc..

Ex.: O Diretor de Marketing está interessado em conhecer o comportamento das vendas de cada um de seus produtos.

D)- Distribuição de (ou por) Freqüências - ou Seriação.
– Elementos fixos: local, época e fenômeno

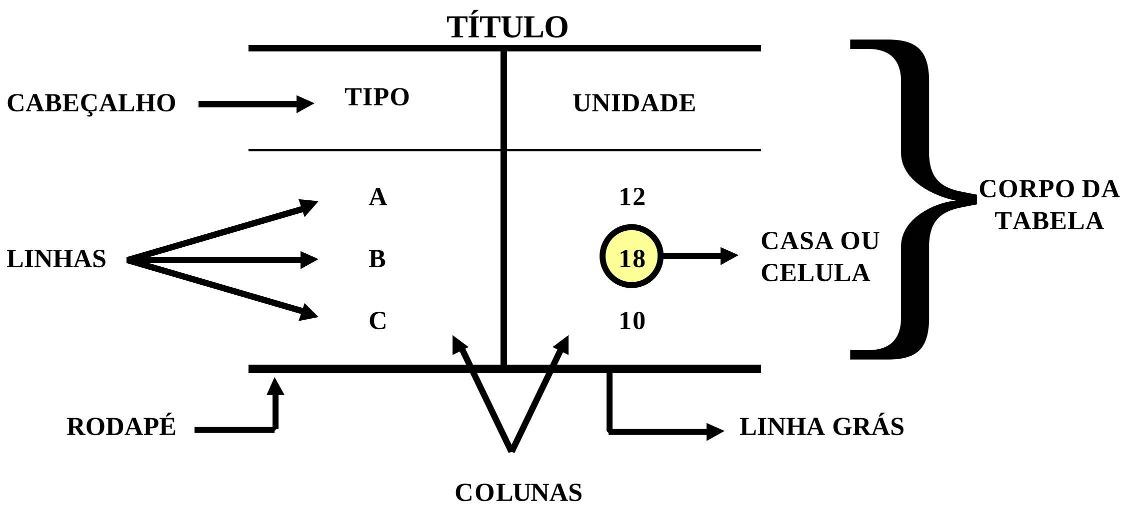
Embora fixo o fenômeno apresenta-se agora através de gradações, isto é, os dados referentes ao fenômeno que se está representando são reunidos de acordo com sua magnitude.

Ex.: Metros:	0 a 10,	10 a 20,	20 a 30,	30 a 40,	40 a 50.
Notas:	0 a 2,	2 a 4,	4 a 6,	6 a 8,	8 a 10.

4.2– Tabelas Estatísticas

As tabelas podem apresentar um grau considerável de organização dos dados. Sua utilidade não se restringe à apresentação organizada e resumida dos valores absolutos coletados segundo uma ordem previamente estabelecida. Em muitos casos se presta a análises de diversos tipos, como avaliações de desempenho no passado e estimativas e projeções para o futuro.

4.2.1– Características Básicas das Tabelas



1–Título

É a parte superior da tabela, na qual se indicam a natureza do fato, o local e a época em que o fenômeno foi observado. (Responder: o que?, quando? e onde?)

Ex.: Produção Brasileira de Trigo - Período 1986 a 2000

2–Cabeçalho

É a parte da tabela que indica a natureza do fenômeno. Especifica o conteúdo das colunas

Ex.: tipo, unidade, quantidade, salário, idade, tonelada, metro, etc..

3–Rodapé

Localiza-se logo após a linha grás que encerra a tabela. Espaço destinado à colocação da fonte de tais informações.

4–Fonte

É a origem das informações da tabela. Localiza-se no rodapé da mesma.

Ex.: IBGE, FGV, Secretaria da Faculdade, etc..

4.2.2– Regras Gerais para Apresentação de Tabelas

1– Cada tabela deve ter significação própria de modo a favorecer a interpretação.
2– Nenhuma casa deve ficar em branco.
3– Evitar a apresentação de tabelas com poucas informações.
4– Nenhuma tabela deverá ser disposta de maneira que a leitura exija a mudança de posição do papel.
5– As tabelas não são fechadas lateralmente.
6– Quando em uma tabela, mais de uma coluna for apresentada sob a mesma especificação, esta deverá ser separada por um conjunto de linhas diferentes.
7– A espessura das linhas do corpo da tabela é proporcional à sua ordem de grandeza
8– Não existem linhas horizontais
9– Sinais: **(-)** quando o valor numérico é nulo, zero
(. . .) quando não se dispõe de dados no momento
(?) quando há dúvidas sobre a exatidão de determinado valor
(0) quando o valor numérico é bem menor do que a unidade utilizada

4.2.3– Arredondamento de Números

Se o número a ser eliminado for 0, 1, 2, 3 ou 4 , arredondamento por falta (permanece o mesmo)
Ex.: 4,273 = 4,27

Se o número a ser eliminado for 6, 7, 8 ou 9, arredondamento por excesso (passa para o seguinte)
Ex.: 4,278 = 4,28

<u>O caso do número 5</u> – (Res. 886 de 06/10/66 - IBGE)

Se o número antecessor à referência 5 for par, ele é mantido. Se for ímpar, será aumentado de uma unidade.
Esta regra só se aplica se o número 5, que será eliminado, vier seguido de zeros ou desacompanhado de valor.
Exs: 3,455 = 3,46
3,465 = 3,46

4.3– Combinações de Séries Estatísticas

4.3.1– Cronológica e Geográfica:

DIA	MUNICIPIOS	
	JUIZ DE FORA	UBÁ
1		
2		
3		

4.3.2– Geográfica e Distribuição:

ZONA	PRODUÇÃO	
	0 a 10	10 a 20
NORTE		
SUL		
LESTE		

4.3.3– Distribuição e Distribuição:

PESO (KG)	ALTURA (m)	
	0 a 10	10 a 20
10 a 20		
20 a 30		
30 a 40		

4.3.4– Geográfica e Geográfica:

ESTADOS	ZONAS	
	NORTE	SUL
MG		
RJ		
SP		

4.3.5– Especificativa, Cronológica e Distribuição:

PROFISSÕES	1995	1996
	500 a 700	700 a 900
ADM. EMPRESA		
CONTADOR		
ADVOGADO		

4.4– Distribuição por Freqüências

Freqüentemente, o estudo de um determinado fenômeno requer a coleta de uma grande massa de dados numéricos. Difícil de ser tratada se esses dados não forem organizados e condensados em uma tabela. Acontece normalmente que, ao coletar os dados referentes ao fenômeno objeto de estudo, o analista se defronta com valores que se repetem algumas vezes. Por isso é necessário agrupar os dados em tabelas de distribuição de freqüências, que é uma das formas mais usadas para sintetizar os dados.

Para a construção de uma tabela de freqüências, é conveniente adotar-se um roteiro que, embora baseado em critérios relativamente arbitrários, facilita e torna mais operacional o trabalho de quem irá montar a tabela. O roteiro proposto consta dos seguintes passos:

4.4.1– Dados Brutos

Ex.: 13, 18, 10, 10, 17, 07, 12, 13, 00, 19, 01, 18, 02, 08, 04
13, 19, 05, 07, 14, 19, 02, 03, 10, 02, 09, 12, 08, 18.

4.4.2– Dados Ordenados (Rol)

Ex.: 00, 01, 02, 02, 02, 03, 04, 05, 07, 07, 08, 08, 09, 10, 10,
10, 12, 12, 13, 13, 13, 14, 17, 18, 18, 18, 19, 19, 19.

4.4.3– Formação da Série

METRO	fi
00 a 05	7
05 a 10	6
10 a 15	9
15 a 20	7
Σ ou **N** ou **n**	**29**

Classes de Freqüência
(CLASSES)

Freqüências de Classes
(FREQÜÊNCIAS)

Limites de Classes

OBS.: Simbologias usadas nas distribuições por freqüências, para determinação da amplitude de classe:

1 ├── **5** ou **1** a **5**	- O limite inferior pertence à classe e o superior não
1 ──┤ **5**	- O limite inferior não pertence à classe e o superior sim
1 ├──┤ **5**	- Os dois limites pertencem a classe
1 ── **5**	- Os dois limites <u>*não*</u> pertencem a classe

4.5– Determinação do Número de Classes.

O número de classes numa distribuição de freqüências é representado por **K**. É importante que a distribuição conte com um número adequado de classes. Se esse número for escasso, os dados originais ficarão tão condensados que pouca informação se poderá extrair da tabela. Se, por outro lado, forem utilizadas muitas classes, haverá algumas com freqüências nulas ou muito pequena, e o resultado será uma distribuição irregular e prejudicial à interpretação do fenômeno como um todo.

Para determinar o número de classe há diversos métodos. A regra de Sturges, um dos métodos, estabelece que o número de classes é igual a:

4.5.1– Fórmula de Sturges:

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

onde: **N** = Número de observações
K = Número de classes

Exemplos:

a- Foram realizados 100 testes de medidas lineares observando-se que o menor valor encontrado foi 2 m e o maior 49 m. Através da fórmula de Sturges determine o número de classes e monte uma tabela.

b- Foram realizados 100 testes de medidas lineares observando-se que o menor valor encontrado foi 2 m e o maior 53 m. Através da fórmula de Sturges determine o número de classes e monte uma tabela.

c- Fazer a letra **'a'** do exercício da página 81.

Observação: O número de classes (determinado pela fórmula de Sturges) e a amplitude, são usados como base para a montagem de uma tabela.
Podemos aumentar ou diminuir o número de classes e arredondar uma amplitude decimal. <u>Use o bom senso</u>

4.6 – Algumas Abreviaturas Usadas nas Distribuições

Xi	**fi**	PM	Ic =c =h	fiac↑	fiac↓	fr	frac↑	frac↓	fi %	fi%ac↑	fi%ac↓
00 a 05	15	2,5	5	15	140	0,1071	0,1071	1,0000	10,71	10,71	100,00
05 a 10	20		5								
10 a 15	30		5								
15 a 20	40		5								
20 a 25	25		5								
25 a 30	10		5								
å	**140**	-	-	-	-	**1,0000**	-	-	**100,00**	-	-

fi = *Freqüência absoluta simples*

PM = *Ponto Médio*

$$PM = \left(\frac{Limite\ inferior\ +\ Limite\ superior}{2} \right)$$ ***CUIDADO***

Ex: $PM = \frac{5+10}{2} = 7{,}5$ *ou metade da amplitude + Limite inferior*

h = *Ic* = *c* = *Intervalo de classe*

h = *Limite superior - Limite inferior* ***CUIDADO***

Ex.: 10 – 5 = 5 (amplitude também é vista no sentido vertical)

fac – = *Freqüência absoluta acumulada crescente*

fac ‾ = *Freqüência absoluta acumulada decrescente*

fr = *Freqüência relativa simples*

$fr = \frac{fi}{Sfi}$ Ex: $fr = \frac{15}{140} = 0{,}1071$

fi % = *Freqüência absoluta percentual*

$fi\% = \frac{fi}{Sfi} \cdot 100$ Ex: $fi\% = \frac{15}{140} \cdot 100 = 10{,}71\%$

CAPÍTULO 5

<u>Representação Gráfica</u>

5.1– Vantagens

a- Causam melhor impressão visual
b- Em conjunto com as tabelas, facilitam a análise e a interpretação

5.2– Desvantagens

a- Demora na confecção
b- Valores arredondados
c- Pequeno número de elementos

5.3– Tipos e Utilização

Existem três tipos de gráficos, classificados quanto ao critério da forma:

a) *Diagramas*
Os diagramas são gráficos geométricos dispostos em duas dimensões. Os diagramas são os gráficos mais usados na representação de séries estatísticas e se apresentam através de uma grande variedade de títulos.

b) *Cartogramas*
Os cartogramas são ilustrações relativas a cartas geográficas, largamente difundidas em Geografia, História e Demografia.

c) *Estereogramas*
Os estereogramas representam volumes e são apresentados em três dimensões. Muitas vezes são confeccionados em cartolina ou madeira, quando não desenhados em perspectiva.

5.4– Classificação dos Gráficos

É possível distinguir, de certo modo arbitrariamente, ***dois*** objetivos que justificariam o emprego de gráficos:

1º- os gráficos são usados para apresentar visualmente dados numéricos, proporcionando maior facilidade e rapidez de compreensão dos mesmos;

2º- apresentar conclusões ou resultados de uma análise.

5.4.1– Gráficos de Informação

São gráficos destinados principalmente ao público em geral, objetivando proporcionar uma visualização rápida e clara da intensidade das modalidades e dos valores relativos ao fenômeno observado. São gráficos tipicamente expositivos, devendo, por conseguinte, ser o mais completo possível, dispensando comentários explicativos adicionais. Nesses gráficos, não se deve prescindir dos títulos, escritos em letra de forma. Já as legendas podem ser omitidas, desde que as informações desejadas estejam presentes, possibilitando a completa interpretação do gráfico.

Alguns tipos de gráficos de informação (Vamos confeccionar os que estão em **negrito**)

- Gráfico de Barras
- Gráfico de Barras Compostas
- Gráfico de Barras Agrupadas
- Gráfico de Barras Bidirecionais
- **Gráfico em Colunas**
- Gráfico em Colunas Remontadas

- **Gráfico em Colunas Bidirecionais**
- Gráfico em Colunas Superpostas
- **Gráficos Lineares**
- Gráfico de Porcentagem Complementar
- Gráficos em Faixas
- Gráficos em Coordenadas Polares
- Gráficos Pictóricos (Pictogramas)
- Estereogramas
- **Gráficos em Setores**
- Gráficos Triangulares

5.4.2– Gráficos de Análise

Os gráficos de análise prestam-se melhor ao trabalho estatístico, fornecendo elementos úteis à fase de análise dos dados, sem deixar de ser também informativos.

Quando se usam gráficos para apresentar os resultados de uma análise, esses freqüentemente vêm acompanhados de uma tabela. Inclui-se, muitas vezes, um texto dissertativo, chamando a atenção do leitor para os pontos principais revelados pelo gráfico ou pela tabela. Muitos relatórios administrativos, econômicos ou de qualquer outra natureza combinam as três formas de apresentação de dados. Isto porque, na prática, poucas pessoas têm habilidade com números, e as que têm dificuldade consultarão, via de regra, apenas o gráfico.

Alguns tipos de gráficos de análise. (Vamos confeccionar os que estão em **negrito**)

- **Histogramas**
- **Poligonal Característica**
- **Polígono de Freqüências**
- **Polígono Freq. Acumuladas (Ogiva de Galton)**
- Gráfico em Hastes (Bastão)
- Gráfico em Escala
- **Curvas de Freqüências (Assimetrias)**
- Curva de Lorenz (*Índice de Gini*)

5.5– Construção de Gráficos

a) Gráfico de Colunas ou Barras

Têm por finalidade comparar grandezas, por meio de retângulos de igual largura, porém de alturas proporcionais às respectivas grandezas. Cada coluna (ou Barra) representa a intensidade de uma modalidade do atributo.

São empregados nas séries especificativas geográficas e cronológicas.

Características

1- Formato de retângulos cujas alturas são proporcionais aos dados da tabela
2- Bases com qualquer dimensão desde que haja proporcionalidade/harmonia no gráfico

Observação: Todas as bases devem ser iguais inclusive os intervalos de separação que podem ser de 1/2 a 2/3 da base.

3- O gráfico <u>pode ou não</u> ser fechado por um quadro denominado cercadura
4- O título e a fonte podem ser colocados em qualquer posição.

Exemplo:

INDÚSTRIA NU - JUIZ DE FORA -200X

PADRÃO	PEÇAS FABRICADAS
A	7
B	5
C	10
D	4
E	6
F	8
G	3
å	**43**

b) Gráfico em Colunas Bidirecionais

Este gráfico é muito utilizado quando se deseja representar, graficamente, quantidades positivas e negativas. ***Sua principal característica é a analise de dados opostos.***

Ex.: Ativo e passivo – débito e crédito – importação e exportação – entrada e saída

Exemplo:

MOVIMENTO DE VEÍCULOS - PEDÁGIO
MATIAS BARBOSA - MG - RODOVIA BR 040 – 200X

MÊS	SENTIDO		
	JUIZ DE FORA	TRÊS RIOS	
JAN	600	300	
FEV	680	200	
MAR	750	150	
ABR	540	800	
MAI	390	900	
JUN	450	100	
å	**3410**	**2450**	

c) Gráficos em Setores (Circular ou Retangular)

Os gráficos em setores ou setogramas são usados para representar valores absolutos ou porcentagens complementares.

Usado quando se pretende comparar as diversas partes de um todo.

Empregados para séries especificativas e geográficas.

Características

1- É construído em uma circunferência de raio qualquer
2- Os valores são proporcionais aos setores circulares.
3- É permitido inscrições nos setores.

Observações: - Neste Gráfico não se utiliza escala, pois totaliza 360º
- Acima de 5 subdivisões a comparação torna-se muito difícil.

Exemplo:

TECELAGEM MARIMAR LTDA - SÃO PAULO - SP - JAN/0X

COR DO TECIDO	METROS	GRAUS
VERMELHO	480	
VERDE	320	
AZUL	260	
AMARELO	220	
VIOLETA	160	
å	**1.440**	**360**

FONTE: TINTURARIA

d) Gráfico Linear ou Gráfico em Linhas ou Diagrama Cartesiano

Os gráficos lineares são freqüentemente usados para a representação de séries de tempo, quando um dos fatores for o tempo.

As linhas são particularmente mais eficientes do que as colunas, quando existem intensas flutuações nas séries ou quando há necessidade de se representarem várias séries em um mesmo gráfico.

Características

1- Representa a série cronológica, requerendo, entretanto, que a série apresente 5 ou mais informações para que ocorra uma melhor visualização.
2- A abscissa é dividida segundo os intervalos de tempo.
3- A altura é função dessa largura, que foi escolhida arbitrariamente.

EXEMPLO 1:

MATRÍCULAS INICIAIS DO COLÉGIO NAVAL - R. JANEIRO – BRASIL

PERÍODO	Nº DE MATRÍCULAS	
1990	816	
1991	904	
1992	1203	
1993	1147	
1994	1239	
1995	1565	
1996	1620	
1997	1833	
1998	1910	
1999	1890	
2000	1903	

Planejamento do Gráfico:

Exemplo 2:

TEMPERATURA DO AR - NITERÓI - RJ - BRASIL 200X

MÊS	TEMPERATURA MÉDIA			
	MÉDIA DAS MÁXIMAS	MÉDIA DAS MÍNIMAS		
JAN	32,6	23,3		
FEV	35,0	23,5		
MAR	32,2	22,1		
ABR	28,5	20,2		
MAI	26,5	18,6		
JUN	27,3	16,6		
JUL	26,5	16,5		
AGO	25,7	16,2		
SET	26,5	17,6		
OUT	27,8	20,0		
NOV	28,3	20,7		
DEZ	31,9	22,6		

e) Histograma

Empregado na série distribuição por freqüência.

É representado na distribuição pelos limites de classes (inferior e superior).

É um gráfico formado por um conjunto de retângulos (colunas) justapostos, próprio das séries distribuição por freqüências, de forma que a área de cada retângulo seja proporcional à freqüência da classe que ele representa. Assim sendo, a soma dos valores correspondentes às áreas dos retângulos será sempre igual à freqüência total.

Exemplo:

COLÉGIO FRANCISCA XAVIER - CLASSES DE NOTAS - 2º ANO - 200X

NOTAS	fi	
0 ├── 1	5	
1 ├── 2	9	
2 ├── 3	13	
3 ├── 4	17	
4 ├── 5	32	
5 ├── 6	45	
6 ├── 7	28	
7 ├── 8	16	
8 ├── 9	11	
9 ├── 10	4	
å	**180**	

f) Poligonal Característica

É a representação do contorno do histograma.

g) Polígono de Freqüências

Empregado na série distribuição por freqüência.

É a linha poligonal fechada que une ordenadas traçadas dos ***pontos médios*** das classes. Sua construção é feita, quase sempre, acompanhando a do histograma.

Exemplo 1:

MALHARIA MIRASOL LTDA - NÚMERO DE EMPREGADOS POR CLASSES SALARIAIS - 200X

SALÁRIOS	Nº DE EMPREGADOS		
125 ├─┤ 149	76		
150 ├─┤ 174	149		
175 ├─┤ 199	51		
200 ├─┤ 224	38		
225 ├─┤ 249	27		
250 ├─┤ 274	12		
275 ├─┤ 299	7		
å	**360**		

Exemplo 2:

Com base na tabela anterior, construa, no mesmo gráfico, um histograma e um polígono de freqüência.

h) Polígono de Freqüências Acumuladas ou Ogiva de Galton ***(Sir Francis Galton – 1822-1911)***

Empregado na série distribuição por freqüência.

A Ogiva de Galton ou Polígono de Freqüências Acumuladas tem por finalidade a representação gráfica das tabelas de freqüências acumuladas.

Na Ogiva, podemos representar qualquer tipo de freqüência acumulada, quer seja relativa ou percentual, crescente ou decrescente, mantendo sempre o eixo das abscissas e alterando a escala do eixo das ordenadas, conforme o tipo dessa freqüência.

Para sua construção, marcamos na abscissa os valores da variável e na ordenada as freqüências acumuladas.

Observações:

- A ogiva é a união dos pontos máximos em fac↑ ou mínimos em fac↓ que formam uma curva.

- Na intercessão dos polígonos temos um valor central representativo da distribuição, demonstrando que 50% dos valores observados estão acima e 50% abaixo deste ponto.

EXEMPLO:

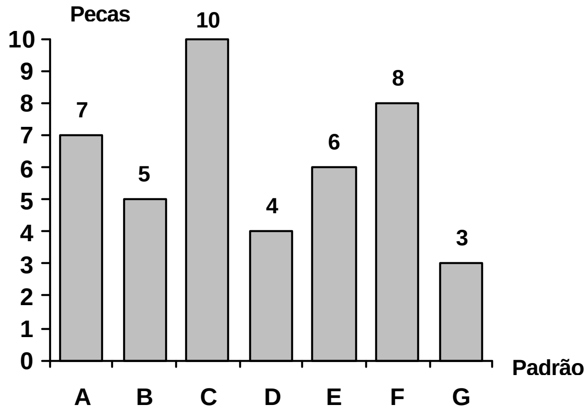
ESTATURA DOS ALUNOS DA 2ª SÉRIE DO COLÉGIO MADRE CABRINI - J. DE FORA - 200X

ESTATURA (cm)	fi				
150 |— 156	5				
156 |— 162	4				
162 |— 168	19				
168 |— 174	18				
174 |— 180	14				
180 |— 186	12				
186 |— 192	4				
å	**76**				-

RESOLUÇÃO DOS EXEMPLOS DE GRÁFICOS

INDÚSTRIA NU - JUIZ DE FORA - 200X

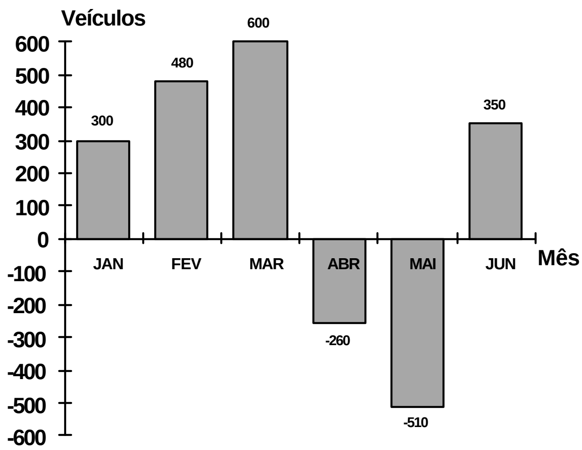
PADRÃO	PEÇAS FABRICADAS
A	7
B	5
C	10
D	4
E	6
F	8
G	3
TOTAL	43



MOVIMENTO DE VEÍCULOS - PEDÁGIO

MATIAS BARBOSA - MG - RODOVIA BR 040 - 200X

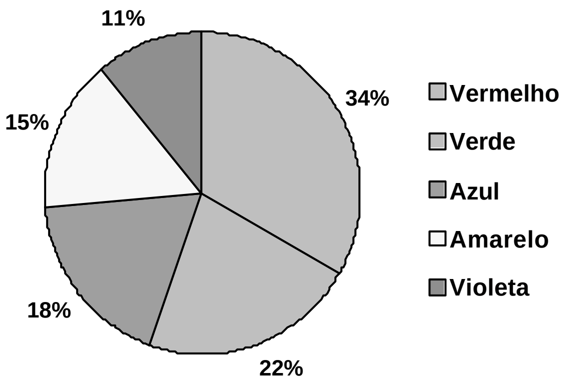
MÊS	SENTIDO		*DIFERENÇA*
	JUIZ DE FORA	TRÊS RIOS	
JAN	600	300	*300*
FEV	680	200	*480*
MAR	750	150	*600*
ABR	540	800	*-260*
MAI	390	900	*-510*
JUN	450	100	*350*
TOTAL	3410	2450	



TECELAGEM MARIMAR LTDA

São Paulo - SP - JAN/02

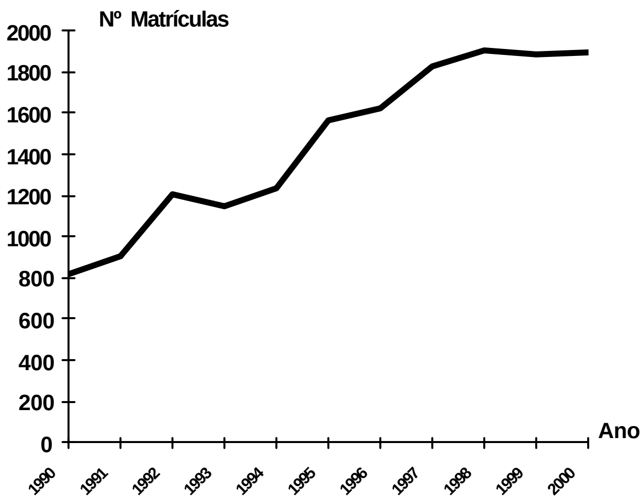
C.TECIDO	METROS	*Graus*
Vermelho	480	*120º=34%*
Verde	320	*80º=22%*
Azul	260	*65º=18%*
Amarelo	220	*55º=15%*
Violeta	160	*40º=11%*
TOTAL	1.440	*360º=100%*



MATRÍCULAS INICIAIS DO COLÉGIO NAVAL

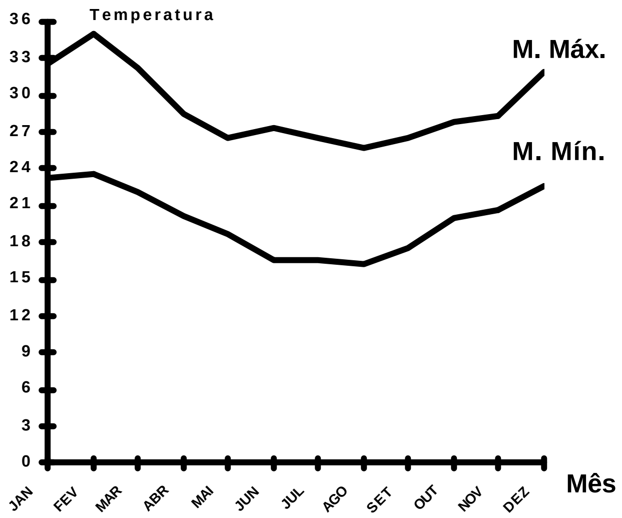
RIO DE JANEIRO - BRASIL

PERÍODO	Nº Matriculas	*ESC. 1 : 200*
1990	816	*4,1 cm*
1991	904	*4,5 cm*
1992	1203	*6,0 cm*
1993	1147	*5,7 cm*
1994	1239	*6,2 cm*
1995	1565	*7,8 cm*
1996	1620	*8,1 cm*
1997	1833	*9,2 cm*
1998	1910	*9,6 cm*
1999	1890	*9,4 cm*
2000	1903	*9,5 cm*



TEMP. AR-NITERÓI - RJ - BRASIL 200X

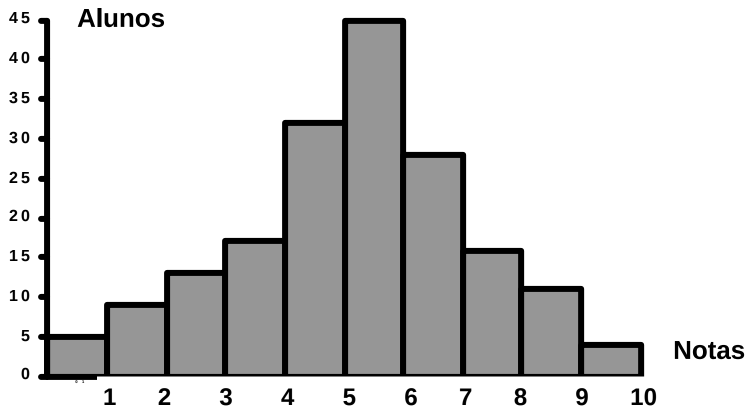
MÊS	TEMPERATURA MÉDIA		ESC 1 : 3	
	MÉDIA DAS MÁXIMAS	MÉDIA DAS MÍNIMAS	*MÉDIAS*	
			MÁX.	*MÍN.*
JAN	32,6	23,3	*10,9*	*7,8*
FEV	35,0	23,5	*11,7*	*7,8*
MAR	32,2	22,1	*10,7*	*7,4*
ABR	28,5	20,2	*9,5*	*6,7*
MAI	26,5	18,6	*8,8*	*6,2*
JUN	27,3	16,6	*9,1*	*5,5*
JUL	26,5	16,5	*8,8*	*5,5*
AGO	25,7	16,2	*8,6*	*5,4*
SET	26,5	17,6	*8,8*	*5,9*
OUT	27,8	20,0	*9,3*	*6,7*
NOV	28,3	20,7	*9,4*	*6,9*
DEZ	31,9	22,6	*10,6*	*7,5*



Colégio Franscisca Xavier

Classes Notas - 2º Ano - 200X

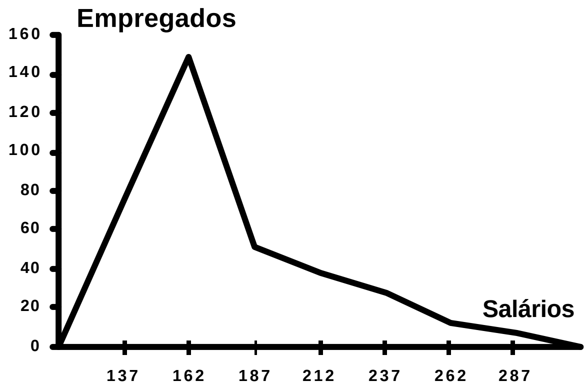
NOTAS	fi	*ESC 1 : 5*
0 ⊢ 1	5	*1,0 cm*
1 ⊢ 2	9	*1,8 cm*
2 ⊢ 3	13	*2,6 cm*
3 ⊢ 4	17	*3,4 cm*
4 ⊢ 5	32	*6,4 cm*
5 ⊢ 6	45	*9,0 cm*
6 ⊢ 7	28	*5,6 cm*
7 ⊢ 8	16	*3,2 cm*
8 ⊢ 9	11	*2,2 cm*
9 ⊢10	4	*0,8 cm*
TOTAL	180	-



MALHARIA MIRASOL LTDA

EMPREGADOS / CLASSES SALARIAIS - 200X

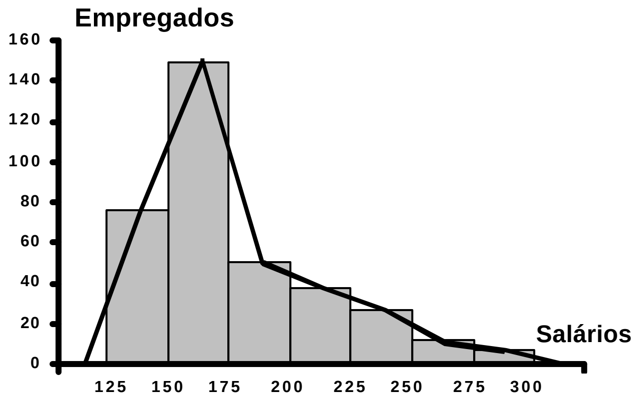
SALÁRIOS	Nº EMPREG.	*PM*	*ESC 1 : 20*
125 ├┤ 149	76	*137*	*3,8 cm*
150 ├┤ 174	149	*162*	*7,4 cm*
175 ├┤ 199	51	*187*	*2,6 cm*
200 ├┤ 224	38	*212*	*1,9 cm*
225 ├┤ 249	27	*237*	*1,4 cm*
250 ├┤ 274	12	*262*	*0,6 cm*
275 ├┤ 299	7	*287*	*0,4 cm*
TOTAL	360	-	-



MALHARIA MIRASOL LTDA

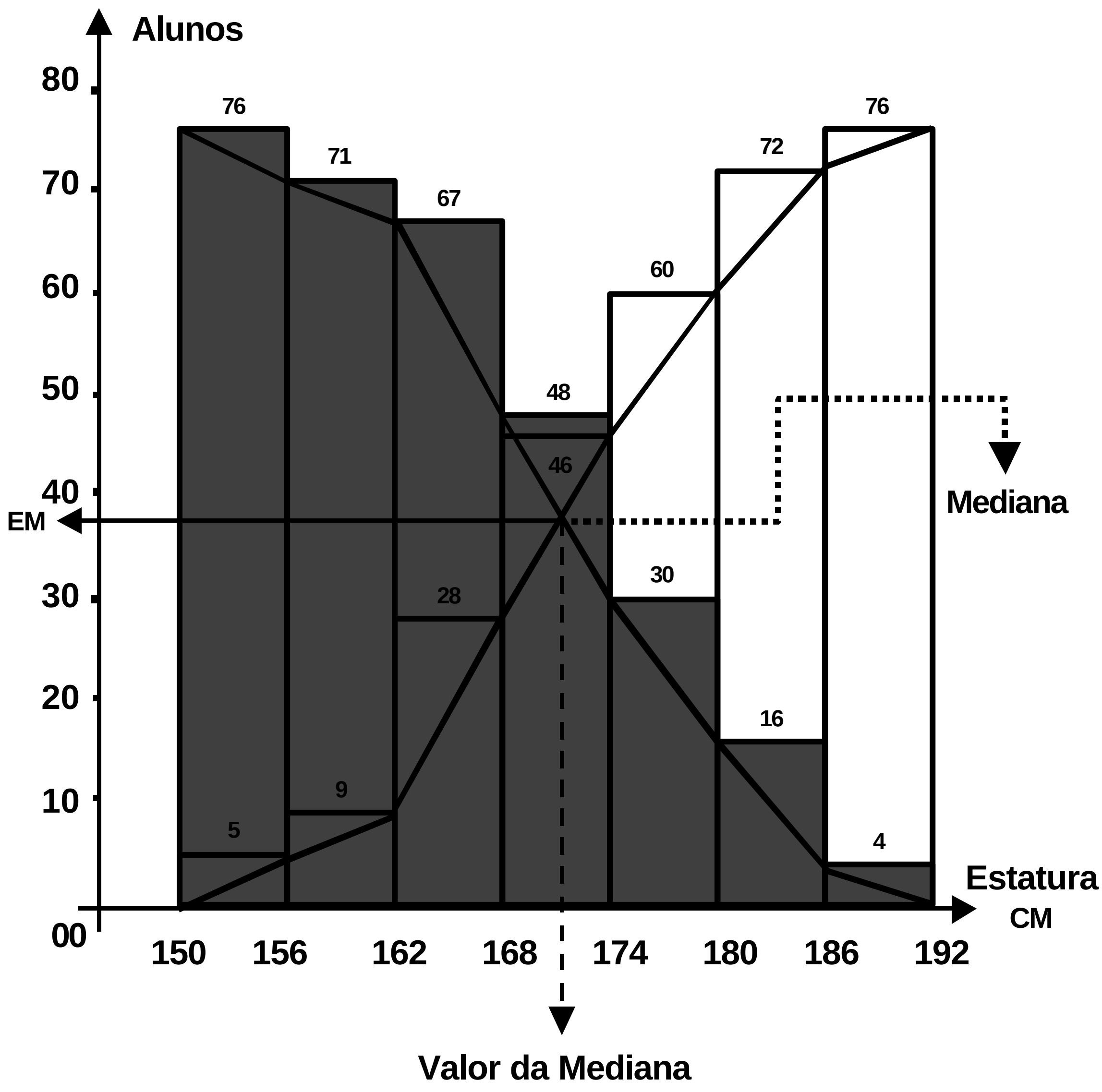
EMPREGADOS / CLASSES SALARIAIS - 200X

SALÁRIOS	Nº EMPREG.	*PM*	*ESC 1 : 20*
125├┤149	76	*137*	*3,8 cm*
150├┤174	149	*162*	*7,4 cm*
175├┤199	51	*187*	*2,6 cm*
200├┤224	38	*212*	*1,9 cm*
225├┤249	27	*237*	*1,4 cm*
250├┤274	12	*262*	*0,6 cm*
275├┤299	7	*287*	*0,4 cm*
TOTAL	360	-	-



Estatura dos alunos da 2ª série do Colégio Madre Cabrini
Rio Pomba - MG - 200X

Estatura (cm)	fi	fiac↑	fiac↓	Esc.1:10 (fiac↑)	Esc.1:10 (fiac↓)
150 a 156	**5**	*5*	*76*	*0,5*	*7,6*
156 a 162	**4**	*9*	*71*	*0,9*	*7,1*
162 a 168	**19**	*28*	*67*	*2,8*	*6,7*
168 a 174	**18**	*46*	*48*	*4,6*	*4,8*
174 a 180	**14**	*60*	*30*	*6,0*	*3,0*
180 a 186	**12**	*72*	*16*	*7,2*	*1,6*
186 a 192	**4**	*76*	*4*	*7,6*	*0,4*
TOTAL	**76**	-	-	-	-



Exercícios:

1 - Gráfico de Colunas

FGTS - ARRECADAÇÃO BRUTA - 200X

MÊS	R$ (MILHÕES)	
MAR	34.888	
ABR	52.334	
MAI	85.023	
JUN	95.254	
JUL	136.126	
AGO	162.643	
å	**566.268**	

2 - Gráfico em Setor - Circular

CONSUMO INDUSTRIAL DE ENERGIA ELÉTRICA DO BRASIL - 200X

EMPRESAS	KWH(Milhões)	GRAUS
SP – Light	13.617	
Cemig	6.763	
RJ – Light	3.226	
Chesf	1.183	
Cesp	1.258	
å	**26.047**	**360**

3 - Gráfico Linear

PRODUÇÃO DE PNEUMÁTICOS - SÃO PAULO - 1997 A 2002

PERÍODO	PNEUS (1000)	
1997	176	
1998	152	
1999	183	
2000	171	
2001	195	
2002	294	
å	**1.171**	**-**

4 - Histograma e Polígono de Freqüência ***(Separados e juntos no mesmo gráfico)***

ESTATURA DOS ALUNOS DO COLÉGIO DOS CAPUCHINHOS - J. FORA - 200X

ESTATURA (cm)	ALUNOS		
150 |—| 155	05		
156 |—| 161	09		
162 |—| 167	19		
168 |—| 173	18		
174 |—| 179	14		
180 |—| 185	12		
186 |—| 191	04		
å	**81**		

5 - Ogiva de Galton

NÚMERO DE EMPREGADOS POR CLASSE SALARIAL - ALFA BETA LTDA - J. FORA - 200X

SALÁRIOS (Em Reais)	NÚMERO DE FUNCIONÁRIOS				
125 |— 150	76				
150 |— 175	149				
175 |— 200	51				
200 |— 225	38				
225 |— 250	27				
250 |— 275	12				
275 |— 300	7				
å	**360**				

CAPÍTULO 6

Medidas de Tendência Central

Vimos, nos capítulos precedentes, que através de uma distribuição de freqüências se estabelece um sistema de classificação que descreve o padrão de variação de um determinado fenômeno estatístico. Ocorre, todavia, que poderia ser muito difícil trabalhar com a distribuição de freqüências completa, razão pela qual costuma-se lançar mão de determinadas medidas. Essas medidas sumarizam certas características importantes da distribuição de freqüências. Há diversas medidas que possibilitam condensar as informações dentro da fase analítica da Estatística Descritiva.

6.1– Média Aritmética ($\overline{X}$ ou Me)

A medida de tendência central mais comumente usada para descrever resumidamente uma distribuição de freqüências é a média, ou mais propriamente, a média aritmética.

É o valor único que representa todos os demais valores de uma série.

Pode ser: Simples e
Ponderada.

OBS.: Existem várias tipos de médias:
aritmética, geométrica, harmônica, quadrática, cúbica e biquadrática.

a) Média Aritmética Simples Para Dados <u>Não</u> Agrupados em Classes de Freqüências

Seja o conjunto X = { X_1, X_2, X_3..., X_n }
onde n = número de valores assumidos pela variável.

$$\boxed{\overline{X} = \frac{\mathbf{SXi}}{\mathbf{n}}}$$

Ex.: X = { 2, 5, 7, 6 }

b) Média Aritmética Para Valores Isolados Ponderados <u>Não</u> Agrupados em Classes de Freqüências

Quando $X_1, X_2, X_3 ..., X_n$ tiverem, respectivamente, os pesos $f_1, f_2, f_3 . . . , f_n$, a $\overline{X}$ será:

$$\overline{X} = \frac{[\ (X_1\ f_1) + (X_2\ f_2) + (X_3\ f_3) + ... + (X_n\ f_n)\ ]}{f_1 + f_2 + f_3 + ... + f_n}$$

$$\boxed{\overline{X} = \frac{\mathbf{SXifi}}{\mathbf{Sfi}}}$$

Ex.: X = { 6, 5, 5, 5, 3, 3, 3, 3, 3, 2, 2, 2, 1, 1, 1 }

c) Propriedades da Média Aritmética

1ª Propriedade- **A soma algébrica dos desvios de um conjunto de números tomados em relação à média aritmética é nula (é zero).**

seja o conjunto $X = \{X_1, X_2, X_3 \ldots, X_n\}$
$\overline{X}$ = Média Aritmética
di = Diferença entre cada valor e sua média aritmética. X_1
$d_1 = X_1 - \overline{X}$
$d_2 = X_2 - \overline{X}$ e etc.
Somando membro a membro, temos:
$d_1 + d_2 + d_3 + ... + d_n = X_1 + X_2 + X_3 + ... + X_n - n\overline{X} = 0$

Exemplo:

Xi	**di = (Xi - $\overline{C}$)**
2	
5	
7	
6	
å = 20	

2ª Propriedade - **Multiplicando-se ou dividindo-se cada elemento de um conjunto de números por um valor constante e arbitrário, a média aritmética fica multiplicada ou dividida por essa constante.**

Seja o conjunto $X = \{X_1, X_2, X_3 \ldots, X_n\}$
K uma constante qualquer.

$$\therefore K \cdot \overline{X} = K \cdot \frac{\Sigma Xi}{\Sigma fi}$$

Ex.: X = { 2, 5, 7, 6 }
$\overline{X}$ = 5
K = 6

3ª Propriedade - **Somando-se ou subtraindo-se um valor constante e arbitrário a cada um dos elementos de um conjunto de números, a média aritmética fica somada ou subtraída por essa constante.**

Seja o conjunto $X = \{X_1, X_2, X_3 \ldots, X_n\}$
K = constante qualquer

$$\therefore \left(\overline{X} \pm K\right) = \frac{\Sigma(Xi \pm K)}{\Sigma fi}$$

Ex.: X = { 2, 5, 7, 6 }
$\overline{X}$ = 5
K = 6

4ª *Propriedade* - A soma dos quadrados dos desvios, tomados em relação à média aritmética, é um mínimo (é o menor valor possível de se encontrar)

$\Sigma(Xi - \overline{X})^2 fi \quad < \quad \Sigma(Xi - K)^2 fi$
onde: K = uma constante qualquer

Ex.: Xi = { 2, 5, 7, 6 }
$\overline{X} = 5$
K = 4

Xi	**$Xi - \overline{X}$**	**$(Ci - \overline{C})^2$**	**Xi - K**	**$(Ci - K)^2$**
2				
5				
7				
6				
Σ = 20				

d) Média Aritmética Para Dados Agrupados em Classes de Freqüências (Processo Longo)

É definido como sendo o quociente entre a soma dos produtos das freqüências pelos pontos médios de cada classe e a soma de todas as freqüências.

$$\boxed{\overline{C} = \frac{\Sigma fiPM}{\Sigma fi}}$$

Exemplo:

ESTATURA (cm)	**fi**		
150 ├— 156	05		
156 ├— 162	04		
162 ├— 168	19		
168 ├— 174	18		
174 ├— 180	14		
180 ├— 186	12		
186 ├— 192	04		
Σ	**76**		

e) Média Aritmética Para Dados Agrupados em Classes de Freqüências (Processo Breve)

Com base na segunda e terceira propriedades é possível se calcular a média aritmética por este processo. ***"Usado somente quando os intervalos de classe (h) forem constantes".***

$$\boxed{\overline{C} = K_0 + h \cdot \left(\frac{\Sigma difi}{\Sigma fi} \right)}$$

onde: K_0 = ponto médio da classe escolhida (prefira o PM da classe de maior freqüência)

h = amplitude de classe

$$di = \frac{PM - K_0}{h}$$

Exemplo:

ESTATURA (cm)	fi			
150 ├── 156	05			
156 ├── 162	04			
162 ├── 168	19			
168 ├── 174	18			
174 ├── 180	14			
180 ├── 186	12			
186 ├── 192	04			
å	**76**			

Feito apenas para explicar.
Não é necessário aplicar a fórmula do $\mathbf{d_i}$
e tão pouco abrir uma coluna para PM.

Observações: Seqüência de passos para cálculo da média pelo processo breve:

1º - imaginar um dos pontos médios (PM) como se fosse a média falsa;
2º - em uma coluna qualquer (di) colocar "0" na casa cuja linha corresponda ao PM escolhido;
3º - completar a coluna (di) inscrevendo abaixo do "0" a seqüência crescente de números positivos (1, 2, 3, 4 ...) e acima os negativos (-1, -2, -3, -4...);
4º - efetuar os produtos das freqüências pelos valores da coluna (di);
5º - somar estes produtos e aplicar a fórmula.

Exercício:

Com base na tabela abaixo calcule:

a) $\overline{X}$, pelo Processo Longo **R.: 5,18**

b) $\overline{X}$, pelo Processo Breve **R.: 5,18**

NOTAS	**fi**				
$0 \vdash 1$	05				
$1 \vdash 2$	09				
$2 \vdash 3$	13				
$3 \vdash 4$	17				
$4 \vdash 5$	32				
$5 \vdash 6$	45				
$6 \vdash 7$	28				
$7 \vdash 8$	16				
$8 \vdash 9$	11				
$9 \vdash 10$	04				
å	**180**				

6.2– Mediana $(\tilde{X}\ \text{ou}\ Md)$

É o valor central de um rol (amontoado de valores organizados).

É a medida que divide um conjunto de dados em duas partes iguais.

6.2.1– Mediana para Dados Não Agrupados em Classes de Freqüências

Primeiro ordenamos os valores de forma crescente ou decrescente.

O elemento mediano será dado pela seguinte expressão:

$$\boxed{EM = \frac{N+1}{2}}$$

onde: EM = elemento mediano (indica a posição da mediana)

N = número de elementos do conjunto observado.

1º Ex.: X = { 3, 5, 7, 10, 15, 20, 23 }

2º Ex.: Y = { 3, 5, 15, 10, 7, 20 }

6.2.2– Mediana Para Dados Agrupados em Classes de Freqüências

a) 1º Processo:

Utiliza-se o modelo:

$$\boxed{\tilde{C} = Li + h \cdot \left(\frac{EM - {'fac}\uparrow}{fi} \right)}$$

onde:

Li = limite inferior da classe mediana.

h = amplitude da classe mediana.

EM = elemento mediano = $\frac{\Sigma fi}{2} = \frac{\Sigma fi + 1}{2}$ {Posição em que se encontra a mediana dentro das classes (olhar a 'fac↑)}

'fac↑ = freqüência absoluta acumulada crescente anterior à classe mediana.

fi = freqüência absoluta simples da classe mediana.

Exemplo:

NOTAS	fi	
0 ├— 2	27	
2 ├— 4	16	
4 ├— 6	34	
6 ├— 8	17	
8 ├—10	16	
å	**110**	-

b) 2º Processo:

É possível calcular a mediana pela representação geométrica as curvas de freqüências absolutas acumuladas ***(Ogiva de Galton)***.

De uma maneira geral, deve-se seguir os seguintes critérios:

b.1-constroi-se um diagrama de linhas sendo que, no eixo das ordenadas (Y) serão representadas as freqüências acumuladas e, no eixo das abscissas (X) as classes correspondentes à variável.

b 2-a partir do EM, traça-se uma paralela ao eixo das abscissas (X) até que esta se intercepte com a linha do diagrama.

b.3-deste ponto de interseção, projeta-se uma perpendicular às abscissas (X) e, neste ponto, onde ela se intercepta o eixo horizontal, chamar-se-á Mediana.

Exemplo: Com base na tabela abaixo, ache a mediana utilizando-se do 2º Processo

NOTAS	fi		
0 ├— 2	27		
2 ├— 4	16		
4 ├— 6	34		
6 ├— 8	17		
8 ├—10	16		
å	**110**		

Exercício:

Com base na tabela abaixo, calcule:

a) $\overline{X}$, pelo Processo Longo **R.: 54,010 kg**

b) $\overline{X}$, pelo Processo Breve **R.:**

c) $\tilde{X}$, pela fórmula **R.: 53,375 kg**

Peso (kg)	**fi**			
39 |— 44	03			
44 |— 49	08			
49 |— 54	16			
54 |— 59	12			
59 |— 64	07			
64 |— 69	03			
69 |— 75	01			
å	**50**			

6.3- Moda ou Norma ou Modo ou Tipo Dominante ou Média Densa, Etc.. $(\hat{C}$ ou Mo$)$

Com base no significado, Karl Pearson a introduziu na estatística no século XIX.

É definida como sendo o valor ou valores que ocorrem com maior freqüência, ou seja, é o valor em torno do qual é mais densa a concentração de observações.

Um conjunto de valores pode apresentar mais de uma moda, neste caso dizemos plurimodal, caso contrário, será unimodal, ou ainda, amodal, quando todos os valores das variáveis apresentarem uma mesma freqüência.

6.3.1– Moda para Dados Não Agrupados em Classes de Freqüências

Basta verificar aquele valor que aparece com mais freqüência.

Ex.: $X_1 = \{ 1, 2, 3, 4, 5, 6, 7, 8 \}$ Conjunto amodal

$X_2 = \{ 5, 10, 10, 15, 20, 25, 30 \}$ Conjunto unimodal: $\hat{X} = 10$

$X_3 = \{ 5, 10, 10, 15, 20, 20, 25 \}$ Conjunto plurimodal ou bimodal: $\hat{X}a = 10$

$\hat{X}b = 20$

6.3.2– Moda Para Valores Isolados Ponderados Não Agrupados em Classes de Freqüências

Quando os dados estiverem dispostos em uma tabela de freqüência, não agrupados em classes, a localização da moda é imediata, bastando para isto, verificar na tabela o valor predominante.

Exemplo:

ESTATURA (cm)	QUANTIDADE DE ALUNOS
170	3
172	5
175	7
178	10
180	15
185	11
190	4
å	**55**

$\hat{X} =$

6.3.3– Moda Para Dados Agrupados em Classes de Freqüências

Por não ser identificada facilmente, utilizamos vários processos para sua obtenção.

Primeiro identificamos a classe que contém a maior freqüência (classe modal).

a) Moda Bruta $(\hat{C}_B)$: consiste em se tomar o ponto médio da classe de maior freqüência

Exemplo:

NOTAS	fi	
0 ⊢ 2	27	
2 ⊢ 4	16	
4 ⊢ 6	34	
6 ⊢ 8	17	
8 ⊢ 10	16	
$\sum$	**110**	

$\hat{C}_B =$

b) Moda de Czuber $\left(\hat{C}_C\right)$: é a abscissa do ponto que divide a classe modal em 2 partes proporcionais às diferenças entre a freqüência da classe modal e a das respectivas classes adjacentes.

$$\boxed{\hat{C}_C = Li + h \cdot \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)}$$

ou

$$\boxed{\hat{C}_C = Li + h \cdot \left[\frac{f\,max - fant}{2 \cdot f\,max - (fant + fpost)}\right]}$$

(fmáx - fant) + (fmáx - fpost)

onde:

Li = limite inferior da classe modal

h = amplitude da classe modal

Δ_1 = fmáx - fant = freqüência absoluta simples máxima, menos a freqüência absoluta simples anterior à classe modal

Δ_2 = fmáx - fpost = freqüência absoluta simples máxima, menos a freqüência absoluta simples posterior à classe modal

Exemplo:

NOTAS	fi
0 ⊢ 2	27
2 ⊢ 4	16
4 ⊢ 6	34
6 ⊢ 8	17
8 ⊢ 10	16
$\sum$	**110**

c) Moda de King $\left(\hat{C}_K\right)$

$$\boxed{\hat{C}_K = Li + h \cdot \left(\frac{fpost}{fant + fpost}\right)}$$

Exemplo:

NOTAS	fi
0 ├— 2	27
2 ├— 4	16
4 ├— 6	34
6 ├— 8	17
8 ├—10	16
å	**110**

Observação: Moda de Czuber e de King - Processo Gráfico (**HISTOGRAMA**)

Exemplo:

NOTAS	fi	
0 ├— 2	27	
2 ├— 4	16	
4 ├— 6	34	
6 ├— 8	17	
8 ├—10	16	
å	**110**	

d) **Moda de Pearson** $\left(\hat{C}_P\right)$**:** É uma relação empírica entre a Média, a Mediana e a Moda.

OBS.: Pressupõe uma distribuição fracamente assimétrica, unimodal, com grande número de observações e pequena amplitude.

$$\boxed{\hat{C}_P = \left(3 \times \tilde{C}\right) - \left(2 \times \overline{C}\right)}$$

Exemplo:

NOTAS	fi			
0 ├— 2	27			
2 ├— 4	16			
4 ├— 6	34			
6 ├— 8	17			
8 ├—10	16			
å	**110**			

Exercícios:

1º- Com base na tabela abaixo calcule:

a) a Moda Bruta; **R.: 56,500 kg**

b) a Moda de Czuber (pela fórmula); **R.: 56,963 kg**

c) a Moda de King (pela fórmula); **R.: 56,838 kg**

d) a Moda de Pearson.
Média = 57,450 kg
Mediana = 57,438 kg
R.: 57,414 kg

Peso (kg)	**fi**			
39 ├── 44	04			
44 ├── 49	08			
49 ├── 54	16			
54 ├── 59	32			
59 ├── 64	21			
64 ├── 69	15			
69 ├── 74	04			
å	**100**			

2º- Com base na tabela abaixo, faça o que se pede:

a) calcule a Moda de Czuber (pela fórmula); **R.: 59,000 kg**

b) demonstre a moda de Czuber através de gráfico;

c) calcule a Moda de King (pela fórmula); **R.: 57,077 kg e R.: 60,596 kg**

d) demonstre a moda de King através de gráfico.

Peso (kg)	**fi**	
39 ├── 44	04	
44 ├── 49	08	
49 ├── 54	20	
54 ├── 59	32	
59 ├── 64	32	
64 ├── 69	15	
69 ├── 74	04	
å	**115**	

6.4– Separatrizes:

São valores que ocupam determinados lugares em uma série ordenada.
Seu cálculo é semelhante ao da mediana, substituindo na fórmula somente a posição do elemento a ser estudado.

Classificam-se em:

6.4.1– Quartil - são os três valores que dividem a distribuição em quatro partes iguais

Q_1 = (primeiro quartil) = valor que dele antecede 25% e sucede 75% de todos os itens da distribuição.

Q_2 = (segundo quartil) = valor igual ao da mediana, 50% para cada lado.

Q_3 = (terceiro quartil) = valor que dele antecede 75% e sucede 25% de todos os itens da distribuição.

a) Quartil Para Dados Não Agrupados em Classes de Freqüências:

Observação: ***A primeira providência é colocar os elementos em ordem crescente ou decrescente.***

$$\boxed{q_i = i \cdot \left(\frac{N+1}{4}\right)}$$

onde:

q_i = posição do elemento quartílico

i = 1, 2 ou 3

N = número de observações do conjunto de dados

Exemplos: 1º- X = {80, 107, 93, 97, 102, 85, 110} Calcular os valores de Q_1 Q_2 e Q_3

2º- Y = {80, 85, 93, 97, 102, 107} Calcular os valores de Q_1 Q_2 e Q_3

b) Quartil Para Valores Isolados Ponderados Não Agrupados em Classes de Freqüências

Exemplos: Com base nas tabelas abaixo, calcule as idades para o Q_1 Q_2 e Q_3

IDADE	fi	
20	10	
23	26	
25	08	
27	04	
29	03	
30	03	
35	01	
$\sum$	**55**	

c) Quartil Para Dados Agrupados em Classes de Freqüências

$$\boxed{Q_i = Li + h \cdot \left(\frac{q_i - fac\uparrow}{fi}\right)} \quad \text{e} \quad \boxed{q_i = i \cdot \left(\frac{\sum fi}{4}\right)}$$

onde: Q_i = valor de cada quartil

i = 1, 2 ou 3

Li = limite inferior da classe quartílica

h = amplitude da classe quartílica

q_i = posição do elemento quartílico

$fac\uparrow$ = freqüência absoluta acumulada crescente anterior à classe quartílica

fi = freqüência absoluta simples da classe quartílica

Exemplo: Calcule os valores de Q_1 Q_2 e Q_3, baseando-se na tabela abaixo:

NOTAS	fi	
0 ├— 2	27	
2 ├— 4	16	
4 ├— 6	34	
6 ├— 8	17	
8 ├— 10	16	
$\sum$	**110**	

6.4.2– Decil - São os nove valores que dividem a distribuição em dez partes iguais.

a) Decil Para Dados <u>Não</u> Agrupados em Classes de Freqüências:

Observação: ***A primeira providência é colocar os elementos em ordem crescente ou decrescente.***

$$\boxed{d_i = i \cdot \left(\frac{N+1}{10}\right)}$$

onde:

d_i = posição do elemento decílico

i = 1, 2, 3, 4 , 5, 6, 7, 8 ou 9

N = número de observações do conjunto de dados

Exemplos: 1º- X = {80, 107, 93, 97, 102, 85, 110} Calcular os valores de D_2, D_5 e D_7

2º- Y = {80, 85, 93, 97, 102, 107} Calcular os valores de D_2, D_5 e D_7

b) Decil Para Valores Isolados Ponderados <u>Não</u> Agrupados em Classes de Freqüências

Exemplos: Com base nas tabelas abaixo, calcule as idades para o D_2, D_5 e D_7

IDADE	fi	
20	10	
23	26	
25	08	
27	04	
29	03	
30	03	
35	01	
$\sum$	**55**	

c) Decil Para Dados Agrupados em Classes de Freqüências

$$\boxed{D_i = Li + h \cdot \left(\frac{d_i - \sum fac_{-}}{fi}\right)} \quad \textbf{e} \quad \boxed{d_i = i \cdot \left(\frac{\sum fi}{10}\right)}$$

Exemplo: Calcule os valores de D_2 D_5 e D_7, baseando-se na tabela abaixo:

NOTAS	fi	
0 ├— 2	27	
2 ├— 4	16	
4 ├— 6	34	
6 ├— 8	17	
8 ├— 10	16	
$\sum$	**110**	

6.4.3- Centil ou Percentil - São os noventa e nove valores que dividem a distribuição em cem partes iguais.

a) Percentil Para Dados Não Agrupados em Classes de Freqüências:

Observação: ***A primeira providência é colocar os elementos em ordem crescente ou decrescente.***

$$\boxed{c_i = i \times \left(\frac{N+1}{100}\right)}$$

onde:

c_i = posição do elemento percentílico

i = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, ... ou 99

N = número de observações do conjunto de dados

Exemplos: 1º- X = {80, 107, 93, 97, 102, 85, 110} Calcular os valores de C_{12}, C_{50} e C_{75}

2º- Y = {80, 85, 93, 97, 102, 107} Calcular os valores de C_{25}, C_{50} e C_{95}

b) Percentil Para Valores Isolados Ponderados Não Agrupados em Classes de Freqüências

Exemplos: Com base nas tabelas abaixo, calcule as idades para o C_{25}, C_{50} e C_{75}

IDADE	fi	
20	10	
23	26	
25	08	
27	04	
29	03	
30	03	
35	01	
$\sum$	**55**	

c) Percentil Para Dados Agrupados em Classes de Freqüências

$$\boxed{C_i = Li + h \times \left(\frac{c_i - 'fac_{-}}{fi}\right)} \quad \text{e} \quad \boxed{c_i = i \times \left(\frac{\sum fi}{100}\right)}$$

Exemplo: Calcule os valores de C_{25} C_{50} e C_{75}, baseando-se na tabela abaixo:

NOTAS	fi	
0 ├— 2	27	
2 ├— 4	16	
4 ├— 6	34	
6 ├— 8	17	
8 ├— 10	16	
$\sum$	**110**	

Exercícios:

1º - Dado os conjuntos abaixo calcule:

a) a Média Aritmética **R.: 10 e 14,63**
b) a Mediana **R.: 10 e 14,5**
c) a Moda (identifique o conjunto) **R.: Amodal e Amodal**
d) o valor do Primeiro Quartil **R.: 6,5 e 11,475**
e) o valor do Oitavo Decil **R.: 14,4 e 18,6**
f) o valor do Septuagésimo Sexto Centil **R.: 13,68 e 17,62**

Y = { 6, 12, 15, 7, 10 }

Z = { 10,5; 11,8; 15,4; 16,5; 20,0; 13,6 }

2º - Dada a distribuição abaixo calcule:
a) a Média Aritmética; **R.: 153,55 @ 154 cm**
b) a Mediana; **R.: 153,25 @ 153 cm**
c) as Modas: Bruta; **R.: 152,50 @ 152 cm**
de Czuber; **R.: 152,91 @ 153 cm**
de King; **R.: 152,97 @ 153 cm**
de Pearson); **R.: 152,65 @ 153 cm**
d) o valor do: Primeiro Quartil; **R.: 150,12 @ 150 cm**
Terceiro Quartil; **R.: 157,50 @ 158 cm**
e) o valor do: Quinto Decil; **R.: 153,25 @ 153 cm**
Sétimo Decil; **R.: 156,36 @ 156 cm**
f) o valor do: Sexagésimo Centil; **R.: 154,50 @ 154 cm**
Octogésimo Oitavo Centis. **R.: 161,00 @ 161 cm**

Estatura (cm)	**fi**			
140 ├— 145	09			
145 ├— 150	15			
150 ├— 155	40			
155 ├— 160	22			
160 ├— 165	10			
165 ├— 170	04			
å	**100**			

CAPÍTULO 7

Medidas de Dispersão ou de Flutuação

Muitas vezes somente os cálculos ou apresentações de um valor específico para um conjunto qualquer não são suficientes para caracterizar uma distribuição ou um conjunto de valores.

Exemplo:

DIAS / TURNOS	SEGUNDA FEIRA	TERÇA FEIRA	QUARTA FEIRA	QUINTA FEIRA	SEXTA FEIRA	TOTAL DA PRODUÇÃO	MÉDIA DA PRODUÇÃO
I	150	150	150	150	150	**750**	**150**
II	70	130	150	180	220	**750**	**150**
III	15	67	117	251	300	**750**	**150**

Se basearmos na produção média diária, não teremos como identificar o grau de relacionamento entre as variáveis, visto ser a média semanal nos três turnos igual a 150 peças.

Algumas perguntas são fundamentais para o entendimento das variáveis.
a) A produção é homogênea ?
b) A produção média semanal é suficiente para uma análise estatística ?
c) A produção diária dos três turnos é compatível com a produção média semanal ?

Tais informações são obtidas através do estudo das medidas de dispersão que permitem a análise de até que ponto estes valores apresentam oscilações para mais ou para menos, em relação a uma medida de posição fixada.

7.1– Amplitude Total ou Intervalo Total Para Dados Não Agrupados em Classes de Freqüências (AT ou IT)

É a diferença entre o maior e o menor valor de uma distribuição ou de um conjunto de valores.

$$\boxed{AT = IT = Xi_{max} - Xi_{min}}$$

Exemplo: X = { 67, 117, 35, 15, 42, 20, 231, 300 } AT =

Observação: Quando os dados estiverem agrupados em classes de freqüência utiliza-se os seguintes processos:

a) $$\boxed{AT = Ls_{máx} - Li_{min}}$$

b) $$\boxed{AT = PM_{\text{Última Classe}} - PM_{\text{Primeira Classe}}}$$

Exemplo:

Xi	fi	PM
05 ├── 10	05	7,5
10 ├── 15	10	12,5
15 ├── 20	20	17,5
20 ├── 25	10	22,5
25 ├── 30	05	27,5
å	**50**	**-**

AT = 30 – 5 = 25 OU AT = 27,5 – 7,5 = 20

7.2– Desvio Quartil ou Amplitude Semi-interquartílica Para Dados Agrupados em Classes de Freqüências (D_q)

Esta medida é calculada pela semi-diferença entre Q_3 **e** Q_1 cuja amplitude em torno da mediana abrange 50% dos valores mais centrais da distribuição e será mais exata quanto mais simétrica for a distribuição

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

Exemplo: Com base na tabela abaixo calcule o desvio quartílico:

Xi	fi	
05 ├── 10	05	
10 ├── 15	10	
15 ├── 20	20	
20 ├── 25	10	
25 ├── 30	05	
$\sum$	**50**	

DISTRIBUIÇÃO SIMÉTRICA

7.2.1– Gráfico *BOX PLOT* ou Diagrama de Tukey – (Gráfico-Caixa)

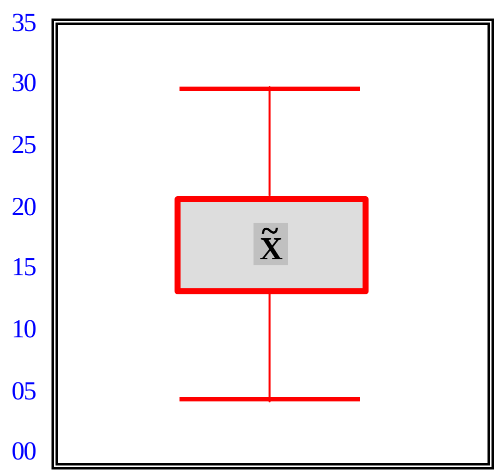
OBS.: **John Wilder Tukey** (1915 / jul-2000)

A introdução do box plot na metodologia de descrição de dados é relativamente recente, embora esse gráfico seja baseado nos quartis, que são medidas muito antigas. Mas, antes de definir box plot, é preciso definir mínimo e máximo, valores essenciais para o desenho do gráfico

Para desenhar um box plot:

a) desenhe um segmento de reta em posição vertical, para representar a amplitude dos dados;
b) marque, nesse segmento, o primeiro, o segundo e o terceiro quartis;
c) desenhe um retângulo (box) de maneira que o lado superior e o lado inferior passem exatamente sobre os pontos que marcam o primeiro e o terceiro quartis;
d) marque com um ponto o local da mediana.

Ex.: Com base no exemplo anterior, $Q_1 = 13{,}75$ $Q_2 = 17{,}50$ $Q_3 = 21{,}25$



IMPORTANTE

O retângulo do box plot é dado pela distância interquartílica; contém 50% dos dados do conjunto. Esses dados se distribuem em torno da mediana.

7.3– Desvio Médio (DM)

O desvio médio ou média dos desvios é igual à média aritmética dos valores absolutos dos desvios tomados em relação a média ou à mediana

a) Desvio Médio Para Dados <u>Não</u> Agrupados em Classes de Freqüências

$$\boxed{\mathbf{DM = \frac{\Sigma|di|}{n}}}$$

onde: $di = |Xi - \overline{X}|$

n = número de observações

Exemplo: X = {2, 15, 8, 5, 20}

b) Desvio Médio Para Dados Agrupados em Classes de Freqüências

$$\boxed{\mathbf{DM = \frac{\Sigma|di|fi}{\Sigma fi}}}$$

onde: $di = |PM - \overline{X}|$

Exemplo:

Xi	fi				
05 ├— 10	05				
10 ├— 15	10				
15 ├— 20	20				
20 ├— 25	10				
25 ├— 30	05				
Σ	**50**				

Observação: Características do Desvio Médio

1- Depende de todos os valores da distribuição

2- Seu cálculo pode ser efetuado a partir da média ou da mediana

3- Não leva em consideração a existência de desvios negativos

7.4– Desvio Padrão (S da letra Sigma)

É a medida de dispersão mais usada, porém não tem uma interpretação física, como ocorre com a média, mediana, moda e os quantis. Contudo, é possível interpretá-lo de forma analítica.

É a raiz quadrada da média aritmética dos quadrados dos desvios, estes tomados a partir da média aritmética.

Desvio padrão mede a dispersão absoluta ou a variabilidade de uma distribuição.

a) Desvio Padrão Para Dados <u>Não</u> Agrupados em Classes de Freqüências

$$\boxed{S = \sqrt{\frac{\Sigma\, di^2}{n-1}}} \quad \text{ou} \quad \boxed{S = \sqrt{\frac{\Sigma\, Xi^2 - \frac{(\Sigma\, Xi)^2}{n}}{n-1}}}$$

onde: $di = Xi - \overline{X}$

Exemplo: X = {5, 8, 10, 12, 15}

Xi			
5			
8			
10			
12			
15			
Σ = 50			

b) Desvio Padrão Para Valores Isolados Ponderados <u>Não</u> Agrupados em Classes de Freqüências

$$\boxed{S = \sqrt{\frac{\Sigma\, di^2 fi}{n-1}}} \quad \text{ou} \quad \boxed{S = \sqrt{\frac{\Sigma\, Xi^2 fi - \frac{(\Sigma\, Xifi)^2}{n}}{n-1}}}$$

onde: $di = Xi - \overline{X}$

Exemplo:

Xi	fi						
5	7						
8	10						
10	20						
12	15						
15	5						
Σ	**57**						

c) Desvio Padrão Para Dados Agrupados em Classes de Freqüências

c.1– Processo Longo

$$\boxed{S = \sqrt{\frac{\Sigma\, di^2 fi}{n - 1}}} \quad \text{ou} \quad \boxed{S = \sqrt{\frac{\Sigma PM^2 fi - \frac{(\Sigma\, PMfi)^2}{n}}{n - 1}}}$$

onde: $di = PM - \overline{X}$

Exemplo:

PESO (KG)	fi							
39,5 ├— 44,5	3							
44,5 ├— 49,5	8							
49,5 ├— 54,5	16							
54,5 ├— 59,5	12							
59,5 ├— 64,5	7							
64,5 ├— 69,5	3							
69,5 ├— 74,5	1							
Σ	**50**							

c.2– Processo Breve

Usado somente quando os intervalos de classe forem constantes. **(h iguais)**

$$\boxed{S = h \times \sqrt{\frac{\Sigma di^2 fi - \frac{(\Sigma\, difi)^2}{n}}{n - 1}}}$$

onde: $di = \dfrac{PM - K_0}{h}$ (este **di** é o do cálculo da média pelo processo breve)

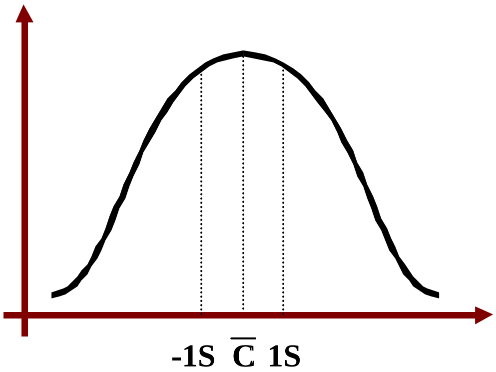
Exemplo:

PESO (KG)	fi				
39,5 ├— 44,5	3				
44,5 ├— 49,5	8				
49,5 ├— 54,5	16				
54,5 ├— 59,5	12				
59,5 ├— 64,5	7				
64,5 ├— 69,5	3				
69,5 ├— 74,5	1				
Σ	**50**				

d) Características do Desvio Padrão

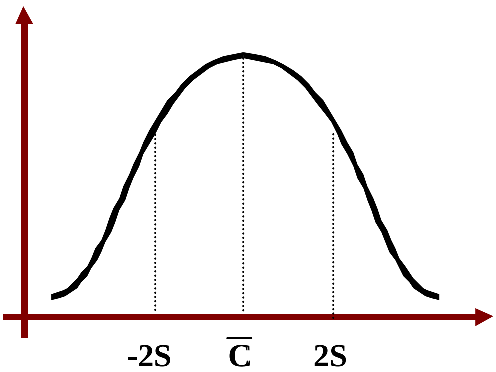
Observação: Considere toda distribuição como sendo "NORMAL" (teórica); aquela que é simétrica e mesocúrtica, ou seja, tem o coeficiente de curtose igual a 0,263.

PRIMEIRO CASO



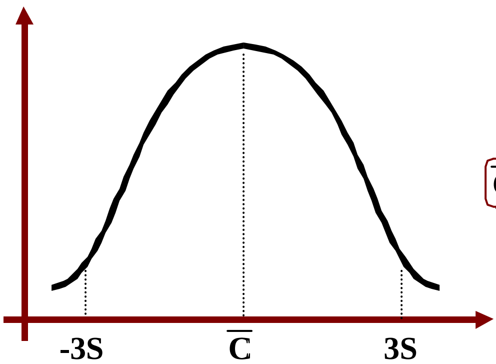
$\overline{C} \pm 1S = 68{,}27\%$ **dos itens da distribuição**

SEGUNDO CASO



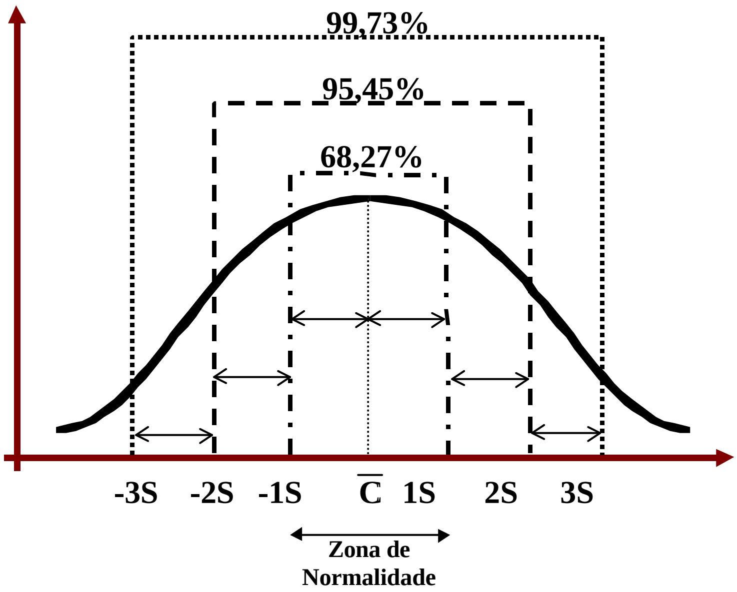
$\overline{C} \pm 2S = 95{,}45\%$ **dos itens da distribuição**

TERCEIRO CASO



$\overline{C} \pm 3S = 99{,}73\%$ **dos itens da distribuição**

RESUMO DAS CARACTERÍSTICAS DO DESVIO PADRÃO



Exemplo de Aplicação da "Zona de Normalidade"

100 recém nascidos - sexo masculino
$\overline{X}$ = 48 cm
S = 2 cm

Altura normal = $\overline{X} \pm 1S = 48 \pm 2 \Rightarrow$ ou seja, uma criança com estatura variando de 46 a 50 cm será considerada de estatura normal.

Demonstração do Primeiro Caso das Características do Desvio Padrão

PESO (KG)	fi	di	difi	di^2	di^2fi
39,5 ├── 44,5	3	-2	-6	4	12
44,5 ├── 49,5	8	-1	-8	1	8
49,5 ├── 54,5	16	0	0	0	0
54,5 ├── 59,5	12	1	12	1	12
59,5 ├── 64,5	7	2	14	4	28
64,5 ├── 69,5	3	3	9	9	27
69,5 ├── 74,5	1	4	4	16	16
å	**50**	-	**25**	-	**103**

$$\overline{X} = 52 + 5 \cdot \left(\frac{25}{50}\right) = 54{,}500 \text{ kg}$$

$$S = 5 \cdot \sqrt{\frac{103 - \frac{(25)^2}{50}}{49}} = 5 \cdot \sqrt{1{,}84694} = 5 \cdot 1{,}359 \cong 6{,}795 \text{ kg}$$

$\bar{X} \pm 1S$ = 68,27% (é o que queremos provar)

No exemplo a $\bar{X} \pm 1S$ é igual a 54,500 kg + 6,795 kg = 61,295 kg e
54,500 kg - 6,795 kg = 47,705 kg

Portanto 47,705 kg é o início do intervalo e 61,295 kg é o fim

49,500 kg (Ls da 2ª classe)	-	47,705 kg	(início do intervalo)	=	1,795 kg
61,295 kg (final do intervalo)	-	59,500 kg	(Li da 5ª classe)	=	1,795 kg

Utilizando-se uma regra de três, teremos:

8 pessoas estão para 5 quilos, (que é a amplitude de classe), assim como
X pessoas estão para 1,795 kg, (que é a quantidade em quilos utilizados da 2ª classe

Portanto X = 2,872 pessoas (quantidade de freqüência, da 2ª classe que será usada)

7 pessoas estão para 5 quilos, (que é a amplitude de classe), assim como
X pessoas estão para 1,795 kg, (que é a quantidade em quilos utilizados da 5ª classe

Portanto X = 2,513 pessoas (quantidade de freqüência, da 5ª classe que será usada)

Concluindo, no intervalo compreendido entre a $\bar{X} \pm 1S$ existem 33,385 pessoas (2,872 da 2ª classe + 16 da 3ª classe + 12 da 4ª classe + 2,513 da 5ª classe) que representa, aproximadamente, 66,8%, do total de pessoas da distribuição.

DEMONSTRANDO: 50 pessoas = 100%
33,385 pessoas = X% donde X = 66,8%

Repare que não encontramos 68,27%, pois a distribuição não é "NORMAL".

7.5– Coeficiente de Variação

É uma medida de dispersão relativa que indica a relação percentual entre o desvio padrão e a média dos dados. *Serve de termo de comparação entre duas ou mais situações diferentes*.

$$\boxed{CV = \left(\frac{S}{\bar{X}}\right) \times 100}$$

Exemplo: S = 8,795 Kg
$\bar{X}$ = 54,500 Kg

$$CV = \frac{8,795}{54,500} \cdot 100 = 16,14\%$$

S = R$ 81.473,39
$\bar{X}$ = R$541.326,53

$$CV = \frac{81.473,39}{541.326,53} \cdot 100 = 15,05\%$$

7.6 – Erro Padrão da Média

O Erro Padrão da amostra é definido como:

$$\boxed{\mathbf{S_{\overline{x}} = \frac{S}{\sqrt{n}}}}$$

O Erro Padrão da Média é uma estatística que "corrige" a inexatidão da média como caracterizadora de um conjunto. Pela expressão acima, podemos perceber que:

a) quanto maior o desvio padrão, pior é a média como medida característica da amostra;

b) quanto maior a amostra, menor será o erro padrão e, teoricamente, a média torna-se, para grandes amostras, uma medida característica da tendência central dos dados.

A expressão acima traduz matematicamente que a média pode não ser característica de um conjunto, em termos de representatividade, se o conjunto não for homogêneo. A Irregularidade do conjunto é expressa pelo desvio padrão – quanto maior o desvio padrão, menos homogênea é a amostra. Por outro lado, conjuntos maiores tendem a possuir médias mais características, embora isso não seja aplicável a todos os casos.

Exemplo: Com base nos dados da última tabela, calcule o erro padrão da média.

S = 6,795 kg
n = 50 unidades

$$S_{\overline{x}} = \frac{6{,}795}{\sqrt{50}} = \frac{6{,}795}{7{,}07} = 0{,}96 \text{ kg}$$

Exercício:

Dada a tabela abaixo, pede-se:

a) a receita média;
b) a receita mediana;
c) o desvio médio;
d) a receita na faixa de 25% a 75% dos itens da distribuição;
e) a variabilidade da receita em torno da média para um desvio padrão;
f) a quantidade de empresas cuja receita esteja compreendida entre a $\overline{X} \pm 1S$ (Suponha que a distribuição seja "NORMAL");
g) o percentual de empresas cuja receita situa-se entre a $\overline{X} \pm 1DM$;
h) a variação percentual da receita;
i) o erro padrão da média.

Receita (em 1.000 R$)	**fi**								
350 —┤ 400	05								
400 —┤ 450	09								
450 —┤ 500	15								
500 —┤ 550	25								
550 —┤ 600	18								
600 —┤ 650	16								
650 —┤ 700	10								
å	**98**								

RESPOSTAS:

a) R$ 541.330,00;
b) R$ 540.000,00;
c) R$ 66.970,00;
d) Q_1 = R$ 485.000,00 e Q_3 = R$ 604.690,00, portanto,
É a receita que varia de R$ 485.000,00 a R$ 604.690,00;
e) S = 81,47
A variabilidade da receita em torno da média para um desvio padrão é a que oscila entre R$459.860,00 e R$ 622.800,00;
f) ± 67 empresas;
g) 54,44%;
h) 15,05%;
i) R$ 8.230,00

CAPÍTULO 8

Momentos

Os momentos podem ser caracterizados como quantidades numéricas, calculadas a partir de uma distribuição de freqüências (ou de probabilidades), e que são utilizadas para fornecer descrições resumidas da distribuição estudada. Dentro da ampla classe dos momentos estão incluídas três importantes medidas estudadas anteriormente: a média, a variância (que é desvio padrão ao quadrado) e, por conseqüência, o próprio desvio padrão.

Como vemos, a noção de momento é **genérica** e abrange igualmente aquelas três medidas. Apenas nós as tratamos separadamente devido à sua grande importância no contexto da Estatística Descritiva.

8.1– Momentos

8.1.1– Momento Natural (Absoluto) de Ordem ‘r’

O momento natural de ordem *‘r’* de um conjunto de números é definido da seguinte forma:

$$\boxed{m'_r = \frac{\Sigma PM^r fi}{\Sigma fi}}$$

onde: ‘r’ é um número inteiro e positivo

Ex.: 1 = Momento natural de primeira ordem (ou primeiro momento natural).
2 = Momento natural de segunda ordem (ou segundo momento natural)...

Exemplo: Dada a tabela abaixo, calcular os momentos de primeira, segunda, terceira e quarta ordens da seguinte distribuição de freqüências:

Classes	fi	PM	PM^1fi	PM^2fi	PM^3fi	PM^4fi
10 —| 20	2	15	30	450	6.750	101.250
20 —| 30	4	25	100	2.500	62.500	1.562.500
30 —| 40	5	35	175	6.125	214.375	7.503.125
40 —| 50	8	45	360	16.200	729.000	32.805.000
50 —| 60	5	55	275	15.125	831.875	45.753.125
60 —| 70	4	65	260	16.900	1.098.500	71.402.500
70 —| 80	2	75	150	11.250	843.750	63.281.250
Σ	**30**	-	1350	68.550	3.786.750	222.408.750

$$m'_1 = \frac{1.350}{30} = 45 \qquad m'_2 = \frac{68.550}{30} = 2.285$$

$$m'_3 = \frac{3.786.750}{30} = 126.225 \qquad m'_4 = \frac{222.408.750}{30} = 7.413.625$$

8.1.2– Momento de Ordem ‘r’ em Relação a uma Origem Qualquer “X_o”

O momento de ordem ***‘r’*** em relação à origem arbitrária $\boldsymbol{X_o}$ de um conjunto de números é definido da seguinte forma:

$$\boxed{{}_{x_o}\mathbf{m}_r = \frac{\mathbf{S(PM - X_o)^r fi}}{\mathbf{Sfi}}}$$

onde: ‘r’ é um número inteiro e positivo
X_o é uma constante qualquer

Ex.: 1 = Momento de primeira ordem em relação à origem $X_0 = 40$.
2 = Momento de segunda ordem em relação à origem $X_0 = 40$...

Exemplo: Dada a tabela abaixo, calcular os momentos de primeira, segunda e terceira ordens em relação à origem $X_o = 40$.

Classes	fi	PM	(PM-X_o)	(PM-X_o)fi	$(PM-X_o)^2fi$	$(PM-X_o)^3fi$
10 —┤ 20	2	15	15-40=-25	-50	1.250	-31.250
20 —┤ 30	4	25	25-40=-15	-60	900	-13.500
30 —┤ 40	5	35	35-40=-05	-25	125	-625
40 —┤ 50	8	45	45-40= 05	40	200	1.000
50 —┤ 60	5	55	55-40= 15	75	1.125	16.875
60 —┤ 70	4	65	65-40= 25	100	2.500	62.500
70 —┤ 80	2	75	75-40= 35	70	2.450	85.750
å	**30**	-	-	150	8.550	120.750

$${}_{40}m_1 = \frac{150}{30} = 5 \qquad {}_{40}m_2 = \frac{8.550}{30} = 285 \qquad {}_{40}m_3 = \frac{120.750}{30} = 4.025$$

8.1.3– Momento Centrado na Média de Ordem ‘r’

O momento de ordem ***‘r’*** centrado na $\boldsymbol{\overline{C}}$ de um conjunto de números é definido fazendo $\boldsymbol{X_o = \overline{C}}$, ou seja:

$$\boxed{\mathbf{m}_r = \frac{\mathbf{S(PM - \overline{X})^r fi}}{\mathbf{Sfi}}}$$

onde: ‘r’ é um número inteiro e positivo

Ex.: 1 = Momento centrado de primeira ordem.
2 = Momento centrado de segunda ordem...

OBS.: Para ***r = 2***, o momento centrado na média, corresponde à variância da distribuição **(S^2)**.

Exemplo: Usando os dados da distribuição de freqüências abaixo, calcular os quatro primeiros momentos centrados na média.

Classes	**fi**	**PM**	**PMfi**	**(PM-$\overline{X}$)fi**	**(PM-$\overline{X}$)^{2}fi**	**(PM-$\overline{X}$)^{3}fi**	**(PM-$\overline{X}$)^{4}fi**
10 —| 20	2	15	30	(-30).2=-60	1.800	-54.000	1.620.000
20 —| 30	4	25	100	(-20).4=-80	1.600	-32.000	640.000
30 —| 40	5	35	175	(-10).5=-50	500	-5.000	50.000
40 —| 50	8	45	360	0.8 = 0	0	0	0
50 —| 60	5	55	275	10.5 = 50	500	5.000	50.000
60 —| 70	4	65	260	20.4 = 80	1.600	32.000	640.000
70 —| 80	2	75	150	30.2 = 60	1.800	54.000	1.620.000
å	**30**	-	1.350	0	7.800	0	4.620.000

$$\overline{X} = \frac{1.350}{30} = 45$$

$$m_1 = \frac{0}{30} = 0 \qquad m_2 = \frac{7.800}{30} = 260 \qquad m_3 = \frac{0}{30} = 0 \qquad m_4 = \frac{4.620.000}{30} = 154.000$$

Exercício:

Dada a tabela abaixo, calcule:

A) os momentos naturais de primeira e segunda ordens;

B) os momentos de segunda e terceira ordens em relação à origem $X_o = 10$;

C) os momentos de segunda, terceira e quarta ordens centrados na média.

Classes	**fi**									
00 ├── 04	1									
04 ├── 08	4									
08 ├── 12	5									
12 ├── 16	7									
16 ├── 20	8									
20 ├── 24	6									
24 ├── 28	3									
28 ├── 32	2									
å	**36**									

RESPOSTAS:

A) $m'_1 = 16{,}33$; $m'_2 = 315{,}11$;

B) ${}_{10}m_2 = 88{,}44$; ${}_{10}m_3 = 1.178{,}67$;

C) $m_2 = 48{,}44$; $m_3 = 54{,}67$; $m_4 = 5.548{,}44$; **(Fazendo a média = 16)**

CAPÍTULO 9

Medidas de Assimetria e Curtose

As medidas de assimetria e curtose são as que restam para completarmos o quadro das estatísticas descritivas, que proporcionam, juntamente com as medidas de posição e de dispersão, a descrição e compreensão completas da distribuição de freqüências estudada.

Como já foi dito anteriormente, as distribuições de freqüências não diferem apenas quanto ao valor médio e à variabilidade, como também quanto à sua forma. Do ponto de vista desse último aspecto, as características mais importantes são o ***grau de deformação*** (assimetria) e o ***grau de achatamento ou afilamento*** (curtose) da curva de freqüências ou do histograma. Porém, para estudar as medidas de assimetria e curtose, é necessário o conhecimento de certas quantidades conhecidas como momentos. *(OBS.: Obrigatoriamente, não)*

9.1– Medidas de Assimetria ou Enviesamento (Relação entre $\overline{C}$, $\tilde{C}$ e $\hat{C}$)

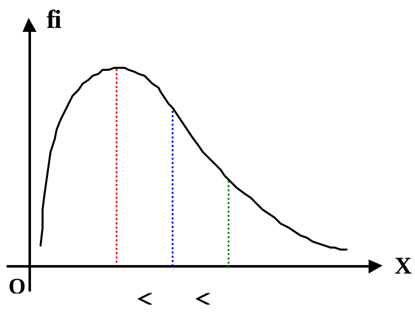
Assimetria, como o próprio nome insinua, significa desvio ou afastamento da simetria. Em outros termos, assimetria é o grau de deformação de uma curva de freqüências.

Para dados agrupados, representados por uma curva de freqüência, as diferenças entre os valores da média, da mediana e da moda, são indicadores da forma da curva em termos de assimetria, ou seja:

9.1.1– Distribuição Com Assimetria Positiva ou Distribuição Unimodal Positivamente Assimétrica

Neste caso a média apresenta um valor maior do que a mediana, e esta será maior do que a moda.

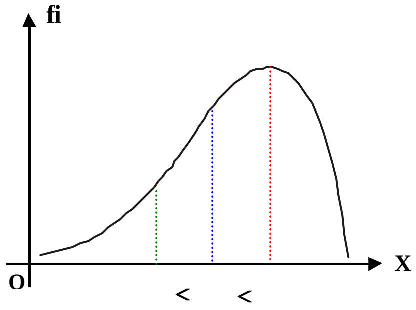
Graficamente, tende à direita:



9.1.2– Distribuição Com Assimetria Negativa ou Distribuição Unimodal Negativamente Assimétrica.

Neste caso a média apresenta um valor menor do que a mediana e esta será menor do que a moda.

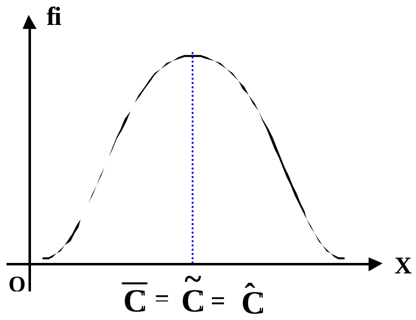
Graficamente tende à esquerda



9.1.3– Distribuição Simétrica ou Distribuição Unimodal Simétrica.

Neste caso, média, mediana e moda apresentam o mesmo valor.

Graficamente é proporcional



9.1.4– Coeficiente de Assimetria

Mede o grau de deformação da curva de distribuição por freqüência.

Segundo Pearson podemos avaliar este grau de deformação pelos seguintes processos:

a) Primeiro Coeficiente: mede a distância, em número de Desvios Padrão, da média até a moda.

$$\boxed{AS = \frac{\overline{C} - \hat{C}_P}{S}}$$

b) Segundo Coeficiente: mede a distância, em número de Desvios Padrão, da média até a mediana.

$$\boxed{AS = \frac{3 \times \left(\overline{C} - \tilde{C}\right)}{S}}$$

Em função do resultado é possível determinar o comportamento de cada curva, portanto, se:

AS = 0, então a distribuição é simétrica

AS > 0, então a distribuição é positivamente assimétrica

AS < 0, então a distribuição é negativamente assimétrica

Exemplo: A

Primeiro Coeficiente de Assimetria

$\overline{X} = 5{,}00$

$\hat{X}_P = 5{,}00 \qquad AS = \dfrac{5{,}00 - 5{,}00}{2{,}70} = 0{,}00$ (AS = 0; a distribuição é simétrica)

$S = 2{,}70$

Segundo Coeficiente de Assimetria

$\overline{X} = 5{,}00$

$\tilde{X} = 5{,}00 \qquad AS = \dfrac{3 \cdot (5{,}00 - 5{,}00)}{2{,}70} = 0{,}00$ (AS = 0; a distribuição é simétrica)

$S = 2{,}70$

Exemplo: B

Primeiro Coeficiente de Assimetria

$\overline{X} = 3{,}89$

$\hat{X}_P = 2{,}72 \qquad AS = \dfrac{3{,}89 - 2{,}72}{2{,}70} = 0{,}43$ (AS > 0; distribuição assimétrica positiva)

$S = 2{,}70$

Segundo Coeficiente de Assimetria

$\overline{X} = 3{,}89$

$\tilde{X} = 3{,}50 \qquad AS = \dfrac{3 \cdot (3{,}89 - 3{,}50)}{2{,}70} = 0{,}43$ (AS > 0; distribuição assimétrica positiva)

$S = 2{,}70$

Exemplo: C

Primeiro Coeficiente de Assimetria

$\overline{X} = 6{,}11$

$\hat{X}_P = 7{,}28 \qquad AS = \dfrac{6{,}11 - 7{,}28}{2{,}70} = -0{,}43$ (AS < 0; distribuição assimétrica negativa)

$S = 2{,}70$

Segundo Coeficiente de Assimetria

$\overline{X} = 6{,}11$

$\tilde{X} = 6{,}50 \qquad AS = \dfrac{3 \cdot (6{,}11 - 6{,}50)}{2{,}70} = -0{,}43$ (AS < 0; distribuição assimétrica negativa)

$S = 2{,}70$

9.2– Coeficiente Momento de Assimetria (e_M)

Outra medida utilizada par avaliar a assimetria de uma distribuição de freqüências é o coeficiente momento de assimetria, calculado com base nos momentos centrados da segunda, terceira, e quarta ordens, que é definido por:

$$\boxed{e_{M_1} = \frac{\sqrt{b_1}(b_2+3)}{2(5b_2 - 6b_1 - 9)}} \quad \text{ou} \quad \boxed{e_{M_2} = \sqrt{b_1} = \frac{m_3}{\left(\sqrt{m_2}\right)^3} = \frac{m_3}{S^3}}$$

Onde: $b_1 = \dfrac{m_3^2}{m_2^3}$ e $b_2 = \dfrac{m_4}{m_2^2}$ Onde: $m_2 = S^2$ e $\sqrt{m_2} = S$

Exemplo: (Os momentos abaixo já foram calculados anteriormente – página 71)

Classes	fi	PM	PMfi	$(PM-\overline{X})fi$	$(PM-\overline{X})^2fi$	$(PM-\overline{X})^3fi$	$(PM-\overline{X})^4fi$
10 ⊣ 20	2	15	30	(-30).2=-60	1.800	-54.000	1.620.000
20 ⊣ 30	4	25	100	(-20).4=-80	1.600	-32.000	640.000
30 ⊣ 40	5	35	175	(-10).5=-50	500	-5.000	50.000
40 ⊣ 50	8	45	360	0.8 = 0	0	0	0
50 ⊣ 60	5	55	275	10.5 = 50	500	5.000	50.000
60 ⊣ 70	4	65	260	20.4 = 80	1.600	32.000	640.000
70 ⊣ 80	2	75	150	30.2 = 60	1.800	54.000	1.620.000
å	30	-	1.350	0	7.800	0	4.620.000

$$\overline{X} = \frac{1.350}{30} = 45$$

$$m_1 = \frac{0}{30} = 0 \qquad m_2 = \frac{7.800}{30} = 260 \qquad m_3 = \frac{0}{30} = 0 \qquad m_4 = \frac{4.620.000}{30} = 154.000$$

$$b_1 = \frac{0}{260^3} = 0 \qquad \text{e} \qquad b_2 = \frac{154.000}{260^2} = \frac{154.000}{67.600} = 2,278$$

$$e_{M_1} = \frac{\sqrt{0}\cdot(2,278+3)}{2\cdot(5\cdot 2,278 - 6\cdot 0 - 9)} = \frac{0}{2\cdot(11,39 - 0 - 9)} = \frac{0}{2\cdot(2,39)} = \frac{0}{4,78} = 0$$

$$e_{M_2} = \sqrt{0} = \frac{0}{\left(\sqrt{260}\right)^3} = 0$$

Exercícios:

1º) Calcular o momento de assimetria (pelas duas fórmulas) da distribuição dos coeficientes de inteligência, da tabela abaixo:

Quoc. Int.	fi	PM	PM.fi	PM- $\bar{X}$	(PM- $\bar{X}$).fi	(PM- $\bar{X}$)^2.fi	(PM- $\bar{X}$)^3.fi	(PM- $\bar{X}$)^4.fi
68 a 72	4	70	280	-25,97	-103,87	2.697,07	-70.033,95	1.818.548,14
72 a 76	9	74	666	-21,97	-197,70	4.342,81	-95.397,06	2.095.555,41
76 a 80	16	78	1.248	-17,97	-287,47	5.164,82	-92.794,56	1.667.208,917
80 a 84	28	82	2.296	-13,97	-391,07	5.461,90	-76.284,51	1.065.440,26
84 a 88	45	86	3.870	-9,97	-448,50	4.470,05	-44.551,50	444.029,93
88 a 92	66	90	5.940	-5,97	-393,80	2.349,67	-14.019,72	83.650,98
92 a 96	85	94	7.990	-1,97	-167,17	328,76	-646,56	1.271,57
96 a 100	72	98	7.056	2,03	146,40	297,68	605,28	1.230,74
100 a 104	54	102	5.508	6,03	325,80	1.965,66	11.859,48	71.552,21
104 a 108	38	106	4.028	10,03	381,27	3.825,38	38.381,27	385.092,06
108 a 112	27	110	2.970	14,03	378,90	5.317,23	74.618,46	1.047.145,74
112 a 116	18	114	2.052	18,03	324,60	5.853,62	105.560,28	1.903.603,73
116 a 120	11	118	1.298	22,03	242,37	5.340,15	117.661,21	2.592.468,60
120 a 124	5	122	610	26,03	130,17	3.388,67	88.218,43	2.296.619,89
124 a 128	2	126	252	30,03	60,07	1.804,00	54.180,20	1.627.212,01
å	**480**		**46.064**	**30,50**	**0,00**	**52.607,47**	**97.356,76**	**17.100.630,18**

Respostas:

Média	=	95,9667
Momento Centrado na Média de primeira ordem (m_1)	=	Zero
Momento Centrado na Média de segunda ordem (m_2)	=	119,5989
Momento Centrado na Média de terceira ordem (m_3)	=	202,8266
Momento Centrado na Média de quarta ordem (m_4)	=	35.625,3129
Valor de b_1	=	0,0332487
Valor de b_2	=	2,9659141
Valor do C. Momento de Assimetria 1ª fórmula	=	0,0835952
Valor do C. Momento de Assimetria 2ª fórmula	=	0,1967829

2º) Com base nas tabelas da página seguinte, calcule:

a) Média
b) Mediana
c) Moda de Pearson
d) Desvio Padrão
e) Momento absoluto de 1ª ordem
f) Momento de 2ª ordem em relação à origem $X_o = 15$
g) Momento centrado de 1ª ordem
h) Momento centrado de 2ª ordem
i) Momento centrado de 3ª ordem
j) Momento centrado de 4ª ordem
k) Coeficiente de Variação
l) Coeficiente de Assimetria
m) Coeficiente Momento de Assimetria – 1ª fórmula
n) Coeficiente Momento de Assimetria – 2ª fórmula

Tabela ‘A’

Receita	fi									
150 a 200	2									
200 a 250	6									
250 a 300	7									
300 a 350	12									
350 a 400	10									
400 a 450	8									
450 a 500	5									
S	50									

Tabela ‘B’

Peso (kg)	fi									
2 a 4	2									
4 a 6	6									
6 a 8	14									
8 a 10	10									
10 a 12	8									
12 a 14	6									
14 a 16	4									
S	50									

Tabela ‘C’

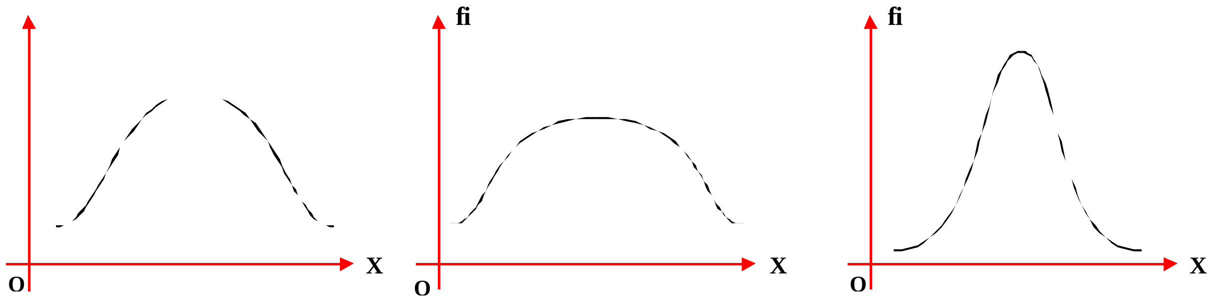
Classes	fi									
10 a 20	2									
20 a 30	6									
30 a 40	7									
40 a 50	12									
50 a 60	10									
60 a 70	8									
70 a 80	5									
S	50									

9.3– Medidas de Curtose

A curtose ou excesso indica até que ponto a curva de freqüências de uma distribuição se apresenta mais afilada ou mais achatada do que uma curva-padrão, denominada curva normal.

Observação: A curva normal (curva padrão), apresenta um coeficiente de curtose igual a 0,263 e recebe o nome de mesocúrtica. Se o coeficiente for maior que 0,263 recebe o nome de platicúrtica e se menor que 0,263 chamará leptocúrtica.

Graficamente:



Normal ou Mesocúrtica
C = 0,263

Platicúrtica ou Achatada
C > 0,263

Leptocúrtica ou Afilada
C < 0,263

9.3.1– Cálculo do Coeficiente de Curtose

$$\boxed{C = \frac{\frac{Q_3 - Q_1}{2}}{C_{90} - C_{10}}} \quad \textit{ou} \quad \boxed{C = \frac{D_q}{C_{90} - C_{10}}} \quad \textit{ou} \quad \boxed{C = \frac{Q_3 - Q_1}{2 \times (C_{90} - C_{10})}}$$

Exemplo: Com os dados abaixo calcule o coeficiente de curtose

$Q_3 = 6{,}75 \qquad C_{90} = 8{,}20$

$Q_1 = 3{,}25 \qquad C_{10} = 1{,}80$

$C = \dfrac{6{,}75 - 3{,}25}{2 \cdot (8{,}20 - 1{,}80)} = 0{,}273$ portanto, **curva platicúrtica**

9.4– Coeficiente Momento de Curtose

O coeficiente momento de curtose utiliza-se do quociente entre o momento centrado de quarta ordem e o quadrado do momento centrado de segunda ordem (variância), ou seja:

$$\boxed{b_2 = \frac{m_4}{m_2^2}} \quad \textbf{ou} \quad \boxed{b_2 = \frac{m_4}{S^4}}$$

Observação: A quantidade b_2 aparece na fórmula do coeficiente momento de assimetria apresentado anteriormente (item 9.2)

Se $b_2 = 3$ ➔ distribuição ou curva mesocúrtica

Se $b_2 < 3$ ➔ distribuição ou curva platicúrtica

Se $b_2 > 3$ ➔ distribuição ou curva leptocúrtica

Exemplo: Calcular o coeficiente momento de curtose da distribuição dos coeficientes de inteligência (página 77)

Recorreremos aos seguintes valores já calculados (página 77)

$m_4 = 35.626,31$

$m_2 = S^2 = 109,60$ ➔ $m_2^2 = 12.012,16$

Então,

$$b_2 = \frac{m_4}{m_2^2} = \frac{35.626,31}{12.012,16} = 2,96 \cong 3,0 \longrightarrow \text{Distribuição Mesocúrtica}$$

A curtose pode ser medida igualmente pelo coeficiente c_2:

$$\boxed{\mathbf{c_2 = b_2 - 3}}$$

Teríamos então:

$c_2 = 0$ ➔ distribuição ou curva mesocúrtica

$c_2 < 0$ ➔ distribuição ou curva platicúrtica

$c_2 > 0$ ➔ distribuição ou curva leptocúrtica

Exercício

Os dados abaixo referem-se a renda nominal de 60 famílias (valores em R$1,00)

400	**350**	**370**	**375**	**399**	**405**	**360**	**408**	**430**	**385**
390	**390**	**385**	**360**	**397**	**400**	**406**	**440**	**415**	**410**
400	**410**	**382**	**340**	**360**	**370**	**380**	**370**	**413**	**390**
400	**390**	**355**	**375**	**427**	**397**	**413**	**430**	**357**	**340**
350	**410**	**420**	**360**	**403**	**382**	**390**	**425**	**420**	**404**
420	**410**	**411**	**404**	**397**	**420**	**404**	**421**	**440**	**380**

Pede-se:

a) montar uma tabela de distribuição por freqüência usando a fórmula de "Sturges";

b) a média aritmética; **(R$ 395,75)**

c) a mediana; **(R$ 400,00)**

d) a moda de Karl Pearson; **(R$ 408,50)**

e) o primeiro e o terceiro quartis; o décimo e o nonagésimo percentis; **(R$ 378,33; R$412,50; R$ 360,00; R$ 426,25)**

f) o desvio médio; **(R$ 19,75)**

g) o desvio quartílico; **(R$ 17,09)**

h) construir um gráfico box plot;

i) o desvio padrão (processo longo e breve); **(R$ 24,09)**

j) o percentual de famílias cujas rendas situam-se entre a $\overline{X} \pm 1\,\mathrm{DM}$; **(56,67%)**

q) o número de famílias cujas rendas situam-se entre a $\overline{X} \pm 1\,S$ (suponha que a distribuição seja normal). **(± 41 famílias)**

l) coeficiente de variação; **(6,09%)**

m) o erro padrão da média; **(R$ 3,11)**

n) os momentos de primeira, segunda, terceira e quarta ordens centrados na média; **(R$ 0,00; R$ 570,69; R$ – 4.192,28; R$ 776.752,02)**

o) o coeficiente de assimetria de Pearson (primeiro e segundo). Faça a curva e determine o seu comportamento; **(– 0,529; Assimetria Negativa)**

p) o coeficiente momento de assimetria; **(0,09)**

q) determinar o grau de achatamento da curva, identificando-a pelo nome e pela curva; **(0,258; Leptocúrtica)**

r) o coeficiente momento de curtose; **(R$ 2,38)**

OBSERVAÇÃO: Trabalhe sempre com duas casas depois da vírgula, fazendo arredondamento estatístico, onde for possível.

CAPÍTULO 10

Probabilidade

10.1- Conceito e Caracterização

A Teoria das Probabilidades estuda os fenômenos aleatórios com vários resultados possíveis, quantificando as suas possibilidades de ocorrência. Com base na teoria das probabilidades, jamais será possível dizer o que vai ocorrer num experimento aleatório – pois isso dependerá sempre do acaso; no entanto, ela permite prever o que pode ocorrer e ainda dimensiona a chance de ocorrência de cada uma das possibilidades. Entende-se por "chance" a medida da ocorrência das circunstâncias favoráveis.

A utilização de probabilidades indica que existe um elemento do acaso, ou de incerteza, quanto à ocorrência ou não de um evento futuro. Assim é que, em muitos casos, pode ser praticamente impossível afirmar por antecipação o que ocorrerá; mas é possível dizer o que pode ocorrer. Assim sendo, as probabilidades servem para exprimir a chance de ocorrência de um determinado evento.

O termo probabilidade é usado de modo muito amplo na conversação diária para sugerir um certo grau de incerteza sobre o que ocorreu no passado, o que ocorrerá no futuro ou o que está ocorrendo no presente. O torcedor de certo time pode apostar contra ele porque sua "probabilidade" de ganhar é pequena. O aluno poderá ficar contente porque acha que sua "probabilidade" de obter bons resultados nas provas é grande.

A idéia de probabilidade desempenha papel importante em muitas situações que envolvam uma tomada de decisão. Suponhamos que um empresário deseja lançar um novo produto no mercado. Ele precisará de informações sobre a "probabilidade" de sucesso para seu novo produto.

10.2- Terminologias

Toda teoria está fundamentada numa terminologia própria. No caso da teoria das probabilidades, são usadas as seguintes expressões:

a) Experimento
É o fato ou fenômeno que está sendo estudado.

Ex.: O lançamento de uma moeda; a extração de uma carta de um baralho; a análise do clima e o estudo da economia.

b) Espaço Amostral ou Conjunto Universo (W = S = Space**)**
É o conjunto de todos os resultados possíveis de um certo experimento.

Ex.: 1º) Se o evento consistir no lançamento de um dado, qual será o espaço amostral?

W = {1, 2, 3, 4, 5, 6} ou seja **W =**

2º) Se o experimento consistir no lançamento de 3 moedas consecutivas, qual será o espaço amostral?

c) Evento Elementar
É cada um dos resultados possíveis do espaço amostral do experimento.

Ex.: Cara, num lançamento de moeda; Ás, numa extração de carta.

Classificação dos eventos:
Dois ou mais eventos elementares de certo espaço amostral são ditos:

1º) Evento Simples
Formado por um único elemento do espaço amostral (**W = S**).

Ex.: ocorrência da face 3 no lançamento de um dado.

2º) Evento Composto
Formado por mais de um elemento do espaço amostral (**W = S**).

Ex.: ocorrência de face par no lançamento de um dado.

3º) Evento Certo
É aquele que ocorre em qualquer realização do experimento.

Ex.: no lançamento de um dado fatalmente sairá a face 1, 2, 3, 4, 5 ou 6

4º) Evento Impossível
É aquele que não ocorre em qualquer realização do experimento.

Ex.: No lançamento de um dado sair a face 7

5º) Evento Complementar
Para um evento **A** qualquer, o complementar de **A**, denotado por $\overline{\mathbf{A}}$ é dado por $\overline{\mathbf{A}} = \mathbf{S} - \mathbf{A}$, ou seja, é um outro conjunto formado pelos elementos que pertencem a **S** e não pertencem a **A**. O resultado da reunião de **A** e $\overline{\mathbf{A}}$ é exatamente o espaço amostral.

Ex.: Coroa é complementar de cara (e vice-versa); o conjunto de cartas de paus, ouros e copas é complementar do conjunto de espadas.

6º) Evento Mutuamente Exclusivo
Caracteriza-se quando dois ou mais eventos não podem ocorrer simultaneamente, ou seja, a ocorrência de um exclui a possibilidade de ocorrência do outro e vice-versa.

Ex.: 1º) Se a carta é de copas, então ela não é de ouro; se o tempo está nublado, então não há sol.

Ex.: 2º) Ocorrência de face menor que 2 ou maior que 5 no lançamento de um dado.

P (face menor que 2) = $\frac{1}{6}$ = 0,167 ou 16,7%

P (face maior que 5) = $\frac{1}{6}$ = 0,167 ou 16,7%

Observe que os eventos “face menor que 2 ou face maior que 5” são mutuamente excludentes, pois a ocorrência de um, impossibilita a ocorrência do outro. Porém, não são complementares, pois não esgotam todos os resultados possíveis do experimento.
Eventualmente poderão esgotar todos os resultados possíveis, nesse caso serão chamados de **mutuamente excludentes e exaustivos**.

7º) Evento Independente
Dizemos que dois ou mais eventos são independentes quando não exercem ações recíprocas, comportando-se cada um de maneira que lhe é própria sem influenciar os demais.
Caracteriza-se, portanto, quando a ocorrência de um evento não for afetada pela ocorrência do outro, sendo a recíproca verdadeira.

Ex.: Consideremos o lançamento de duas moedas:
Temos: S = {Ca, Ca; Ca, Co; Co, Co; Co, Ca}
Os resultados dos eventos são independentes de uma moeda para outra.

8º) Evento Condicionado

Quando associados dois ou mais eventos a um experimento aleatório qualquer dizemos que eles são condicionados a outro evento B do mesmo experimento.
Caracteriza-se quando a ocorrência de um evento A qualquer dependa da ocorrência de outro evento B.

Ex.: 1º) retirada, sem reposição, de duas cartas vermelhas de um baralho completo.

Ex.: 2º) uma caixa contém 20 bolas numeradas de 1 a 20. Pretendo sortear a bola 5 e a bola 8. Tiro uma bola e verifico que é a bola 8, ________________________________
____
____

10.3- Regras para combinação de Probabilidades

Todos os problemas que envolvem probabilidades podem ser resolvidos, basicamente a partir de dois teoremas fundamentais.

10.3.1- Teorema da adição (ou)

a) para eventos mutuamente exclusivos

P (A ou B) = P (A) + P (B)

Ex.: Qual a probabilidade de ocorrer "dama" ou "valete" ao retirarmos uma única carta de um baralho completo ?

b) para eventos não mutuamente exclusivos

P (A ou B) = P (A) + P (B) – P (A . B)

Ex.: Qual a probabilidade de ocorrer "ás" ou "carta de ouros" ao retirarmos um única carta de um baralho completo ?

10.3.2- Teorema da multiplicação (e)

a) para eventos condicionados

P (A e B) = P (A) . [P (A) . P (B / A)]

Ex.: Retira-se, sem reposição, duas cartas de um baralho completo, qual a probabilidade de ambas serem espadas ?

b) para eventos independentes

P (A e B) = P (A) . P (B)

Ex.: Qual a probabilidade de acertarmos os dois primeiros jogos da loteria esportiva utilizando palpite simples ?

10.4- Axiomas das Probabilidades

A probabilidade de um evento **A** qualquer, denotada por **P (A)**, é dada por um quociente em que o numerador é o número de casos favoráveis à ocorrência do evento, e o denominador o número de casos possíveis (espaço amostral), ou seja, a probabilidade simples de um evento acontecer é a razão entre o número de casos favoráveis e o número de casos possíveis

$$\boxed{\mathbf{Pr = P(A) = P(e) = P = \frac{h}{n}}}$$

onde:
h = número de casos favoráveis
n = número de casos possíveis

Ex.: Joga-se um dado uma vez, qual a probabilidade de sair a face 2 ?

$$P(2) = \frac{1}{6} = 0{,}1666 \cong 16{,}67\%$$

Observação: **a)** A probabilidade de um evento certo é igual a 1 ou 100%

Ex.: $P(E) = \frac{6}{6} = 1 = 100\%$

b) A probabilidade **(P)** de sair face 2 no lançamento de um dado é 1/6, obviamente, a probabilidade contrária **(q)**, ou seja, a de não sair a face 2 é 5/6, portanto:

$$q = \frac{n-h}{n} = \frac{5}{6} \longrightarrow q = \frac{n}{n} - \frac{h}{n} \longrightarrow q = 1 - Pr \;\therefore\; Pr + q = 1$$

Logo: 1º Axioma ➔ $0 \le Pr \le 1$
2º Axioma ➔ Pr(S) = 1
3º Axioma ➔ Se A e B são eventos mutuamente exclusivos, então:
P(A E B) = P(A) + P(B)

<u>Exercícios de aplicação</u>:

01) Uma caixa contém bolas numeradas de 1 a 15. Retira-se uma bola ao acaso, qual a probabilidade de se obter uma bola cujo número seja múltiplo de 4?

02) Retira-se uma carta de um baralho de 52 cartas. Qual a probabilidade de se obter um rei de ouros?

03) Joga-se um dado duas vezes. Qual a probabilidade de se obter a face 5 em ambas as jogadas?

04) Uma caixa contém 100 peças das quais 5 são defeituosas. Seleciona-se ao acaso uma peça, que não é recolocada e, seleciona-se outra peça. Qual a probabilidade de que ambas as peças retiradas sejam defeituosas?

10.5- Análise Combinatória

Quando a contagem direta do número de possibilidades é muito trabalhosa, podemos nos valer da análise combinatória para determinar os números de casos favoráveis e/ou possíveis dos experimentos estudados. Para tanto, é necessário primeiro classificar adequadamente o agrupamento e depois aplicar a fórmula correta.

10.5.1-Permutação

Permutar é (re)ordenar os elementos de um conjunto numa seqüência previamente definida. As permutações podem ser:

a) **Permutação sem Repetição**
Conjuntos com elementos distintos

Ex.: Três membros de uma organização social se ofereceram como voluntários, para compor a diretoria, para o próximo ano, assumindo as funções de Presidente, Tesoureiro e Secretário. Qual o número de maneiras pelas quais os três podem assumir tais cargos ?

b) **Permutação com Repetição**
Conjuntos com alguns elementos iguais entre si.

Ex.: Quantos anagramas distintos têm na palavra MISSISSIPI? (MILONE: 120)

Exercício de aplicação:

Quantos números de 4 algarismos podem ser formados com os 10 algarismos 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9: (SPIEGEL:31)
a) admitindo-se repetições;
b) não se admitindo repetições;
c) não se admitindo repetições e o último algarismo devendo ser zero.

10.5.2- Arranjo

Quando a permutação é feita com apenas uma parte dos elementos do espaço amostral, ou seja, por arranjo se entende o número total de permutações possíveis nos subconjuntos de *‘r’* elementos de um conjunto composto por *‘n’* elementos. Os arranjos são calculados pela seguinte expressão:

$$\boxed{A_n^r = \frac{n!}{(n - r)!}}$$

Onde: n = número total de elementos
r = o que se pretende agrupar

Ex.: Com base no exemplo anterior, suponha que existam 10 membros na organização social e que nenhuma indicação tenha sido feita para os cargos de Presidente, Tesoureiro e Secretário. Qual é o número, de diferentes disposições de 3 membros eleitos entre os 10 membros do clube, que poderão ser formadas?

Exercício de aplicação:

Cinco pessoas constituem a junta diretora de uma empresa. Suponha que somente três destes diretores sejam convidados a representar a empresa num banquete.
Quantos arranjos diferentes seriam possíveis para compor este trio?

10.5.3- Combinação

Quando as escolhas se distinguem só pela qualidade e não pela ordem dos elementos.

Para se calcular o número de combinações possíveis de *‘n’* elementos, tomando-se *‘r’* de cada vez, usa-se a expressão:

$$\boxed{C_n^r = \frac{n!}{r!(n - r)!}}$$

Ex.: Suponhamos que 3 membros de uma pequena organização social, de 10 membros, venham a ser escolhidos para formar uma comissão.
Qual é o número, de diferentes grupos de 3 pessoas, que pode ser escolhido ?

<u>Exercício de aplicação</u>:

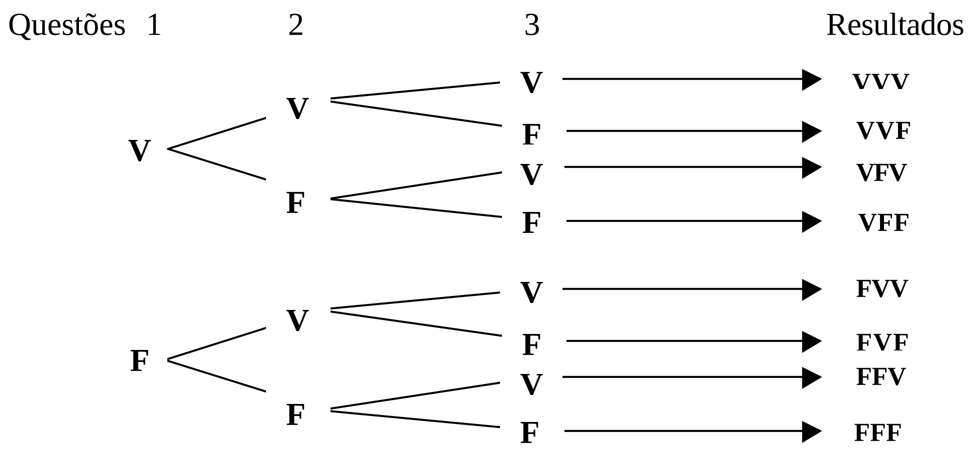
De 5 matemáticos e 7 físicos deve-se constituir uma comissão de 2 matemáticos e 3 físicos. De quantas maneiras podemos formar a comissão se: (SPIEGEL:33)

a) qualquer matemático e qualquer físico pode ser incluído;
b) determinado físico deve fazer parte da comissão;
c) dois determinados matemáticos não devem pertencer à comissão.

10.6- Diagrama da Árvore

Quando o número de "pontos" do espaço amostral é relativamente pequeno, é possível a sua contagem direta, utilizando o chamado diagrama da árvore (ou de decisão), que consiste em representar graficamente todas as possíveis variantes de uma dada situação. Recebe esse nome porque sua figura característica se assemelha a uma árvore, com suas ramificações partindo de cada uma das possibilidades originais e intermediárias. Apesar de ser um processo gráfico facilmente "mecanizável", sua aplicação se restringe a eventos simples, uma vez que a sua complexidade é diretamente proporcional ao número de possibilidades de ramificação.

Ex.: Um estudante deve responder um teste do tipo verdadeiro (V) ou falso (F). Se considerarmos apenas três questões, qual a probabilidade dele acertar todo o teste?



O diagrama nos diz que ele tem uma chance a favor e sete contra, se considerarmos apenas três questões, portanto, 1/8 = 12,5%.

Outra forma, é partir da definição de probabilidade. Com base nela podemos esboçar a seguinte solução analítica:

Número de Questões	Respostas Possíveis	Probabilidade de Acerto
1	**V ou F**	1/2 = 50,0%
2	**VV, VF, FV ou FF**	1/4 = 25,0%
3	**VVV, VVF, VFF, VFV, FVF, FVV, FFV ou FFF**	1/8 = 12,5%

Como se vê, a resposta do diagrama foi aqui confirmada.

<u>Exercício de aplicação</u>:

Uma caixa A contém 10 peças perfeitas e 3 defeituosas. Outra caixa B contém 8 peças perfeitas e 5 defeituosas. Sorteando-se uma das caixas ao acaso, qual a probabilidade de que seja retirada uma peça defeituosa?

10.7- Teorema de Bayes (Reverendo “Thomas Bayes” – 1702 / 1761)

Com base na probabilidade condicional, é possível calcular a probabilidade de um dado evento B ocorrer após certo evento A ter ocorrido. O que o teorema de Bayes possibilita é a quantificação de certo evento A ter sido provocado por B, C ou D. Assim, se A pode ter sido provocado por B, C ou D e quer-se quantificar a chance dele ter sido produzido por D em particular, basta relacionar as chances de produção por D em relação à chance de ter sido produzido por B, C ou D.

$$\boxed{P(A/D) = \frac{P(A) \times P(D/A)}{P(A) \times P(D/A) + P(B) \times P(D/B) + P(C) \times P(D/C)}}$$

<u>Exercício de aplicação:</u>

Uma peça é manufaturada por 3 fábricas.
- A fábrica 1 produz o dobro da 2.
- 2% das peças da fábrica 1 são defeituosas.
- 2% das peças da fábrica 2 são defeituosas.
- 4% das peças da fábrica 3 são defeituosas.
- As fábricas 2 e 3 produzem o mesmo número de peças.

Uma peça é extraída ao acaso e é defeituosa.
Qual a probabilidade da peça ser da fábrica 3?

EXERCÍCIOS

ESPAÇO AMOSTRAL

1º) Se o experimento consistir em acertar um alvo com 3 tiros consecutivos, qual será o espaço amostral? (Demonstre como chegar ao resultado)

EVENTOS

2º) A probabilidade de um indivíduo A estar vivo daqui a 10 anos é de 70%. A probabilidade de outro indivíduo B estar vivo daqui a 10 anos é 20%. Qual a probabilidade de que daqui a 10 anos:
a) ambos estejam vivos?;
b) ambos estejam mortos?;
c) pelo menos 1 esteja vivo?;
d) pelo menos 1 esteja morto?

3º) Uma caixa contém 100 peças das quais 5 são defeituosas. Seleciona-se ao acaso uma peça e depois de recolocada seleciona-se outra. Qual a probabilidade de que ambas as peças retiradas sejam defeituosas?

4º) Uma sala tem 12 moças e 4 rapazes. Sorteando-se ao acaso uma comissão de 3 pessoas, qual a probabilidade de que a comissão seja formada só por moças?

5º) Retira-se, ao acaso, uma carta de um baralho de 52 cartas. Qual a probabilidade de se obter um "ás" ou um "rei"?

6º) Retira-se, ao acaso, uma carta de um baralho de 52 cartas. Qual a probabilidade de se obter um "valete" ou uma carta de "copas"?

7º) De um baralho comum de 52 cartas retirou-se uma carta, verificando-se que é vermelha. Qual a probabilidade de essa carta ser uma figura?

8º) Uma urna possui 8 bolas, sendo 6 azuis e 2 verdes.
Sorteando-se uma bola ao acaso, qual a probabilidade de se retirar:
a- uma bola azul;
b- uma bola verde.

9º) Na jogada de um dado honesto determine a probabilidade de se obter:
a- um número par;
b- um número menor do que 3;
c- um número maior ou igual a 3;
d- um número maior do que 6;
e- um número menor do que 10.

10º) Determine a probabilidade de se jogar um dado e se obter:
a- um número par ou múltiplo de 3;
b- um número ímpar ou múltiplo de 6.

11º) Extraindo uma carta de um baralho honesto, determine a probabilidade de se obter:
a- um rei ou uma carta de ouros;
b- uma figura ou uma carta de ouros;
c- uma figura ou um sete;
d- um sete de ouros, ou um rei ou uma dama vermelha.

12º) Uma urna contém 20 bolas (10 azuis, 6 verdes e 4 vermelhas).
Sorteando-se bolas sucessivamente, determine a probabilidade de se obter:
a- a 1ª azul, a 2ª verde sem reposição da 1ª ;
b- a 1ª azul, a 2ª verde com reposição da 1ª ;
c- a 1ª azul, a 2ª verde, a 3ª vermelha, com reposição da 1ª e da 2ª .

13º) Um casal pretende ter 4 filhos. Considere iguais a ½a probabilidade de se ter um filho do sexo masculino ou do sexo feminino.
Determine a probabilidade do casal ter:
a- 4 Mulheres
b- 3 Mulheres e 1 Homem em qualquer ordem
c- 2 Mulheres e 2 Homens em qualquer ordem
d- 1 Mulher e 3 Homens em qualquer ordem
e- 4 Homens
f- Homem, Mulher, Homem, Mulher, nesta ordem

14º) Uma urna contém 15 bolas numeradas de 1 a 15. Retira-se uma bola.
Qual a probabilidade de se obter uma bola cujo número seja múltiplo de 3 ou 7?

15º) Numa caixa contendo 4 bolas vermelhas, 3 bolas azuis e 3 bolas brancas, queremos saber qual a probabilidade de retirarmos 1 bola vermelha, e sem reposição uma bola branca?

16º) Uma rifa composta por 15 números irá definir o ganhador de dois prêmios sorteados um de cada vez.
Se você adquiriu três números, qual a probabilidade de ganhar os dois prêmios?

17º) Duas bolas são retiradas, sem reposição, de uma urna que contém duas bolas brancas, três bolas pretas e cinco bolas vermelhas.
Determine a probabilidade de que:
a- ambas sejam pretas;
b- ambas sejam vermelhas;
c- ambas sejam da mesma cor;
d- ambas sejam de cores diferentes.

18º) Um piloto de Fórmula Um tem 50% de probabilidade de vencer determinada corrida, quando esta se realiza sob chuva. Caso não chova durante a corrida, sua probabilidade de vitória é de 25%. Se o Serviço de Meteorologia estimar em 30% a probabilidade de que chova durante a corrida, qual é a probabilidade deste piloto ganhar esta corrida?

19º) Um aluno propõe-se a resolver uma questão de um trabalho.
A probabilidade de que consiga resolver a questão sem necessidade de uma pesquisa é de 40%. Caso faça a pesquisa, a probabilidade de que consiga resolver a questão é de 70%.
Se a probabilidade do aluno fazer a pesquisa é de 80%, calcule a probabilidade de que consiga resolver a questão.

20º) Um pesquisador desenvolve sementes de quatro tipos de plantas, P1, P2, P3 e P4.
Plantados canteiros-pilotos destas sementes, a probabilidade de todas germinarem é de 40% para P1, 30% para P2, 25% para P3 e 50% para P4.
Um canteiro-piloto é selecionado ao acaso. Qual é a probabilidade de que todas as sementes plantadas tenham germinado?

21º) Lança-se um par de dados não viciados. Achar a probabilidade da soma ser maior ou igual a 10, se ocorrer face 5 no primeiro dado?

22º) Um lote contém 50 peças boas e 10 defeituosas. Uma peça é escolhida ao acaso e, sem reposição desta, outra peça é escolhida. Determine a probabilidade das duas peças serem defeituosas.

23º) Em uma urna, existem 10 bolas, sendo 3 vermelhas. Em outra urna, existem 12 bolas, sendo 4 vermelhas. Retira-se uma bola de cada urna. Determine a probabilidade de:
a- ambas serem vermelhas;
b- ao menos uma ser vermelha.

24º) Em um lote de 12 objetos existem 4 defeituosos. Seja o experimento retirar-se 2 objetos quaisquer e verificar se são ou não defeituosos. Determine as probabilidades de que:
a- ambos os objetos sejam defeituosos;
b- ambos os objetos não sejam defeituosos;
c- pelo menos um objeto seja defeituoso.

25º) Lançam-se três moedas não viciadas. Encontre a probabilidade de ocorrer três caras se:
a- não se tem nenhuma informação;
b- ocorre cara na primeira;
c- ocorre cara numa das moedas.

26º) São dadas três caixas, com os seguintes conteúdos:
a- a caixa I tem 10 lâmpadas, das quais 4 são defeituosas
b- a caixa II tem 6 lâmpadas, das quais 1 é defeituosa
c- a caixa III tem 6 lâmpadas, das quais 3 são defeituosas
Uma caixa é selecionada ao acaso e desta é escolhida uma lâmpada. Determine a probabilidade desta lâmpada ser defeituosa.

27º) A probabilidade que o aluno A resolva um certo problema é P(A) = 1/2, a que o aluno B o resolva é P(B) = 1/3, e a que o aluno C resolva é P(C) = 1/4. Qual a probabilidade de que:
a- os três resolvam o problema;
b- ao menos um resolva o problema.

28º) Um grupo de 50 moças é classificado de acordo com a cor dos cabelos e dos olhos de cada uma, segundo a tabela:

CABELOS	**OLHOS**	
	AZUIS	**CASTANHOS**
LOIRA	17	9
MORENA	4	14
RUIVA	3	3

a- Se você marca encontro com uma dessas garotas, escolhida ao acaso, qual a probabilidade de dela ser:
I- loira;
II- morena de olhos azuis;
III- morena ou ter olhos azuis.

b- Está chovendo quando você encontra a garota. Seus cabelos estão completamente cobertos, mas você percebe que ela tem olhos castanhos. Qual a probabilidade de que ela seja morena?

29º) Escolhido ao acaso um elemento do conjunto dos divisores positivos de 60, qual a probabilidade de que seja primo?

30º) Três pessoas A, B e C vão participar de um concurso num programa de televisão. O apresentador faz um sorteio entre A e B e, em seguida, faz um sorteio entre C e o vencedor do 1º sorteio, para decidir quem iniciará o concurso. Se em cada sorteio as duas pessoas têm a mesma "chance" de ganhar, qual é a probabilidade de A iniciar o concurso?

ANÁLISE COMBINATÓRIA

31º) Um representante de vendas deve visitar seis cidades durante uma viagem.

a- Se há dez cidades na área geográfica que vai visitar, quantos grupos diferentes de seis cidades pode ele visitar?

b- Suponhamos que existam dez cidades na região que ele visitará e suponhamos, também, que a seqüência das visitas programadas às cidades selecionadas seja importante. Quantas diferentes seqüências existem de seis cidades escolhidas de um grupo de dez?

c- Suponhamos que as seis cidades a visitar já tenham sido escolhidas, mas ainda não se tenha determinado a seqüência na qual serão feitas as visitas. Quantas seqüências existem para as seis cidades escolhidas?

32º) Das dez cidades descritas no problema anterior, suponhamos que seis sejam de fato mercados "primários" para o produto em questão, enquanto as outras quatro são mercados "secundários".
Se o vendedor escolhe aleatoriamente as seis cidades para visitar, qual a probabilidade de que:
a- quatro das cidades sejam mercados primários e dois secundários;
b- que todas as seis cidades sejam mercados primários?

33º) Quatro atletas participam de uma corrida. Quantos resultados existem para o 1º, 2º e 3º lugares?

34º) Dez pessoas, entre elas Antônio e Beatriz, devem ficar em fila. De quantas formas isto pode ser feito se Antônio e Beatriz devem ficar sempre juntos?

35º) Num determinado setor de um hospital, trabalham cinco médicos e dez enfermeiros. Quantas equipes distintas, construídas cada uma de um médico e quatro enfermeiros, podem ser formados nesse setor?

36º) Em um teste de múltipla escolha, com cinco alternativas distintas, sendo uma única correta, o número de modos distintos de ordenar as alternativas de maneira que a única correta não seja nem a primeira nem a última é:

37º) Numa prova oficial, de Fórmula Um, participarão 25 pilotos e, apenas os 6 primeiros colocados ganharão pontos. Considerando que todos os pilotos terão a mesma chance de classificação, qual é o número de maneiras diferentes que poderá ser formado o grupo daqueles que obterão pontos, sem levar em consideração a posição dos 6 primeiros colocados?

38º) Uma urna contém 10 bolas: 6 pretas iguais e 4 brancas iguais. Quantas são as maneiras diferentes de se extrair, uma a uma, as 10 bolas da urna?

39º) De quantas maneiras podemos dispor em uma fileira 5 fichas de cores distintas? (SPIEGEL:30)

40º) De quantas maneiras 10 pessoas podem sentar-se em um banco que só tem 4 lugares? (SPIEGEL:31)

41º) Deseja-se dispor 5 homens e 4 mulheres em fila, de modo que as mulheres ocupem os lugares pares. Quantos são os arranjos possíveis? (SPIEGEL:31)

42º) Suponhamos que 3 membros de uma pequena organização social, de 10 membros, venham a ser escolhidos para formar uma comissão. Se o grupo contém 6 mulheres e 4 homens, qual a probabilidade de que uma comissão escolhida, aleatoriamente, seja composta por 2 mulheres 1 homem?

43º) Uma sala tem 12 moças e 4 rapazes. Sorteando-se ao acaso uma comissão de 3 pessoas, qual a probabilidade de que a comissão seja formada só por moças?

DIAGRAMA DA ÁRVORE

44º) Três máquinas: A, B e C produzem, respectivamente, 50%, 30% e 20% do total de peças fabricadas na indústria.
As porcentagens de peças defeituosas dessas máquinas são, respectivamente, 3%, 4% e 5%.
Se uma peça é selecionada ao acaso de uma das máquinas, qual a probabilidade de que a peça seja defeituosa?

TEOREMA DE BAYES

45º) Temos 3 urnas:
- A contém 3 bolas vermelhas e 5 brancas
- B contém 2 bolas vermelhas e 1 branca
- C contém 2 bolas vermelhas e 3 brancas

Uma urna é sorteada ao acaso e uma bola é retirada.
Se a bola for vermelha, qual a probabilidade de que ela tenha vindo da urna A?

46º) Um pesquisador desenvolve sementes de quatro tipos de plantas, P1, P2, P3 e P4.
Plantados canteiros-pilotos destas sementes, a probabilidade de todas germinarem é de 40% para P1, 30% para P2, 25% para P3 e 50% para P4.

a- Escolhido um canteiro ao acaso, verificou-se que nem todas as sementes haviam germinado.
Calcule a probabilidade de que o canteiro escolhido seja o de sementes de P3.

b- Escolhido um canteiro ao acaso, verificou-se que todas as sementes haviam germinado.
Calcule a probabilidade de que o canteiro escolhido seja o de sementes de P1.

RESPOSTAS

1º) 8					
2º) a) 14%	b) 24%	c) 76%	d) 86%		
3º) 0,25%					
4º) 39,29%					
5º) 15,38%					
6º) 30,77%					
7º) 11,54%					
8º) a) 75%	b) 25%				
9º) a) 50%	b) 33,33%	c) 66,67%	d) 0%	e) 100%	
10º) a) 66,67%	b) 66,67%				
11º) a) 30,77%	b) 42,31%	c) 30,77%	d) 13,46%		
12º) a) 15,79%	b) 15%	c) 3%			
13º) a) 6,25%	b) 25%	c) 37,50%	d) 25%	e) 6,25%	f) 6,25%
14º) 46,67%					
15º) 13,33%					
16º) 2,86%					
17º) a) 6,67%	b) 22,22%	c) 31,11%	d) 68,89%		
18º) 32,50%					
19º) 64%					
20º) 36,25%					
21º) 5,56%					
22º) 2,54%					
23º) a) 10%	b) 53,33%				
24º) a) 9,09%	b) 42,42%	c) 57,58%			
25º) a) 12,50%	b) 25%	c) 25%			
26º) 35,56%					
27º) a) 4,17%	b) 75%				
28º) a) I) 52% II) 8% III) 76%	b) 53,85%				
29º) 25%					
30º) 25%					
31º) a) 210	b) 151.200	c) 720			
32º) a) 43%	b) 0,5%				
33º) 24					
34º) 362.880					
35º) 1.050					
36º) 60					
37º) 177.100					
38º) 210					
39º) 120					
40º) 5.040					
41º) 2.880					
42º) 50%					
43º) 39,29%					
44º) 3,7%					
45º) 26,01%					
46º) a) 29,41%	b) 27,59%				

GLOSSÁRIO DE FÓRMULAS

E S T A T Í S T I C A

01 – FÓRMULA DE "STURGES" PARA DETERMINAÇÃO DO NÚMERO DE CLASSES

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

02 – MÉDIA ARITMÉTICA SIMPLES PARA DADOS <u>NÃO</u> AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{\overline{X} = \frac{\Sigma Xi}{n}}$$

03 – MÉDIA ARITMÉTICA PARA VALORES ISOLADOS PONDERADOS <u>NÃO</u> AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{\overline{X} = \frac{\Sigma Xifi}{\Sigma fi}}$$

04 – MÉDIA ARITMÉTICA PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS (Processo Longo)

$$\boxed{\overline{C} = \frac{\Sigma fiPM}{\Sigma fi}}$$

05 – MÉDIA ARITMÉTICA PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS (Processo Breve)

$$\boxed{\overline{C} = K_0 + h \times \left(\frac{\Sigma difi}{\Sigma fi}\right)} \quad \textbf{onde} \quad \boxed{di = \frac{PM - K_0}{h}}$$

06 – MEDIANA PARA DADOS <u>NÃO</u> AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{EM = \frac{N + 1}{2}}$$

07 – MEDIANA PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{\tilde{C} = Li + h \times \left(\frac{EM - \Sigma fac_{-}}{fi}\right)} \quad \textbf{onde} \quad \boxed{EM = \frac{\Sigma fi}{2}}$$

08 – MODA BRUTA

$$\boxed{\hat{C}_B = \text{PM da Classe de Maior Freqüência}}$$

09 – MODA DE CZUBER

$$\hat{C}_C = Li + h \cdot \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) \quad \text{OU} \quad \hat{C}_C = Li + h \cdot \left[\frac{f\,max - fant}{2 \cdot f\,max - (fant + fpost)} \right]$$

(fmáx - fant) + (fmáx - fpost)

ONDE: $\Delta_1 = f_{max} - f_{ant}$

$\Delta_2 = f_{max} - f_{post}$

10 – MODA DE KING

$$\hat{C}_K = Li + h \cdot \left(\frac{fpost}{fant + fpost} \right)$$

11 – MODA DE KARL PEARSON

$$\hat{C}_P = \left(3 \cdot \tilde{C} \right) - \left(2 \cdot \overline{C} \right)$$

12 – QUARTIL PARA DADOS NÃO AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$q_i = i \cdot \left(\frac{N+1}{4} \right)$$

13 – QUARTIL PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$Q_i = Li + h \cdot \left(\frac{q_i - \text{'fac -}}{fi} \right) \quad \text{onde} \quad q_i = i \cdot \left(\frac{Sfi}{4} \right)$$

14 – DECIL PARA DADOS NÃO AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$d_i = i \cdot \left(\frac{N+1}{10} \right)$$

15 – DECIL PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$D_i = Li + h \cdot \left(\frac{d_i - \text{'fac -}}{fi} \right) \quad \text{onde} \quad d_i = i \cdot \left(\frac{Sfi}{10} \right)$$

16 – CENTIL PARA DADOS NÃO AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$c_i = i \cdot \left(\frac{N+1}{100} \right)$$

17 – CENTIL PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$C_i = Li + h \cdot \left(\frac{c_i - \text{'fac -}}{fi} \right) \quad \text{onde} \quad c_i = i \cdot \left(\frac{Sfi}{100} \right)$$

18 – AMPLITUDE TOTAL OU INTERVALO TOTAL PARA DADOS NÃO AGRUPADOS DE FREQÜÊNCIAS

$$\boxed{AT = IT = Xi_{max} - Xi_{min}}$$

19 – AMPLITUDE TOTAL OU INTERVALO TOTAL PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{AT = Ls_{máx} - Li_{min}} \quad \text{ou} \quad \boxed{AT = PM_{\text{Última Classe}} - PM_{\text{Pr imeira Classe}}}$$

20 – DESVIO QUARTÍLICO PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

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

21 – DESVIO MÉDIO PARA DADOS NÃO AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{DM = \frac{S|di|}{n}} \quad \text{onde} \quad \boxed{di = |X_i - \overline{X}|}$$

22 – DESVIO MÉDIO PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{DM = \frac{S|di|fi}{Sfi}} \quad \text{onde} \quad \boxed{di = |PM - \overline{X}|}$$

23 – DESVIO PADRÃO PARA DADOS NÃO AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{S = \sqrt{\frac{S\,di^2}{n - 1}}} \quad \text{ou} \quad \boxed{S = \sqrt{\frac{S\,Xi^2 - \frac{(S\,Xi)^2}{n}}{n - 1}}} \quad \text{onde} \quad \boxed{di = X_i - \overline{X}}$$

24 – DESVIO PADRÃO PARA VALORES ISOLADOS PONDERADOS NÃO AGRUPADOS EM CLASSES DE FREQÜÊNCIAS

$$\boxed{S = \sqrt{\frac{S\,di^2 fi}{n - 1}}} \quad \text{ou} \quad \boxed{S = \sqrt{\frac{S\,Xi^2 fi - \frac{(S\,Xifi)^2}{n}}{n - 1}}} \quad \text{onde} \quad \boxed{di = X_i - \overline{X}}$$

25 – DESVIO PADRÃO PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS (Processo Longo)

$$\boxed{S = \sqrt{\frac{S\,di^2 fi}{n - 1}}} \quad \text{ou} \quad \boxed{S = \sqrt{\frac{SPM^2 fi - \frac{(S\,PMfi)^2}{n}}{n - 1}}} \quad \text{onde} \quad \boxed{di = PM - \overline{X}}$$

26 – DESVIO PADRÃO PARA DADOS AGRUPADOS EM CLASSES DE FREQÜÊNCIAS (Processo Breve)

$$\boxed{S = h \times \sqrt{\frac{Sdi^2 fi - \frac{(S\,difi)^2}{n}}{n - 1}}} \quad \text{onde} \quad \boxed{di = \frac{PM - K_0}{h}}$$

27 – COEFICIENTE DE VARIAÇÃO

$$\boxed{CV = \left(\frac{S}{\overline{C}}\right) \times 100}$$

28 – ERRO PADRÃO DA MÉDIA

$$\boxed{S_{\overline{X}} = \frac{S}{\sqrt{n}}}$$

29 – MOMENTO NATURAL (ABSOLUTO) DE ORDEM "r"

$$\boxed{m'_r = \frac{SPM^r fi}{Sfi}}$$

30 - MOMENTO DE ORDEM "r" EM RELAÇÃO A UMA ORIGEM QUALQUER "X$_o$"

$$\boxed{{}_{x_o}m_r = \frac{S(PM - X_o)^r fi}{Sfi}}$$

31 – MOMENTO CENTRADO NA MÉDIA DE ORDEM "r"

$$\boxed{m_r = \frac{S(PM - \overline{X})^r fi}{Sfi}}$$

32 – PRIMEIRO COEFICIENTE DE ASSIMETRIA DE KARL PEARSON

$$\boxed{AS = \frac{\overline{C} - \hat{C}_P}{S}}$$

33 – SEGUNDO COEFICIENTE DE ASSIMETRIA DE KARL PEARSON

$$\boxed{AS = \frac{3 \times (\overline{C} - \tilde{C})}{S}}$$

34 – COEFICIENTE MOMENTO DE ASSIMETRIA

$$\boxed{e_{M_1} = \frac{\sqrt{b_1}(b_2 + 3)}{2(5b_2 - 6b_1 - 9)}} \quad \textbf{ou} \quad \boxed{e_{M_2} = \sqrt{b_1} = \frac{m_3}{\left(\sqrt{m_2}\right)^3} = \frac{m_3}{S^3}}$$

Onde: $b_1 = \frac{m_3^2}{m_2^3}$ e $b_2 = \frac{m_4}{m_2^2}$ Onde: $m_2 = S^2$ e $\sqrt{m_2} = S$

35 – COEFICIENTE DE CURTOSE

$$\boxed{C = \frac{\frac{Q_3 - Q_1}{2}}{C_{90} - C_{10}}} \quad \textbf{ou} \quad \boxed{C = \frac{D_q}{C_{90} - C_{10}}} \quad \textbf{ou} \quad \boxed{C = \frac{Q_3 - Q_1}{2 \times (C_{90} - C_{10})}}$$

36 – COEFICIENTE MOMENTO DE CURTOSE

$$\boxed{b_2 = \frac{m_4}{m_2^2}} \quad \textbf{ou} \quad \boxed{b_2 = \frac{m_4}{S^4}} \quad \textbf{ou} \quad \boxed{c_2 = b_2 - 3}$$

37 – ESPAÇO AMOSTRAL

$$\boxed{S = n^k}$$

38 – FÓRMULA GERAL DE "PROBABILIDADES"

$$\boxed{Pr = P(A) = P(e) = P = \frac{h}{n}}$$

39 – PROBABILIDADE CONTRÁRIA OU <u>NÃO</u> EVENTO

$$\boxed{q = 1 - P}$$

40 – PERMUTAÇÃO

$$\boxed{p_n = n!}$$

41 – ARRANJO

$$\boxed{A_n^r = \frac{n!}{(n - r!)}}$$

42 – COMBINAÇÃO

$$\boxed{C_n^r = \frac{n!}{r!(n - r)!}}$$

43 – TEOREMA DE BAYES

$$\boxed{P(A/D) = \frac{P(A) \times P(D/A)}{P(A) \times P(D/A) + P(B) \times P(D/B) + P(C) \times P(D/C)}}$$

BIBLIOGRAFIA

BUSSAB, Wilton O.; MORETTIN, Pedro A.. **Estatística básica métodos quantitativos.** 4. ed. São Paulo: Saraiva, 1999.

CUNHA, Suzana Ezequiel; COUTINHO, Maria Tereza Cunha. **Iniciação à estatística.** 4. ed. Belo Horizonte: Lê, 1979.

CRESPO, Antônio Arnot. **Estatística fácil.** 17. ed. São Paulo: Saraiva, 1999.

DOWNING, Douglas; CLARK, Jeffrey. **Estatística aplicada.** São Paulo: Saraiva, 1998.

FONSECA, Jairo Simon da; MARTINS, Gilberto de Andrade. **Curso de estatística.** 6. ed. São Paulo: Atlas, 1996.

FONSECA, Jairo Simon da; MARTINS, Gilberto de Andrade; TOLEDO, Geraldo de Andrade. **Estatística aplicada.** 2. ed. São Paulo: Atlas, 1985.

HOFFMANN, Rodolfo. **Estatística para economistas.** 3. ed. São Paulo: Pioneira, 1998.

KASMIER, Leonardo J.. **Estatística aplicada a economia e administração.** São Paulo: McGraw-Hill, 1982.

KUME, Hitoshi. **Métodos estatísticos para melhoria da qualidade.** 2. ed. Rio de Janeiro: Gente, 1989.

LOPES, Paulo Afonso. **Probabilidades e estatística.** Rio de Janeiro: Reichmann & Affonso Editores, 1999.

MAIA, Paulo B. et al. **Controle estatístico do processo – CEP.** São Paulo: SHARP, 1992.

MARTINS, Gilberto de Andrade; DONAIRE, Denis. **Princípios de estatística.** 4. ed. São Paulo: Atlas, 1979.

MILONE, Giuseppe; ANGELINI, Flávio. **Estatística geral.** São Paulo: Atlas, 1993.

SILVA, Elio Medeiros da et al. **Estatística (Vol. 1, 2 e 3).** São Paulo: Atlas, 1995.

SPIEGEL, Murray Ralph. **Estatística.** 3. ed. São Paulo: Makron Books, 1993. (Coleção Schaum)

SPIEGEL, Murray Ralph. **Probabilidade e estatística.** São Paulo: McGraw-Hill, 1978. (Coleção Schaum)

SPINELLI, Walter; SOUZA, Maria Helena S. de. **Introdução à estatística.** 3. ed. São Paulo: Ática, 1997.

TOLEDO, Geraldo Luciano; OVALLE, Ivo Izidoro. **Estatística básica.** 2. ed. São Paulo: Atlas, 1995.

TRIOLA, Mário F. **Introdução à estatística.** Rio de Janeiro: LTC Livros Técnicos e Científicos Editora, 1999.

VIEIRA, Sonia. **Princípios de estatística.** São Paulo: Pioneira, 1999.

GRAMÁTICA GREGA

SÍMBOLOS UTILIZADOS NA ESTATÍSTICA

ALFABETO ([1]) - O Alfabeto grego consta de 24 (vinte e quatro) letras:

Maiús-culas	Minús-culas	Nome clássico	Designação em grego	Pronún-cia	Nome moderno
Α	α	alfa	ἄλφα	*a*	alfa
Β	β	beta	βῆτα	*b*	vita
Γ	γ	gama	γάμμα	*g*	gama
Δ	δ	delta	δέλτα	*d*	delta
Ε	ε	èpsilón	ἒ ψιλόν	*ĕ*	épsilon
Ζ	ζ	dzeta	ζῆτα	*dz*	zita
Η	η	eta	ἦτα	*ē*	ita
Θ	ϑ, θ	*theta* ([2])	θῆτα	*th*	*thita*
Ι	ι	iota	ἰῶτα	*i*	iota
Κ	κ	capa	κάππα	*k*	capa
Λ	λ	lambda	λάμβδα	*l*	landa
Μ	μ	*mü* ([2])	μῦ	*m*	mi
Ν	ν	*nü* ([2])	νῦ	*n*	ni
Ξ	ξ	xi (csi)	ξῖ	*x* (cs)	xi (csi)
Ο	ο	òmicrón	ὂ μικρόν	*ŏ*	ómicron
Π	π	pi	πῖ	*p*	pi
Ρ	ρ	ró	ῥῶ	*rh*	ró
Σ	σ, ς	sigma	σῖγμα	*s*	sigma
Τ	τ	tau	ταῦ	*t*	taf
Υ	υ	*üpsilón* ([2])	ὖ ψιλόν	*ü*	ípsilon
Φ	φ	fi	φῖ	*ph*	fi
Χ	χ	*khi* ([2])	χῖ	*kh*	*khi*
Ψ	ψ	psi	ψῖ	*ps*	psi
Ω	ω	omega	ὦ μέγα	*ō*	omega

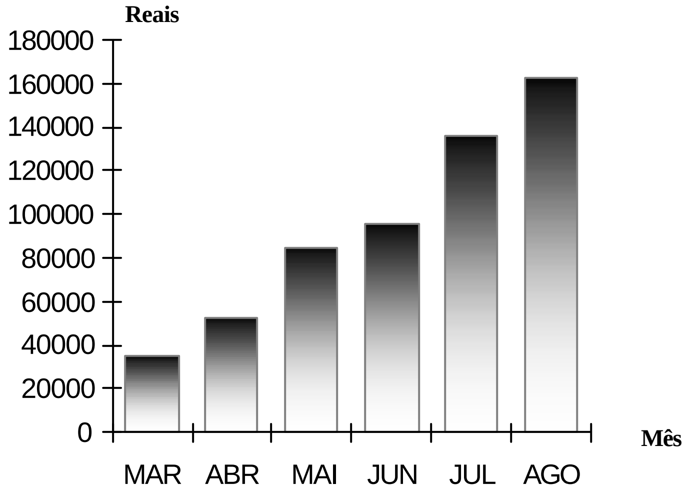
([1]) A palavra *Alfabeto* deriva das 2 (duas) primeiras letras gregas: *alfa* e *beta*.

Respostas dos exercícios sobre gráficos: (Pág. 42)

1 - Gráfico de Colunas

FGTS - ARRECADAÇÃO BRUTA - 200X

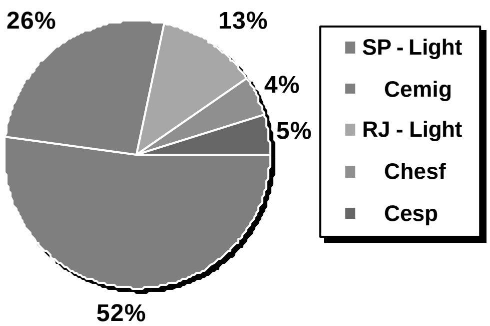
MÊS	R$ (MILHÕES)	ESCALA 1:20.000
MAR	34.888	1,7
ABR	52.334	2,6
MAI	85.023	4,3
JUN	95.254	4,8
JUL	136.126	6,8
AGO	162.643	8,1
å	**566.268**	-



2 - Gráfico em Setor - Circular

CONSUMO INDUSTRIAL DE ENERGIA ELÉTRICA
BRASIL - 200X

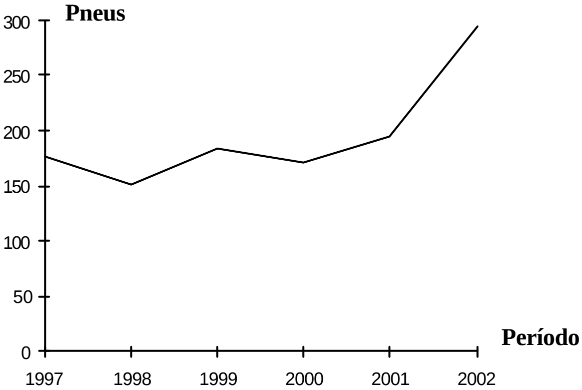
EMPRESAS	KWH(Milhões)	GRAUS	%
SP – Light	13.617	188	52
Cemig	6.763	94	26
RJ – Light	3.226	45	13
Chesf	1.183	16	4
Cesp	1.258	17	5
å	**26.047**	**360**	**100**



3 - Gráfico Linear

PRODUÇÃO DE PNEUMÁTICOS
SÃO PAULO - 1997 A 2002

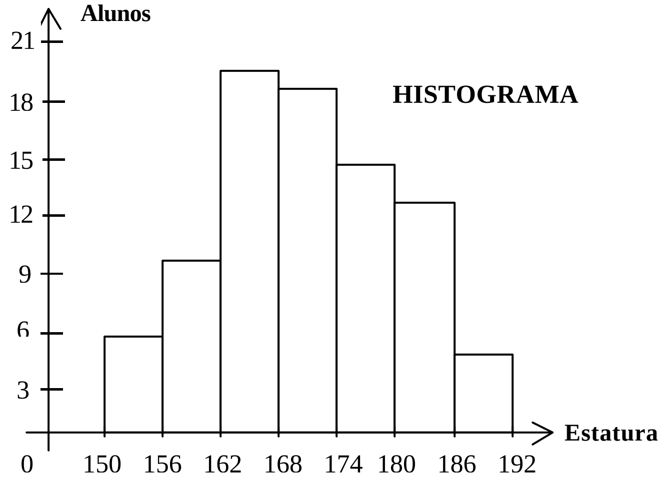
PERÍODO	PNEUS (1000)	ESCALA 1:50
1997	176	3,5
1998	152	3,0
1999	183	3,7
2000	171	3,4
2001	195	3,9
2002	294	5,9
å	**1.171**	-

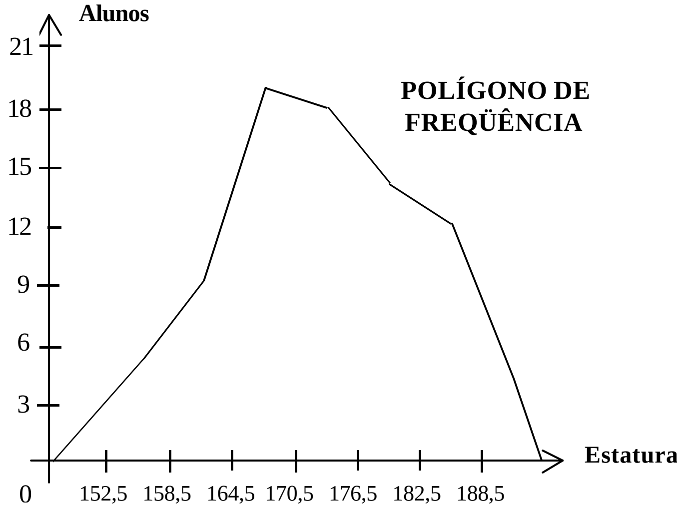


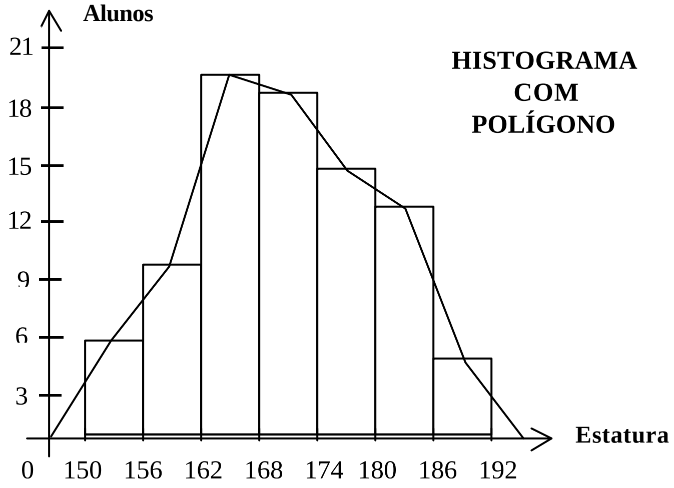
4 - Histograma e Polígono de Freqüência

ESTATURA DOS ALUNOS DO COLÉGIO DOS CAPUCHINHOS - J. FORA - 200X

ESTATURA (cm)	ALUNOS	PM	ESCALA 1:3
150 |—| 155	05	152,5	1,7
156 |—| 161	09	158,5	3,0
162 |—| 167	19	164,5	6,3
168 |—| 173	18	170,5	6,0
174 |—| 179	14	176,5	4,7
180 |—| 185	12	182,5	4,0
186 |—| 191	04	188,5	1,3
å	**81**	-	-



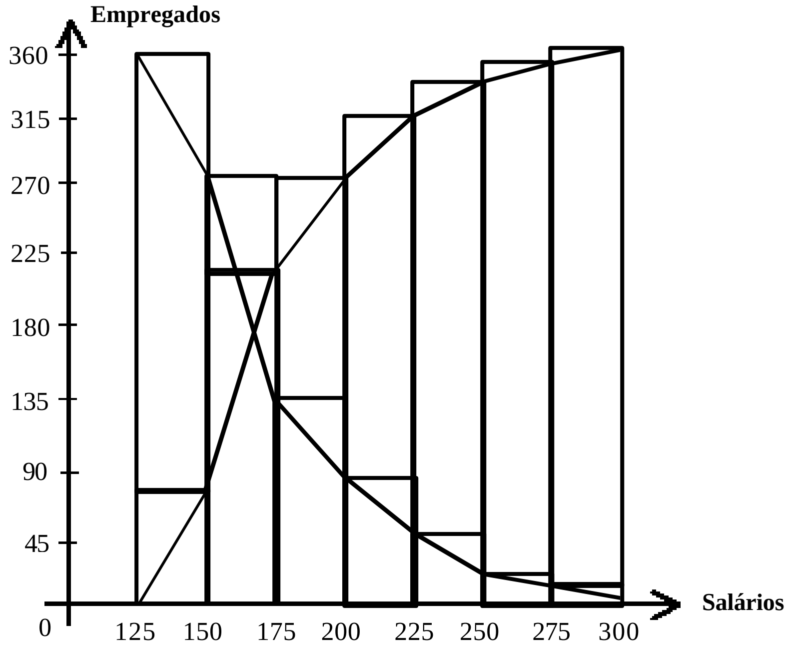




5 - Ogiva de Galton

NÚMERO DE EMPREGADOS POR CLASSE SALARIAL - ALFA BETA LTDA - J. FORA - 200X

SALÁRIOS (Em Reais)	NÚMERO DE FUNCIONÁRIOS	fac-	fac¯	ESCALA 1:45 (fac -)	ESCALA 1:45 (fac¯)
125 150	76	76	360	1,7	8,0
150 175	149	225	284	5,0	6,3
175 200	51	276	135	6,1	3,0
200 225	38	314	84	7,0	1,9
225 250	27	341	46	7,6	1,0
250 275	12	353	19	7,8	0,4
275 300	7	360	7	8,0	0,2
å	**360**	-	-	-	-



Resolução ***correta*** do exercício da página 77

1º Exercício: **Calcular o momento de assimetria (pelas duas fórmulas) da distribuição dos coeficientes de inteligência, da tabela abaixo:**

Quoc. Int.	fi	PM	PM.fi	$PM-\bar{X}$	$(PM-\bar{X}).fi$	$(PM-\bar{X})^2.fi$	$(PM-\bar{X})^3.fi$	$(PM-\bar{X})^4.fi$
68 A 72	4	70	280	-25,97	-103,87	2.697,07	-70.033,95	1.818.548,14
72 A 76	9	74	666	-21,97	-197,70	4.342,81	-95.397,06	2.095.555,41
76 A 80	16	78	1.248	-17,97	-287,47	5.164,82	-92.794,56	1.667.208,917
80 A 84	28	82	2.296	-13,97	-391,07	5.461,90	-76.284,51	1.065.440,26
84 A 88	45	86	3.870	-9,97	-448,50	4.470,05	-44.551,50	444.029,93
88 A 92	66	90	5.940	-5,97	-393,80	2.349,67	-14.019,72	83.650,98
92 A 96	85	94	7.990	-1,97	-167,17	328,76	-646,56	1.271,57
96 A 100	72	98	7.056	2,03	146,40	297,68	605,28	1.230,74
100 A 104	54	102	5.508	6,03	325,80	1.965,66	11.859,48	71.552,21
104 A 108	38	106	4.028	10,03	381,27	3.825,38	38.381,27	385.092,06
108 A 112	27	110	2.970	14,03	378,90	5.317,23	74.618,46	1.047.145,74
112 A 116	18	114	2.052	18,03	324,60	5.853,62	105.560,28	1.903.603,73
116 A 120	11	118	1.298	22,03	242,37	5.340,15	117.661,21	2.592.468,60
120 A 124	5	122	610	26,03	130,17	3.388,67	88.218,43	2.296.619,89
124 A 128	2	126	252	30,03	60,07	1.804,00	54.180,20	1.627.212,01
å	**480**		**46.064**	**30,50**	**0,00**	**52.607,47**	**97.356,76**	**17.100.630,18**

Respostas:

	Média	=	95,9667
	Momento de primeira ordem (m_1)	=	Zero
	Momento de segunda ordem (m_2)	=	109,5989
	Momento de terceira ordem (m_3)	=	202,8266
	Momento de quarta ordem (m_4)	=	35.626,3129
	Valor de b_1	=	0,0312487
	Valor de b_2	=	2,9659141
	Valor do C. Momento de Assimetria 1ª fórmula	=	0,0935952
	Valor do C. Momento de Assimetria 2ª fórmula	=	0,1767729
	Desvio Padrão	=	10,469

RESOLUÇÃO DOS EXERCÍCIOS SUGERIDOS DA PÁGINA 78

Tabela "A"

Receita	fi	fac	PM	PM.fi	$PM - \overline{X}$	$(PM - \overline{X})fi$	$(PM - \overline{X})^2 fi$	$(PM - \overline{X})^3 fi$	$(PM - \overline{X})^4 fi$	$(PM - 15)^2 fi$
150 a 200	2	2	175	350	-166	-332	55.112	-9.148.592	1.518.666.272	51.200
200 a 250	6	8	225	1.350	-116	-696	80.736	-9.365.376	1.086.383.616	264.600
250 a 300	7	15	275	1.925	-66	-462	30.492	-2.012.472	132.823.152	473.200
300 a 350	12	27	325	3.900	-16	-192	3.072	-49.152	786.432	1.153.200
350 a 400	10	37	375	3.750	34	340	11.560	393.040	13.363.360	1.296.000
400 a 450	8	45	425	3.400	84	672	56.448	4.741.632	398.297.088	1.344.800
450 a 500	5	50	475	2.375	134	670	89.780	12.030.520	1.612.089.680	1.058.000
Soma	**50**	-	-	**17.050**	**-112**	**0**	**327.200**	**-3.410.400**	**4.762.409.600**	**5.641.000**

Calcular: **Média, Mediana, Moda de Pearson, Desvio Padrão, CV, AS, m1, m2, m3, m4, b1, b2, em1 e em2.**

média	**mediana**		**m_1**	**m_2**	**m_3**	**m_4**
341,00	**341,67**		**0,00**	**6.544,00**	**-68.208,00**	**95.248.192,00**
moda	**s**	**S**	**b_1**	**b_2**	**em_1**	**em_2**
343,00	**80,89**	**81,72**	**0,01660124**	**2,224181168**	**0,1665053498**	**0,128845834**
AS 1º	**AS 2º**		**CV**	**${}_{15}m_2$**		
-0,025	**-0,025**		**23,96%**	**112.820,00**		

Tabela "B"

Peso (kg)	fi	fac	PM	PM.fi	$PM - \overline{X}$	$(PM - \overline{X})fi$	$(PM - \overline{X})^2 fi$	$(PM - \overline{X})^3 fi$	$(PM - \overline{X})^4 fi$	$(PM - 15)^2 fi$
2 a 4	2	2	3	6	-6	-12	72	-432	2592	288
4 a 6	6	8	5	30	-4	-24	96	-384	1536	600
6 a 8	14	22	7	98	-2	-28	56	-112	224	896
8 a 10	10	32	9	90	0	0	0	0	0	360
10 a 12	8	40	11	88	2	16	32	64	128	128
12 a 14	6	46	13	78	4	24	96	384	1536	24
14 a 16	4	50	15	60	6	24	144	864	5184	0
Soma	**50**	-	-	**450**	**0**	**0**	**496**	**384**	**11200**	**2296**

Calcular: **Média, Mediana, Moda de Pearson, Desvio Padrão, CV, AS, m1, m2, m3, m4, b1, b2, em1 e em2.**

Média	Mediana		**m_1**	**m_2**	**m_3**	**m_4**
9,000	**8,600**		**0,00**	**9,920**	**7,680**	**224,000**
Moda	**s**	**S**	**b_1**	**b_2**	**em_1**	**em_2**
7,800	**3,150**	**3,182**	**0,06042093**	**2,276274714**	**0,3212088475**	**0,245806697**
AS 1º	**AS 2º**		**CV**	**${}_{15}m_2$**		
0,381	**0,381**		**35,36%**	**45,920**		

Tabela "C"

Classes	fi	fac	PM	PM.fi	$PM - \overline{X}$	$(PM - \overline{X})fi$	$(PM - \overline{X})^2 fi$	$(PM - \overline{X})^3 fi$	$(PM - \overline{X})^4 fi$	$(PM - 15)^2 fi$
10 a 20	2	2	15	30	-33,20	-66,40	2.204,48	-73.188,74	2.429.866,04	0,00
20 a 30	6	8	25	150	-23,20	-139,20	3.229,44	-74.923,01	1.738.213,79	600,00
30 a 40	7	15	35	245	-13,20	-92,40	1.219,68	-16.099,78	212.517,04	2.800,00
40 a 50	12	27	45	540	-3,20	-38,40	122,88	-393,22	1.258,29	10.800,00
50 a 60	10	37	55	550	6,80	68,00	462,40	3.144,32	21.381,38	16.000,00
60 a 70	8	45	65	520	16,80	134,40	2.257,92	37.933,06	637.275,34	20.000,00
70 a 80	5	50	75	375	26,80	134,00	3.591,20	96.244,16	2.579.343,49	18.000,00
Soma	**50**	-	-	**2.410**	**-22,40**	**0,00**	**13.088,00**	**-27.283,20**	**7.619.855,36**	**68.200,00**

Calcular: **Média, Mediana, Moda de Pearson, Desvio Padrão, CV, AS, m1, m2, m3, m4, b1, b2, em1 e em2.**

média	**mediana**		**m_1**	**m_2**	**m_3**	**m_4**
48,200	**48,333**		**0,000**	**261,760**	**-545,664**	**152397,107**
moda	**s**	**S**	**b_1**	**b_2**	**em_1**	**em_2**
48,600	**16,179**	**16,343**	**0,01660125**	**2,224181168**	**0,1665053498**	**0,128845834**
AS 1º	**AS 2º**		**CV**	**${}_{15}m_2$**		
-0,025	**-0,025**		**33,91%**	**1.364**		