

Ministério da Educação  
Escola Técnica Aberta do Brasil  
Escola Técnica da Universidade Federal do Paraná

# Estatística Aplicada

Javert Guimarães Falco



Cuiabá, 2008

*Comissão Editorial* Profª Drª Maria Lucia Cavalli Neder - UFMT  
Profª Drª Ana Arlinda de Oliveira - UFMT  
Profª Drª Lucia Helena Vendrusculo Possari - UFMT  
Profª Drª Gleyva Maria Simões de Oliveira - UFMT  
Prof. M. Sc. Oreste Preti - UAB/UFMT  
Profª Esp. Mércia Freire Rocha Cordeiro Machado - ET-UFPR

*Designer Educacional* Maria Cristina de Aguiar Campos

#### Ficha Catalográfica

---

F181e FALCO, Javert Guimarães  
Estatística aplicada / Javert Guimarães Falco.  
Cuiabá: EdUFMT; Curitiba: UFPR, 2008.  
92p.: il.; color.

Inclui bibliografia  
ISBN 9788561819286

1. Estatística 2. Estatística - Textos didáticos 1. I. Título.

CDU - 311 (075)

---

*Revisão* Germano Aleixo Filho

*Capa* Téo de Miranda (*Lay out*)  
Hugo Leonardo Leão Oliveira (arte final)

*Ilustração* Téo de Miranda

*Diagramação* Terencio Francisco de Oliveira  
Elizabeth Kock Carvalho Netto

## PROGRAMA e-TEC BRASIL



Amigo(a) estudante:

O Ministério da Educação vem desenvolvendo Políticas e Programas para expansão da Educação Básica e do Ensino Superior no País. Um dos caminhos encontrados para que essa expansão se efetive com maior rapidez e eficiência é a modalidade a distância. No mundo inteiro são milhões os estudantes que freqüentam cursos a distância. Aqui no Brasil, são mais de 300 mil os matriculados em cursos regulares de Ensino Médio e Superior a distância, oferecidos por instituições públicas e privadas de ensino.

Em 2005, o MEC implantou o **Sistema Universidade Aberta do Brasil** (UAB), hoje, consolidado como o maior programa nacional de formação de professores, em nível superior.

Para expansão e melhoria da educação profissional e fortalecimento do Ensino Médio, o MEC está implementando o **Programa Escola Técnica Aberta do Brasil** (e-Tec Brasil). Espera, assim, oferecer aos jovens das periferias dos grandes centros urbanos e dos municípios do interior do País oportunidades para maior escolaridade, melhores condições de inserção no mundo do trabalho e, dessa forma, com elevado potencial para o desenvolvimento produtivo regional.

O e-Tec é resultado de uma parceria entre a Secretaria de Educação Profissional e Tecnológica (SETEC), a Secretaria de Educação a Distância (SEED) do Ministério da Educação, as universidades e escolas técnicas estaduais e federais.

O Programa apoia a oferta de cursos técnicos de nível médio por parte das escolas públicas de educação profissional federais, estaduais, municipais e, por outro lado, a adequação da infra-estrutura de escolas públicas estaduais e municipais.

Do primeiro Edital do e-Tec Brasil participaram 430 proponentes de adequação de escolas e 74 instituições de ensino técnico, as quais propuseram 147 cursos técnicos de nível médio, abrangendo 14 áreas profissionais. O resultado deste Edital contemplou 193 escolas em 20 unidades federativas. A perspectiva do Programa é que sejam ofertadas 10.000 vagas, em 250 pólos, até 2010.

Assim, a modalidade de Educação a Distância oferece nova interface para a mais expressiva expansão da rede federal de educação tecnológica dos últimos anos: a construção dos novos centros federais (CEFETs), a organização dos Institutos Federais de Educação Tecnológica (IFETs) e de seus *câmpus*.

O Programa e-Tec Brasil vai sendo desenhado na construção coletiva e participação ativa nas ações de democratização e expansão da educação profissional no País, valendo-se dos pilares da educação a distância, sustentados pela formação continuada de professores e pela utilização dos recursos tecnológicos disponíveis.

A equipe que coordena o Programa e-Tec Brasil lhe deseja sucesso na sua formação profissional e na sua caminhada no curso a distância em que está matriculado(a).

Brasília, Ministério da Educação – setembro de 2008.

# CURSO TÉCNICO EM GESTÃO PÚBLICA

Seja bem-vindo(a):

Queremos cumprimentá-lo(a), pois você está começando o Curso Técnico em Gestão Pública do Programa e-Tec Brasil 2008.

O profissional Técnico em Gestão Pública deverá, após o término do Curso, estar integrado às particularidades da gestão pública, tais como:

- a estrutura do Poderes Públicos e hierarquia dos Setores Públicos;
- o processo de elaboração e execução dos planejamentos políticos e econômicos da administração pública, bem como o processo de prestação de contas aos cidadãos;
- o funcionamento dos quadros dos servidores públicos, bem como o processo de contratação e remuneração;
- os processos de comunicações formais utilizados pela administração pública, bem como o processo de aplicação destes recursos;
- o atendimento ao público e o processo de aplicabilidade da legislação, bem como adaptar-se às atualizações devido às mudanças de legislação.

O Programa e-Tec Brasil se constitui instrumento valioso para a formação de seus cidadãos. Esperamos, que, efetivamente, você aproveite esta oportunidade de escrever sua história e seu futuro ao lado de quem está ajudando a transformar nosso País.

Sucesso para você!

Márcia Denise Gomes Machado Carlini,  
Coordenadora do Curso Técnico em Gestão Pública - e-Tec Brasil.



## SUMÁRIO

Estatística Aplicada - Javert Guimarães Falco

---

CONVERSA INICIAL .....	9
UNIDADE 1 - INTRODUÇÃO À ESTATÍSTICA .....	11
UNIDADE 2 - TABELAS E GRÁFICOS .....	21
UNIDADE 3 - DISTRIBUIÇÃO DE FREQUÊNCIAS .....	37
UNIDADE 4 - MEDIDAS .....	55
UNIDADE 5 - PROBABILIDADE .....	81
RETOMANDO A CONVERSA INICIAL .....	89
REFERÊNCIAS BIBLIOGRÁFICAS .....	91





## CONVERSA INICIAL

Caro(a) aluno(a):

Este material foi elaborado a partir de minha experiência na educação presencial e a distância, buscando trazer de forma objetiva, simples e prática os principais conteúdos que serão importantes para seu exercício profissional.

Objetiva-se, nesta disciplina, fundamentar as possibilidades de coletar, organizar, descrever e interpretar dados, enfim, medir fenômenos coletivos.

Os assuntos abordados iniciam-se com um texto introdutório, onde são estudados: panorama histórico, conceitos e variáveis .

Na Unidade II, estudaremos tabelas e gráficos. Na Unidade III, a distribuição de frequência é o nosso foco. Na Unidade IV, as medidas são estudadas em profundidade. Na Unidade V, estudaremos Probabilidade.

Espero que, através dos conteúdos contemplados nesta apostila, somados às aulas expositivas e ao seu esforço pessoal, possamos quebrar, juntos, o paradigma de que a Estatística é uma disciplina "difícil, complicada" e transformá-la em uma disciplina de simples compreensão, útil e aplicável no seu cotidiano profissional.

Muito estudo e consequente sucesso nesta caminhada!

Prof. Javert Guimarães Falco



# UNIDADE

INTRODUÇÃO  
À ESTATÍSTICA

1



Nesta unidade, apresentaremos noções básicas de Estatística e vocabulário correspondente.

O homem é curioso, e como tal, procura investigar sobre tudo aquilo que o cerca. Investigação sugere pesquisa, busca de informações e análise de dados e tudo isto faz pensar em ESTATÍSTICA.

O relacionamento da Estatística com as demais ciências é cada vez mais intenso. Veja, por exemplo, que a estatística auxilia a Genética, nas questões de hereditariedade; é valiosa na Economia, na análise da produtividade, da rentabilidade, nos estudos de viabilidade; é básica para as Ciências Sociais, nas pesquisas sócio-econômicas; é de aplicação intensa na Engenharia Industrial, no controle de qualidade, na comparação de fabricações; é indispensável à Administração, à Programação, à Medicina, à Psicologia, à História, e, de forma direta ou indireta, às demais atividades.

## PANORAMA HISTÓRICO DA ESTATÍSTICA

Historicamente, o desenvolvimento da estatística pode ser entendido a partir de dois fenômenos : a necessidade de governos coletarem dados censitários e o desenvolvimento da teoria do cálculo das probabilidades.

Dados têm sido coletados através de toda a história. Na Antiguidade, vários povos já registravam o número de habitantes, de nascimentos, de óbitos, faziam estimativas das riquezas sociais, distribuíam equitativamente terras aos povos, cobravam impostos e realizavam inquéritos quantitativos por processos que, hoje, chamaríamos de "estatísticas". Na Idade Média colhiam-se informações, geralmente com finalidades tributárias ou bélicas. Atualmente, informações numéricas são necessárias para cidadãos e organizações de qualquer natureza, e de qualquer parte do mundo globalizado.

### O que é Estatística?

"Estatística é um conjunto de métodos e processos quantitativos que serve para estudar e medir os fenômenos coletivos."

Dugé de Bernonville

Em outras palavras, é a ciência que se preocupa com a coleta, a organização, descrição (apresentação), análise e interpretação de dados experimentais e tem como objetivo fundamental o estudo de uma população.

Este estudo pode ser feito de duas maneiras:

Investigando todos os elementos da população;

Amostragem, ou seja, selecionando alguns elementos da população.

### Conceitos Estatísticos

População: Conjunto de indivíduos, objetos ou informações que apresentam pelo

menos uma característica comum, cujo comportamento interessa-nos analisar. Em outras palavras, conjunto de todas as medidas, observações relativas ao estudo de determinado fenômeno.

Como em qualquer estudo estatístico temos em mente estudar uma ou mais características dos elementos de uma população, é importante definir bem essas características de interesse para que sejam delimitados os elementos que pertencem à população e quais os que não pertencem.

Exemplos:

Deseja-se saber se nas indústrias situadas no Estado do Paraná, em 2007, existia algum tipo de controle ambiental.

1- População ou universo: indústrias situadas no Estado do Paraná em 2007.  
Característica: existência ou não de algum tipo de controle ambiental na indústria.

Deseja-se conhecer o consumo total de energia elétrica em MWH nas residências da cidade de Curitiba no ano de 2007.

2- População ou universo: todas as residências que estavam ligadas à rede elétrica em Curitiba, em 2007.

Características: consumo anual de energia elétrica em MWH.

Divisão da população

1-População Finita: apresenta um número limitado de elementos. É possível enumerar todos os elementos componentes.

Exemplo:

Idade dos alunos do curso de Gestão Pública em EAD no Estado do Paraná. População: Todos os alunos de Gestão Pública em EAD No Estado do Paraná. 23-2-1-2-População infinita: apresenta um número ilimitado de elementos. Não é possível enumerar todos os elementos componentes. Entretanto, tal definição existe apenas no campo teórico, uma vez que, na prática, nunca encontraremos populações com infinitos elementos, mas sim, populações com grande número de componentes; e nessas circunstâncias, tais populações são tratadas como se fossem infinitas.

Exemplo:

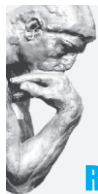
Tipos de bactérias no corpo humano

População: Todas as bactérias existentes no corpo humano.

Em geral, como os universos são grandes, para se investigarem todos os elementos populacionais, para determinarmos a característica, necessita-se muito tempo, e/ou o custo é elevado, e/ ou o processo de investigação leva à destruição do elemento observado, ou, como no caso de populações infinitas, é impossível observar a totalidade da população. Assim, estudar parte da população constitui-se um aspecto fundamental da Estatística.

**Amostragem:** É a coleta das informações de parte da população, chamada amostra, mediante métodos adequados de seleção destas unidades.

**Amostra:** É uma parte (um subconjunto finito) representativa de uma população selecionada segundo métodos adequados. O objetivo é tirar conclusões sobre populações com base nos resultados da amostra. Para isso, é necessário garantir que amostra seja representativa, ou seja, a amostra deve conter as mesmas características básicas da população, no que diz respeito ao fenômeno que desejamos pesquisar.



**PARA  
REFLETIR**

Você já deve ter sido entrevistada ou conhece alguém que já respondeu ao Censo do IBGE. Você considera que o Censo é importante para – além de se conhecerem números relativos à idade, profissão, nascimento, morte etc – serem feitas propostas de melhoria das condições de vida?

**Censo:** É o exame completo de toda população. Quanto maior a amostra, mais precisas e confiáveis deverão ser as induções feitas sobre a população. Logo, os resultados mais perfeitos são obtidos pelo Censo. Na prática, esta conclusão muitas vezes não acontece, pois, o emprego de amostras, com certo rigor técnico, pode levar a resultados mais confiáveis ou até mesmo melhores do que os que seriam obtidos através de um Censo.

As razões de se recorrer a amostras são: menor custo e tempo para levantarem dados; melhor investigação dos elementos observados.

Dados sobre o Brasil podem ser obtidos junto ao Instituto Brasileiro de Geografia e Estatística (IBGE) pelo site [www.ibge.gov.br](http://www.ibge.gov.br)

A Estatística ocupa-se fundamentalmente das propriedades das populações cujas características são passíveis de representação numérica como resultado de medições e contagens. Essas características da população são comumente chamadas de VARIÁVEIS. As características ou variáveis podem ser divididas em dois tipos: qualitativas e quantitativas.

**variáveis qualitativas** - quando o resultado da observação é apresentado na forma de qualidade ou atributo. Dividem-se em:

- **variáveis nominais:** quando podem ser separadas por categorias chamadas de não mensuráveis.  
Exemplo: a cor dos olhos, tipo de acomodação, marcas de carro, sexo etc.
- **variáveis ordinais:** quando os números podem agir como categorias ou ordenações. Como sugere o nome, elas envolvem variáveis que representam algum elemento de ordem. Uma classificação em anos pode ser um exemplo clássico. A classificação deste tipo de variáveis geralmente causa confusão.  
Exemplo: Grau de satisfação da população brasileira com relação ao trabalho de seu presidente (valores de 0 a 5, com 0 indicando totalmente insatisfeito e 5 totalmente satisfeito).

**variáveis quantitativas** - quando o resultado da observação é um número, decorrente de um processo de mensuração ou contagem. Dividem-se em:

● **variáveis contínuas:** são aquelas que podem assumir qualquer valor num certo intervalo (contínuo) da reta real. Não é possível enumerar todos os possíveis valores. Essas variáveis, geralmente, provêm de medições.

Exemplo: A altura dos alunos é uma variável contínua, pois teoricamente, um aluno poderá possuir altura igual a 1,80m, 1,81m, 1,811m, 1,812m ... (medições: peso, estatura, etc.)

● **variáveis discretas:** são aquelas que podem assumir apenas valores inteiros em pontos da reta real. É possível enumerar todos os possíveis valores da variável.

Exemplo: Número de alunos de uma escola, número de mensagens em uma secretária eletrônica etc.

As variáveis podem ser resumidas da seguinte maneira:

<b>Variáveis</b>	Qualitativa	Nominal (sexo, cor dos olhos etc) Ordinal (classe social, grau de instrução etc)
	Quantitativa	Contínua (peso, altura...) Discreta (número de filhos, número de carros etc)

## Divisão da Estatística

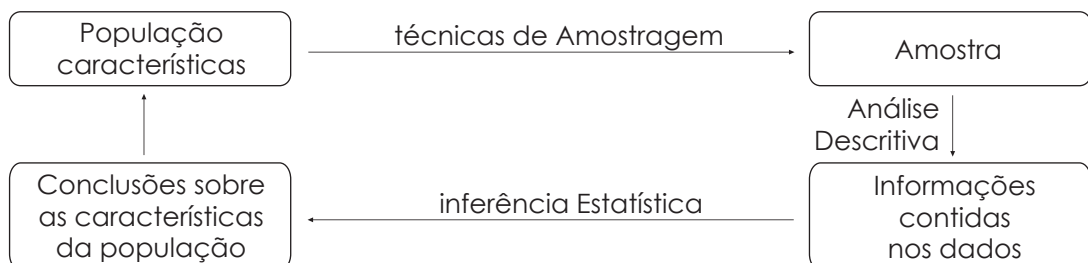
A Teoria Estatística moderna se divide em dois grandes campos:

● **Estatística Descritiva** - é aquela que se preocupa com a coleta, organização, classificação, apresentação, interpretação e análise de dados referentes ao fenômeno através de gráficos e tabelas, além de calcular medidas que permitam descrever o fenômeno.

● **Estatística indutiva** - é aquela que partindo de uma amostra, estabelece hipóteses, tira conclusões sobre a população de origem e que formula previsões, fundamentando-se na teoria das probabilidades. A estatística indutiva cuida da análise e interpretação dos dados. O processo de generalização do método indutivo está associado a uma margem de incerteza. Isto se deve 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.

Para se analisarem os dados de forma estatística podem-se obter os resultados de duas maneiras: através de um censo ou através de uma amostragem (pesquisa em uma amostra).

Para exemplificar essas teorias analise o esquema abaixo:





Exemplos de utilização:

Pesquisa de Mercado, Pesquisa de opinião pública e em praticamente todo experimento.

## **Fases do método Estatístico**

1ª FASE - Definição do Problema: Saber exatamente aquilo que se pretende pesquisar é o mesmo que definir corretamente o problema.

2ª FASE - Planejamento: Como levantar informações? Que dados deverão ser obtidos? Qual levantamento a ser utilizado? Censitário? Por amostragem? E o cronograma de atividades? Os custos envolvidos? etc.

3ª FASE – Coletas de Dados: Fase operacional. É o registro sistemático de dados, com um objetivo determinado.

Coleta de Dados. Após a definição do problema a ser estudado e o estabelecimento do planejamento da pesquisa (forma pela quais os dados serão coletados; cronograma das atividades; custos envolvidos exame das informações disponíveis; delineamento da amostra etc.), o passo seguinte é a coleta de dados, que podem ser de dois tipos:

1- Dados Primários: os dados são obtidos diretamente na fonte originária (coleta direta)  
Exemplo: Preferência dos consumidores por um determinado produto.

Métodos de coleta de dados primários: É importante garantir que a coleta de dados primários seja executada de maneira estatisticamente correta, senão os resultados podem ser tendenciosos.

Observação: O pesquisador não pergunta, observa. Por exemplo: pesquisa de observação para diagnosticar as necessidades de trânsito de uma cidade.

2- Levantamento: É o método mais comum de se coletar dados. O instrumento pode ser um questionário estruturado ou um roteiro de itens em que o entrevistado disserta à vontade sobre cada item da pesquisa.

As três principais formas de levantamento, resumindo as vantagens e desvantagens são:

- Entrevista pessoal: mais flexível e muito caro.
- Telefone: mais barato, penetra em segmentos difíceis, mas é de fácil recusa.
- Questionários (postal, fax ou e-mail): mais lento, média de retorno das respostas muito baixas, mas sem interferência do pesquisador.

Dados Secundários: os dados são obtidos de algo já disposto. Provém da coleta direta.

Exemplo:

Pesquisa sobre a mortalidade infantil, que é feita através de dados colhidos por outras pesquisas.



**PARA  
REFLETIR**

Você já deve ter sido entrevistado(a) para alguma pesquisa. Você sempre respondeu verdadeiramente? Quando a pesquisa é feita por telefone ou e-mail, você tem a mesma disposição para responder?

É mais seguro trabalhar com fontes primárias. O uso da fonte secundária traz o grande risco de erros de transcrição.

4º FASE – Apuração dos Dados: Resumo dos dados através de sua contagem e agrupamento. É a condensação e tabulação de dados.

5º FASE – Apresentação dos Dados: Há duas formas de apresentação, que não se excluem mutuamente. A apresentação tabular, ou seja, é uma apresentação numérica dos dados em linhas e colunas distribuídas de modo ordenado, segundo regras práticas fixadas pelo Conselho Nacional de Estatística e a apresentação gráfica dos dados numéricos que constitui uma apresentação geométrica permitindo uma visão rápida e clara do fenômeno.

6º FASE – Análise e Interpretação dos Dados: A última fase do trabalho estatístico é a mais importante e delicada. Está ligada essencialmente ao cálculo de medidas e coeficientes, cuja finalidade principal é descrever o fenômeno (estatística descritiva). Na estatística indutiva a interpretação dos dados se fundamenta na teoria da probabilidade.

## Em síntese

- **Estatística** é a ciência que se preocupa com a coleta, organização, descrição, análise e interpretação de dados experimentais.
- **População** é um conjunto de indivíduos, objetos ou informações que apresentam pelo menos uma característica comum.
- **Amostra** é uma parte representativa de uma população.

As características das populações são chamadas de variáveis, que podem ser divididas em:

- **Qualitativas:**
  - nominal (sexo, cor dos olhos...)
  - ordinal (classe social, grau de instrução...)
- **Quantitativas:**
  - contínua (peso, altura...)
  - discreta (número de filhos, número de carros...)

**Estatística Descritiva** se preocupa com a coleta, organização, classificação, apresentação, interpretação e análise de dados experimentais.

**Estatística Indutiva** se preocupa com as hipóteses e conclusões sobre a população.

**Responda às questões abaixo e confira as respostas, no gabarito, no final do livro**

**1) Podemos dizer que a Estatística pode ser:**

- a) Construtiva ou referencial
- b) Destrutiva ou direcional
- c) Descritiva ou inferencial
- d) Regressiva ou pontencial
- e) Ingressiva ou preferencial

**2) O que é Estatística Descritiva?**

- a) É o cálculo de medidas que permitirão descrever, com detalhes, o fenômeno que está sendo analisado.
- b) É a parte da Estatística referente à coleta e à tabulação dos dados.
- c) É a parte da Estatística referente às conclusões sobre as fontes de dados.
- d) É a generalização das conclusões sobre as fontes de dados.
- e) É a obtenção dos dados seja através de simples observação ou mediante a utilização de alguma ferramenta.

**3) Assinale a alternativa que expressa corretamente o que é população para a Estatística:**

- a) população é a totalidade de habitantes de uma país.
- b) população é uma parte da amostra.
- c) população é uma reunião de pessoas populares.
- d) população é um conjunto de elementos portadores de, pelo menos, uma característica comum.
- e) população é o conjunto de elementos que responderão a um questionário que fornecerá informações relevantes para uma pesquisa.

**4) A definição correta de amostra é a que consta da alternativa:**

- a) amostra é o conjunto de dados obtidos em uma pesquisa.
- b) amostra é um brinde que ganhamos em um evento.
- c) amostra é o conjunto dos resultados obtidos com uma pesquisa.
- d) amostra é uma tabela de dados obtida com uma pesquisa.
- e) amostra é um subconjunto finito de uma população.

**5) Dentro do universo estatístico "alunos de uma faculdade", as variáveis "cor dos cabelos" e número de irmão" podem ser classificadas, respectivamente, como:**

- a)** 1ª variável quantitativa discreta; 2ª. variável qualitativa.
- b)** 1ª. variável quantitativa contínua; 2ª variável qualitativa.
- c)** 1ª. variável qualitativa; 2ª. variável quantitativa contínua.
- d)** 1ª. variável qualitativa; 2ª variável quantitativa discreta.
- e)** 1ª. variável quantitativa discreta; 2ª. variável qualitativa contínua.

# UNIDADE

TABELAS E  
GRÁFICOS

# 2



Nesta unidade, estudaremos a apresentação tabular que é uma apresentação numérica dos dados. Consiste em dispor os dados em linhas e colunas distribuídos de modo ordenado, segundo algumas regras práticas ditadas pelo Conselho Nacional de Estatística e pelo IBGE.

As tabelas têm a vantagem de conseguir expor, sinteticamente e em um só local, os resultados sobre determinado assunto, de modo a se obter uma visão global mais rápida daquilo que se pretende analisar.

Essa integração de valores que temos nas tabelas, nos permite ainda a utilização de representações gráficas, normalmente, uma forma mais útil e elegante de demonstrar as características que serão analisadas.

## TABELA

É um quadro que resume um conjunto de observações. Uma tabela compõe-se de:

**Corpo** – conjunto de linhas e colunas que contêm informações sobre a variável em estudo;

**Cabeçalho** – parte superior da tabela que especifica o conteúdo das colunas;

**Coluna indicadora** – parte da tabela que especifica o conteúdo das linhas;

**Casa ou célula** – espaço destinado a um só número;

**Título** – conjunto de informações, as mais completas possíveis, respondendo às perguntas: O quê?, Quando?, Onde?, localizado no topo da tabela;

**Fonte** – indicação da entidade responsável pelo fornecimento dos dados ou pela sua

Veja o exemplo abaixo, que mostra os elementos que formam uma tabela.

Tabela 11 – Evolução da renda dos Empreendedores Brasileiros				
Faixa de Renda	2000(%)	2000(%)	2000(%)	2000(%)
Menos de 3 SM	30	39	43	53
De 3 a 6 SM	30	31	34	22
Mais de 6 a 9 SM	14	12	11	8
Mais de 9 a 15 SM	12	12	9	6
Mais de 15 SM	1	3	3	1
Não sabe Recusou	4	6	2	2
Total	100	100	100	100
Fonte: Pesquisas GEM 2000 – 2003				



De acordo com a resolução 886 da Fundação IBGE nas casas ou células, devemos colocar:

- 1) um traço horizontal ( \_ ) quando o valor é zero e não só quanto a natureza das coisas , como que resultado do inquérito;
- 2) três pontos (...) quando não temos os dados.
- 3) um ponto de interrogação ( ? ) quando temos dúvida quanto a exatidão de determinado valor
- 4) zero (0) quando o valor é muito pequeno para ser expresso pela unidade utilizada. Se os valores são expressos em numerais decimais precisamos acrescentar a parte decimal um número correspondente de zeros (0,0; 0,00; 0,00; ..).

## Séries Estatísticas

Denomina-se série estatística toda tabela que apresenta a distribuição de um conjunto de dados estatísticos em função da ÉPOCA, do LOCAL, ou da ESPÉCIE (fenômeno).

Numa série estatística observa-se a existência de três elementos ou fatores: o TEMPO, o ESPAÇO e a ESPÉCIE. Conforme varie um desses elementos, a série estatística classifica-se em TEMPORAL, GEOGRÁFICA e ESPECÍFICA

### - Série temporal, Histórica ou cronológica

É a série cujos dados estão em correspondência com o tempo, ou seja, variam com o tempo.

Exemplo:

Preço do artigo "Y" no  
atacado na cidade  
"X" anos

Preço médio em reais

Anos	Preços Médios em Reais
2003	2,43
2004	2,54
2005	3,01
2006	2,99
2007	2,83

Fonte: Dados Fictícios

### - Série Geográfica, Territorial ou de Localidade

É a série cujos dados estão em correspondência com a região geográfica, ou seja, o elemento variável é o fator geográfico (a região).

Exemplo:

Número de Assaltos na  
Cidade "X" em 2006

Região	Número de Assaltos
Centro	74
Zona Sul	54
Zona Norte	31
Zona Leste	29
Zona Oeste	44

Fonte: Dados Fictícios



## - Série Específica ou categórica

É a série cujos dados estão em correspondência com a espécie, ou seja, variam com o fenômeno.

Exemplo:

Número de Candidatos  
ao vestibular da  
universidade "X"  
em 2006

Área ofertada	Número de Candidatos
Ciências Sociais Aplicadas	2086
Ciências Exatas	1065
Ciências Humanas	1874
Ciências Biológicas	1102
Ciências Tecnológicas	1902

Fonte: Dados Fictícios

## - Séries mistas

As combinações entre as séries anteriores constituem novas séries que são denominadas séries compostas ou mistas e são apresentadas em tabelas de dupla entrada.

Exemplo:

Número de Alunos  
Matriculados nas  
Escolas Particulares na  
cidade "X"

Bairros	2005	2006	2007
Bairro - A	2894	3454	2989
Bairro - B	7075	9876	6543
Bairro - C	1099	3218	2100
Bairro - D	4333	3455	3543
Bairro - E	2976	1765	4098

Fonte: Dados Fictícios

## GRÁFICOS ESTATÍSTICOS

A apresentação gráfica é um complemento importante da apresentação tabular. A vantagem de um gráfico sobre a tabela está em possibilitar uma rápida impressão visual da distribuição dos valores ou das freqüências observadas. Os gráficos propiciam uma idéia inicial mais satisfatória da concentração e dispersão dos valores, uma vez que, através deles, os dados estatísticos se apresentam em termos de grandezas visualmente interpretáveis.

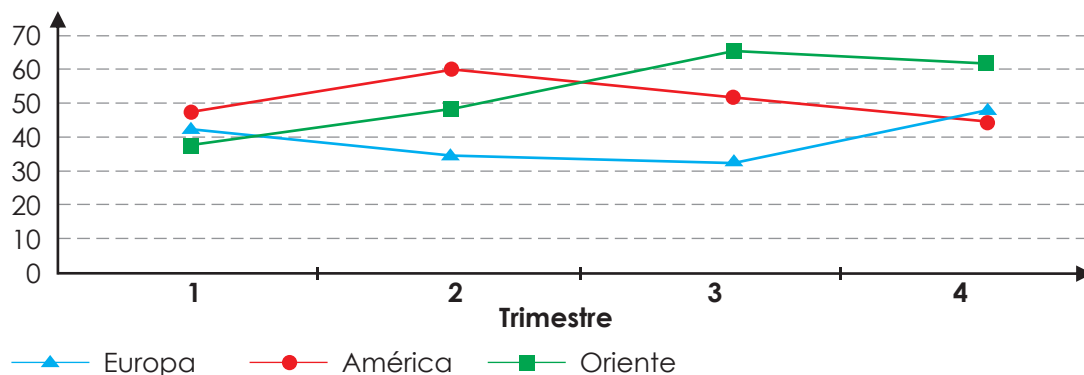
A representação gráfica de um fenômeno deve obedecer a certos requisitos fundamentais, para ser realmente útil:

- **simplicidade:** destituído de detalhes e traços desnecessários;
- **clareza:** possuir uma correta interpretação dos valores representativos do fenômeno em estudo;
- **veracidade:** expressar a verdade sobre o fenômeno em estudo.

## Principais tipos de Gráficos

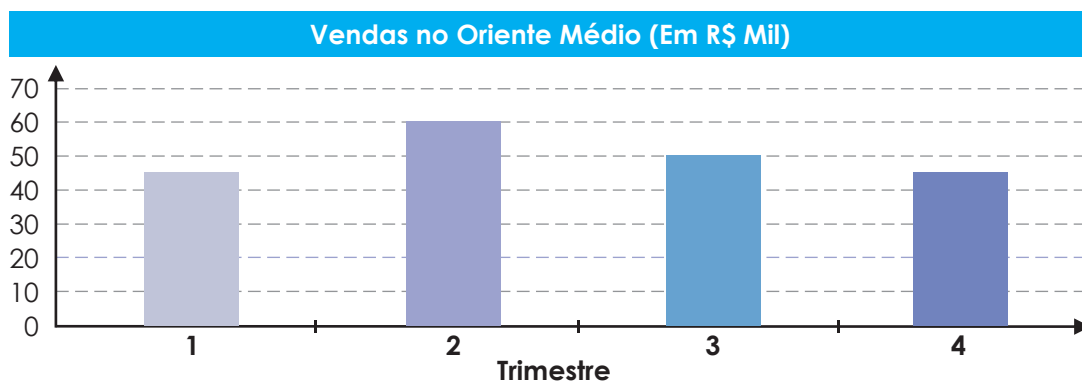
### Gráfico em curvas ou em linhas

São usados para representar séries temporais, principalmente quando a série cobrir um grande número de períodos de tempo.



### Gráfico em Colunas trimestre

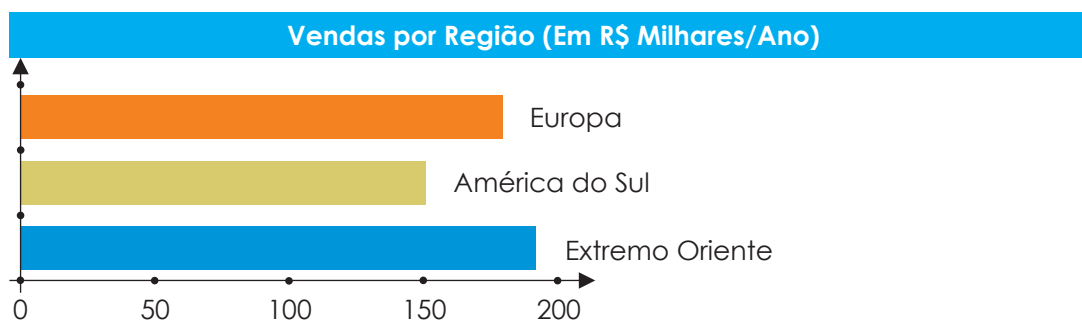
É a representação de uma série estatística através de retângulos, dispostos em colunas (na vertical) ou em retângulos (na horizontal). Este tipo de gráfico representa praticamente qualquer série estatística.



Observação: As regras para a construção são as mesmas do gráfico em curvas. As bases das colunas são iguais e as alturas são proporcionais aos respectivos dados. O espaço entre as colunas pode variar de 1/3 a 2/3 do tamanho da base da coluna.

### Gráfico em Barras

É representado por retângulos dispostos horizontalmente, prevalecendo os mesmos critérios adotados na elaboração de gráfico em coluna.



## Gráfico em Setores

É a representação gráfica de uma série estatística em um círculo de raio qualquer, por meio de setores com ângulos centrais proporcionais às ocorrências. É utilizado quando se pretende comparar cada valor da série com o total.

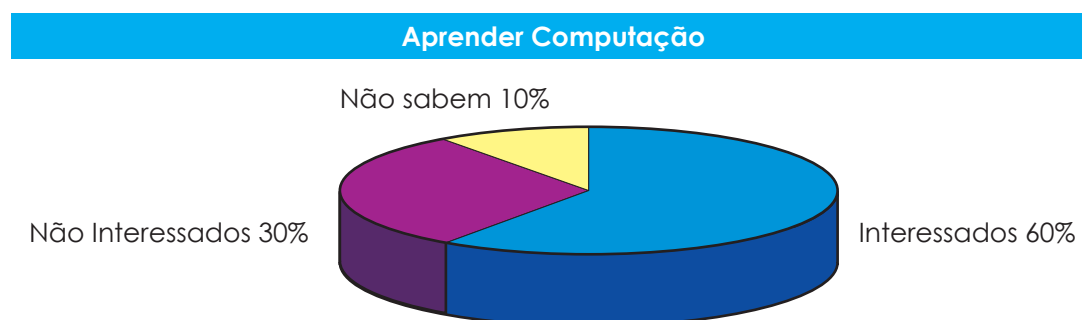
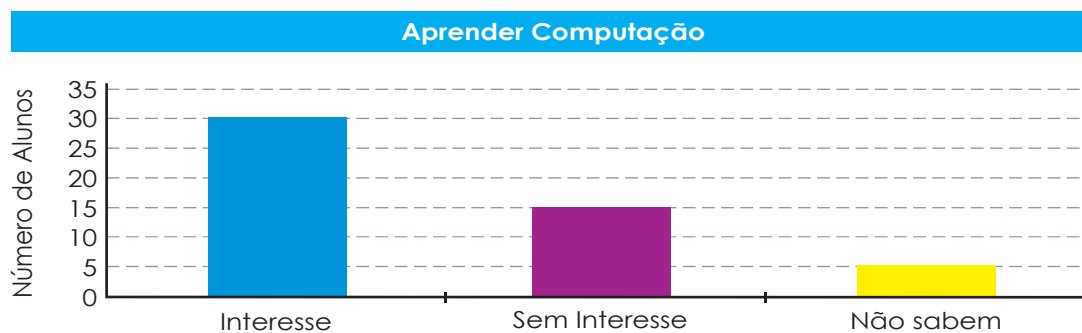
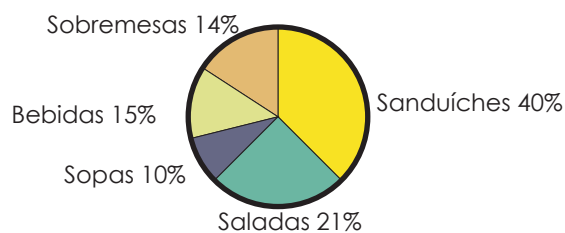
O total da série corresponde a  $360^\circ$  (total de graus de um arco de circunferência). O gráfico em setores representa valores absolutos ou porcentagens complementares.

As séries geográficas, específicas e as categorias em nível nominal são mais representadas em gráficos de setores, desde que não apresentem muitas parcelas (no máximo sete).

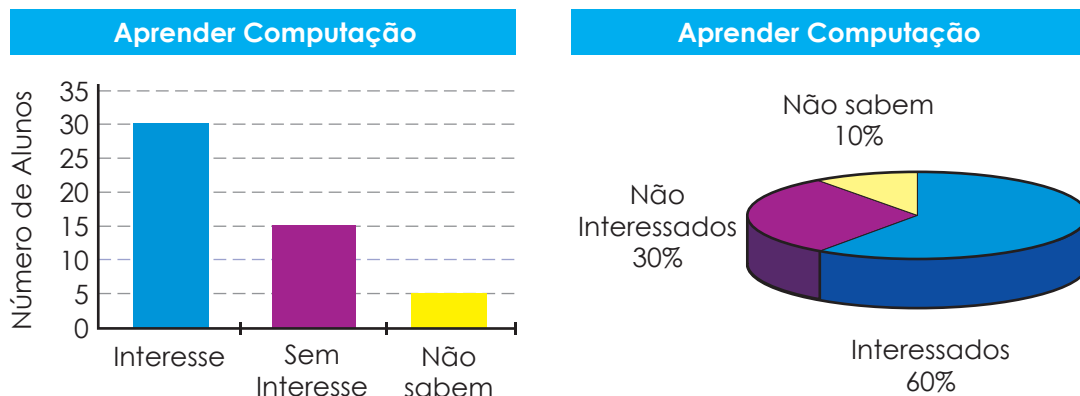
Cada parcela componente do total será expressa em graus, calculada através de uma regra de três:

$$\begin{array}{l} \text{Total} - 360^\circ \\ \text{Parte} - x^\circ \end{array}$$

Exemplo Prático: Em uma amostra com alunos do Ensino Médio, quando perguntados sobre o interesse em aprender computação, obteve-se como respostas: 30 alunos manifestaram interesse, 15 não demonstraram interesse e 5 não sabem. Representar graficamente:



Comparando os dois gráficos



Observe acima que comparamos os gráficos. Essa prática, é feita quando desenhamos dois gráficos, lado a lado, para podermos estabelecer melhor a comparação de um fenômeno.

## NOÇÕES BÁSICAS SOBRE AS PLANILHAS NO EXCEL

O Excel é uma poderosa ferramenta para se organizar informações, fazer cálculos e criar gráficos. Seus documentos são chamados de planilhas, que são grandes tabelas onde podem ser colocados textos, números, figuras, gráficos, fórmulas e muito mais. Conhecer o Excel e seus fundamentos básicos é certamente um grande diferencial para os profissionais de qualquer área.

### A tela principal

Para abrir o Excel, clique no botão Iniciar, Todos os programas e Microsoft Excel. Você verá, então, a tela principal, com menus parecidos com os do Word, alguns botões, a barra de fórmulas, uma planilha em branco, as abas de planilhas e uma linha de informações.

O que é e para que serve uma planilha.

	A	B	C
1			
2			
3			
4			
5			
6			
7			

Você deve notar, ao olhar para a planilha em branco que se abre quando iniciamos o Excel, que ela é formada por retângulos, denominados células. Elas podem conter textos, fórmulas e números. Cada célula é localizada pela letra que está acima dela e pelo número que está do seu lado esquerdo. Assim, em cima e à esquerda temos a célula A1, ao lado direito dela a célula B1 e assim por diante.

Para você compreender como funciona esta planilha, vamos criar um exemplo simples:

- Clique sobre a célula A1 e digite: MINHAS NOTAS;
- Depois, clique sobre a célula A3 e digite: Matemática.
- Pressione Enter, para ir para a linha de baixo (ou clique sobre a célula A4) ;
- Digite Português Enter novamente, Informática;
- Enter de novo e Física.

Agora, vamos aumentar um pouco a largura da coluna A. Para isso, passe com o mouse entre as letras A e B até que o cursor mude para uma barra vertical com duas setas ao lado. Agora, clique e arraste para a direita, até que as palavras que digitamos caibam com certa folga na coluna. O mesmo poderia ser feito com o menu Formatar, Coluna, Largura e digitando o valor 14, por exemplo.

A nossa planilha deve estar assim:

	A	
1	MINHAS NOTAS	
2		
3	Matemática	
4	Português	
5	Informática	
6	Física	
7		
8		

Agora, na célula B2, digite Prova 1, na célula C2, digite Prova 2 e na célula D2, digite Média. Agora, complete com as seguintes notas:

	A	B	C	D
1	MINHAS NOTAS			
2		Prova 1	Prova 2	Média
3	Matemática	95	86	
4	Português	90	100	
5	Informática	100	100	
6	Física	80	89	
7				
8				

Vamos, então, digitar a nossa primeira fórmula: em D3, digite:

$$=(B3+C3)/2$$





Quando você pressionar enter, o valor da média aritmética entre 95 e 86 irá aparecer na célula D3. Note que para digitar uma fórmula, você começa digitando o sinal de igual, depois a operação que você quer fazer com as células. Você somou os valores que estavam nas células B3 e C3, e depois dividiu o resultado por 2 para fazer a média. A divisão é representada pela barra ( / ) e os parênteses são necessários para que o Excel calcule primeiro a soma e só depois divida por 2. Caso não usássemos os parênteses, o valor da célula B3 seria somado à metade do valor da célula C3, o que não desejamos

Operação	Excel
Soma de A com B	A+B
Subtração entre A e B	A - B
Multiplicação de A por B	A*B
Divisão de A por B	A/B
A elevado a B	A^B




Bem, agora vamos usar os recursos do Excel para que não tenhamos que digitar todas as fórmulas! Clique sobre a célula D3. Você notou que ao redor dela aparece um retângulo mais grosso com um pequeno quadrado do lado direito e embaixo da célula? Clique nesse quadrado e arraste para baixo, até a célula D6. Veja o que aconteceu: as células D4, D5 e D6 foram automaticamente criadas copiando a fórmula em D3, mas modificando o número da linha! A célula D4, por exemplo, ficou assim:

$= (B4 + C4) / 2$



Este recurso é muito poderoso e economiza um grande trabalho na hora de elaborar planilhas maiores!

Agora, vamos deixar a nossa planilha mais bonita. Primeiro, selecione as células de A3 até A6 clicando sobre a célula A3 e arrastando até A6. Agora pressione o botão  para deixar os nomes das matérias em negrito. Pressione também o botão  para alinhar as matérias pela direita. Selecione agora as células B2 até D2 e pressione novamente  e  para centralizar.

Selecione todas as notas (de B3 até D6) e clique em  para centralizar as notas.

Agora, selecione as células de A1 até D1 e clique no botão , que irá agrupar as células (chama-se mesclar as células) e nos botões  e .

Ainda com as células selecionadas, clique na célula de tamanho da fonte  10 e selecione o tamanho 14.

Agora vamos colocar bordas para melhorar ainda mais o visual. Para isso, selecione as células de A1 a D6 e clique no triângulo preto pequeno ao lado do botão  , para escolher o tipo de borda e clique em .

A nossa planilha está assim:

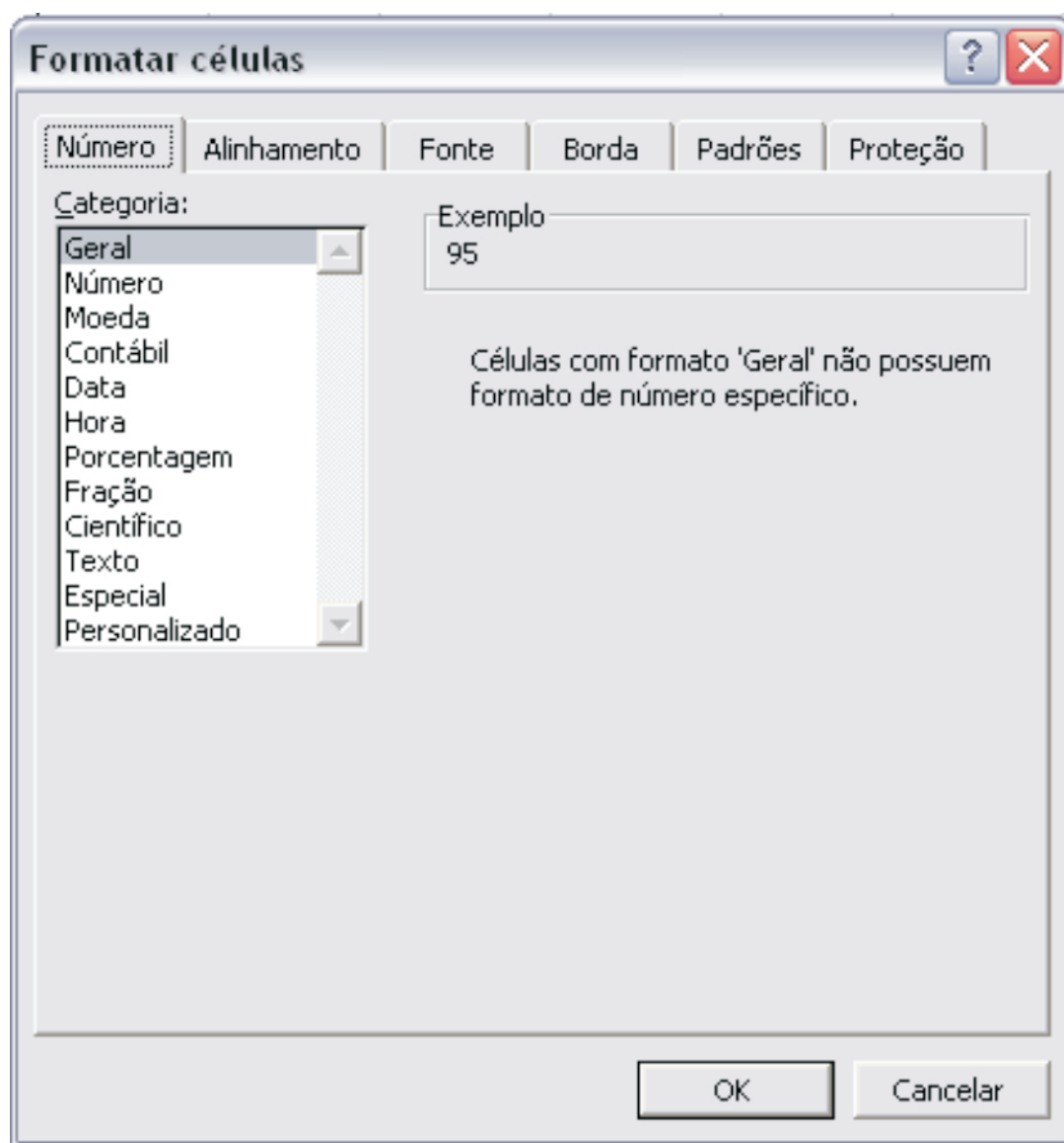
	A	B	C	D
1	<b>MINHAS NOTAS</b>			
2		<b>Prova 1</b>	<b>Prova 2</b>	<b>Média</b>
3	<b>Matemática</b>	95	86	90,5
4	<b>Português</b>	90	100	95
5	<b>Informática</b>	100	100	100
6	<b>Física</b>	80	89	84,5
7				
8				



Agora experimente mudar um número, por exemplo, trocar o 80 de Física por 100. Veja que a média se altera automaticamente. Esse é um recurso muito interessante, pois torna os dados das planilhas sempre atualizados! Não se esqueça de salvar a sua planilha, clicando em Arquivo, Salvar. O procedimento é igual ao do Word.


## Formatação de números e ordenação alfabética

Abra a planilha que você salvou, com as notas, caso a tenha fechado. Agora, selecione as células de B3 até D6. Vá até o menu Formatar, Células. Abre-se uma janela que contém várias opções para formatação das células



Algumas das operações que fizemos para embelezar a nossa planilha estão também disponíveis ali. Mas clique em Número, e em casas decimais abaixo para 1. Clique em OK para ver o resultado. Agora as notas aparecem com uma casa decimal. Isso ajuda na visualização, pois todas as notas têm agora o mesmo aspecto.


Explore também as outras opções desta janela, formatando os números como Moeda, Porcentagem, alterando a fonte (tipo de letra utilizado) etc.

Agora, experimente selecionar as células A3 até D6 e clicar no botão . Ele serve para ordenar os nomes em ordem alfabética. Agora, a ordem fica Física, Informática, Matemática e Português.

## Cores

Vamos colorir um pouco o nosso trabalho. Para isso, selecione todas as células da planilha clicando no retângulo que aparece à esquerda do A e acima do 1 nos cabeçalhos de linhas e colunas:




Clique, então, no botão com o baldinho . Isso faz com que todas as células sejam pintadas de branco, realçando a nossa planilha. Depois, selecione a célula A1 (que está mesclada com as células à sua direita) e no triângulo preto ao lado do botão do baldinho. Selecione a cor azul pálido, o azul mais de baixo da paleta de cores.

Mude as cores de outras células e tente reproduzir esta visualização:

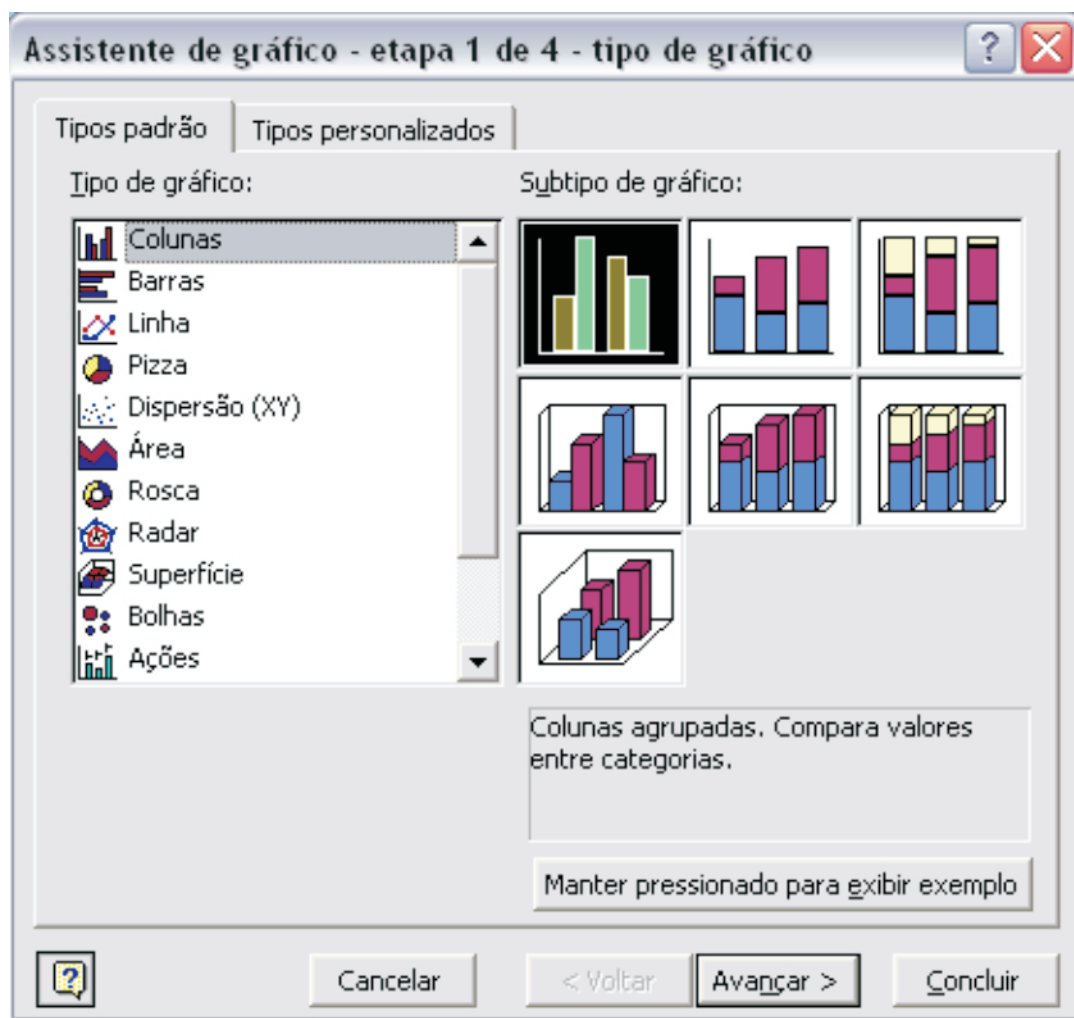
	A	B	C	D	
1	MINHAS NOTAS				
2		Prova 1	Prova 2	Média	
3	Física	100,0	89,0	94,5	
4	Informática	100,0	100,0	100,0	
5	Matemática	95,0	86,0	90,5	
6	Português	90,0	100,0	95,0	
7					
8					

Experimente, também trocar as cores das letras e números com o botão .

## Gráficos no Excel

Finalmente, vamos a um recurso muito utilizado, que é a capacidade do Excel de gerar gráficos a partir dos dados de uma planilha. Para isso, selecione as células de A1. até D6 e clique no botão . Abre-se, então, a janela de criação de gráficos.





Há vários tipos de gráficos que podem ser criados (Colunas, Barras, Linha etc.) e para cada tipo há várias opções (subtipos). Para ver como será o gráfico com o seu conjunto de dados em cada tipo e subtipo, selecione a opção desejada, clique e segure o botão

Manter pressionado para exibir exemplo

Vamos criar um gráfico de colunas com o primeiro sub-tipo (colunas agrupadas). Para isso, clique em Avançar.

Selecione, então, séries em linhas (o que neste caso agrupa as notas da Prova 1, Prova 2 e Média) e clique em Avançar.

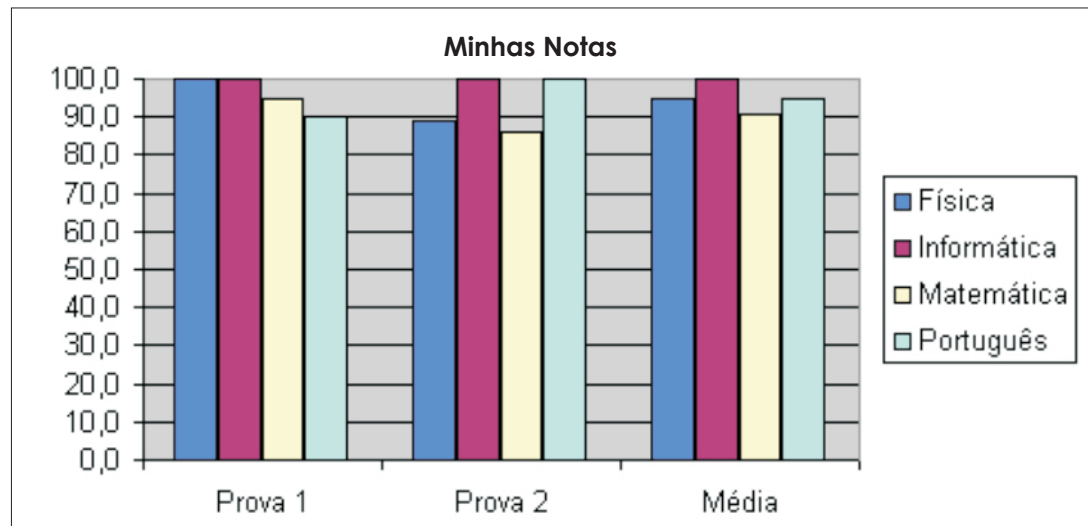
Em Título do gráfico, digite Minhas Notas. Se desejar dar nomes aos eixos x e y, digite-os nos campos indicados. Clique novamente em Avançar e Concluir.

Seu gráfico está criado e é um objeto que pode ser movimentado na planilha,

redimensionado e apagado, se você desejar.

Vamos agora dar o último toque ao gráfico, clicando com o botão direito do mouse sobre o valor 105 do eixo vertical. Depois, clique com o botão esquerdo em Formatar eixo. Clique sobre a palavra Escala, na parte superior. Em mínimo, onde está 75, mude para 0 e em máximo, onde está 105, mude para 100.

O nosso gráfico ficou assim:



Agora você pode explorar o Excel, criando planilhas, fazendo gráficos e experimentando sem medo de errar!

**ATIVIDADE**

### Em síntese

**Tabela** é um quadro que resume um conjunto de observações.

- **Série estatística** é toda a tabela que apresenta a distribuição de um conjunto de dados estatísticos em função da época, local, ou da espécie.

- **As séries podem ser divididas em:**

temporal, histórica ou cronológica  
geográfica, territorial ou de localidade  
específica ou categórica

**Gráfico** é um complemento da tabela e deve apresentar simplicidade, clareza e veracidade.

Os gráficos podem ser classificados em:

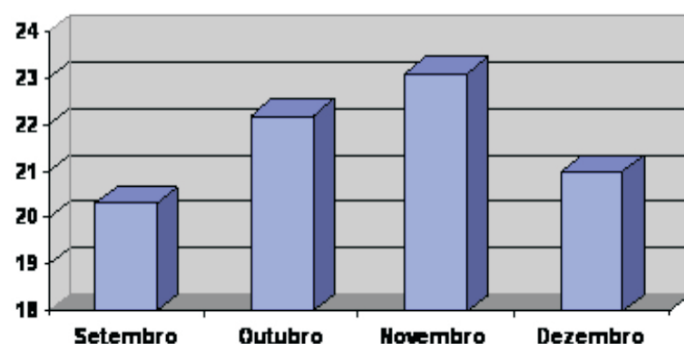
gráfico em curvas ou linhas  
gráfico em colunas  
gráfico em barras  
gráfico em setores

**Responda às questões abaixo e confira as respostas, no gabarito, no final do livro**

**6) Classifique a série abaixo:**

Aplicações em milhões de reais do último quadrimestre de 2007 num determinado banco

Mês	Setembro	Outubro	Novembro	Dezembro
Aplicações	20,3	22,2	23,1	21,0



- a) cronológica
- b) geográfica
- c) específica
- d) distribuição de frequência
- e) nenhuma das anteriores

Fonte: dados fictícios

**7) Classifique a série abaixo:**

Exportação em milhões de dólares por parte de empresas no ano 2006.

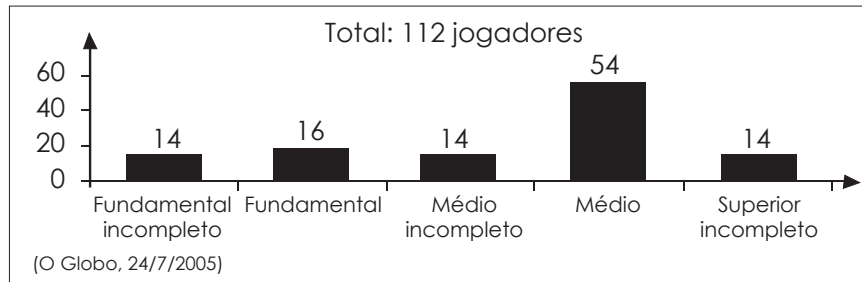
Porte	Exportações
Pequeno	100
Médio	150
Grande	200

- a) cronológica
- b) geográfica
- c) específica
- d) distribuição de frequência
- e) nenhuma das anteriores

Fonte: dados fictícios

8) A escolaridade dos jogadores de futebol nos grandes centros é maior do que se imagina, como mostra a pesquisa abaixo, realizada com os jogadores profissionais dos quatro principais clubes de futebol do rio de Janeiro.

De acordo com esses dados, o percentual dos jogadores dos quatro clubes que concluíram o Ensino médio é de aproximadamente:



- a) 14%.
- b) 48%.
- c) 54%.
- d) 60%.
- e) 68%.

9) Por que é importante fazermos a representação gráfica de uma série de dados estatísticos:

- a) Para que a apresentação fique mais bonita.
- b) Para que seja possível uma melhor visualização dos fenômenos em estudo.
- c) Porque toda apresentação deve ter gráficos.
- d) Porque uma apresentação com gráficos impressiona mais os leitores.
- e) Nenhuma das alternativas está correta.

10) Os dados abaixo referem-se às vendas de uma equipe da empresa-WB. construa o gráfico de linhas da tabela abaixo:

VEND.	MÊS											
	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
A	55	48	56	34	47	32	63	28	39	32	42	67
B	52	52	34	57	51	48	33	36	40	53	42	29
C	45	44	38	50	59	33	25	53	32	41	53	38
D	47	32	59	54	47	70	41	57	51	35	46	57
E	58	63	51	43	37	29	44	26	45	36	37	48

Fonte: dados fictícios

# UNIDADE

DISTRIBUIÇÃO

3



Nesta unidade, abordaremos Distribuição, com o objetivo de produzir visualização dos dados e posterior análise de informações.

## DISTRIBUIÇÃO DE FREQUÊNCIAS

Um conjunto de observações de certo fenômeno, não estando adequadamente organizado, fornece pouca informação de interesse ao pesquisador e ao leitor. Para uma visão rápida e global do fenômeno em estudo é preciso que os dados estejam organizados. Uma das formas de se fazer a organização dos dados coletados em uma pesquisa é através das distribuições de frequência.

### Representação dos Dados (Amostrais ou Popacionais)

Dados brutos: são aqueles que não foram numericamente organizados, ou seja, estão na forma com que foram coletados.

Por exemplo: Números de filhos de um grupo de 50 casais

2	3	0	2	1	1	1	3	2	5
6	1	1	4	0	1	5	6	0	2
1	4	1	3	1	7	6	2	0	1
3	1	3	5	7	1	3	1	1	0
3	0	4	1	2	2	1	2	3	2

**Rol:** é a organização dos dados brutos em ordem de grandeza crescente ou decrescente.

Por exemplo: Números de filhos de um grupo de 50 casais

0	0	0	0	0	0	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	2	2	2	2	2	2	2	2
2	3	3	3	3	3	3	3	3	4
4	4	5	5	5	6	6	6	7	7

Distribuição de frequência sem intervalos de classe: É a simples condensação dos dados conforme as repetições de seus valores. Para um rol de tamanho razoável esta distribuição de frequência é inconveniente, já que exige muito espaço. Veja exemplo abaixo:

Número de pontos obtidos pelos alunos na disciplina-X, colégio-Y, em 2007.

Pontos	Número de Alunos
4	8
5	10
6	7
7	5
8	8
9	5
10	7

Nota: dados hipotéticos

Distribuição de frequência com intervalos de classe: Quando o tamanho da amostra é elevado e o número de variáveis é muito grande, é mais racional efetuar o agrupamento dos valores em vários intervalos de classe.

Veja o Exemplo:

Peso (Kg)	Número de Alunos
40├ 50	30
50├ 60	25
60├ 70	13
70├ 80	12
80├ 90	0

Nota: dados hipotéticos

Vamos, agora, analisar um exemplo prático de coleta de dados e organização destes valores em tabelas de frequências com intervalos de classe.

**Dados Brutos:** Taxas municipais de urbanização (em percentual) no Estado X –2005

8	24	46	13	38	54	44	20	17	14
18	15	30	24	20	8	24	18	9	10
38	79	15	62	23	13	62	18	8	22
11	17	9	35	23	22	37	36	8	13
10	6	92	16	15	23	37	36	8	13
44	17	9	30	26	18	37	43	14	9
28	41	42	35	35	42	71	50	52	17
19	7	28	23	29	29	58	77	72	34
12	40	25	7	32	34	22	7	44	15
9	16	31	30						

Nota: dados hipotéticos

**Rol:** Taxas municipais de urbanização (em percentual) no Estado X – 2005

6	6	7	7	7	8	8	8	8	9
9	9	9	9	10	10	11	12	13	13
13	13	14	14	14	15	15	15	15	16
16	17	17	17	17	18	18	18	18	19
20	20	22	22	22	23	23	23	23	24
24	24	25	26	28	28	29	29	30	30
30	31	32	34	34	34	35	35	35	36
37	37	38	38	40	41	42	42	43	44
44	44	46	50	52	54	58	62	62	71
72	77	79	92						

Nota: dados hipotéticos



## Distribuição de freqüências para dados agrupados em classes:

TAXAS (em %)	Número de Municípios ( $f_i$ )
6  — 16	29
16  — 26	24
26  — 36	16
36  — 46	13
46  — 56	4
56  — 66	3
66  — 76	2
76  — 86	2
86  — 96	1
TOTAL ( $\Sigma$ )	94

Observação:

Representaremos o somatório pela letra grega maiúscula sigma (  $\Sigma$  )

Elementos de uma Distribuição de Freqüência

- **classe:** são intervalos de variação da variável. As classes são representadas simbolicamente por  $i$ , sendo  $i = 1, 2, 3, \dots, k$  (onde  $k$  é o número total de classes da distribuição).

Pelo exemplo prático anterior o intervalo 16 |— 26 define a segunda classe ( $i = 2$ ).

A distribuição é formada por nove classes, podemos afirmar que  $i = 9$ .

O símbolo |— representa um intervalo fechado à esquerda e aberto à direita. Significa que o número à esquerda pertence à classe e o número à direita não pertence.

**Limites de classe:** são os extremos de cada classe. O menor número é o limite inferior da classe (  $li$  ) e o maior número, o limite superior da classe (  $ls$  ).

Por exemplo

Pelo exemplo prático anterior na terceira classe, temos:  $li_3 = 26$  e  $ls_3 = 36$

**Amplitude de um intervalo de classe (h):** é a medida de intervalo que define a classe. Ela é obtida pela diferença entre os limites superior e inferior. Assim:

$$h = ls - li$$

Exemplo: A amplitude da classe descrita acima é 10, pois é a diferença do limite superior com o inferior.

- **Amplitude total (H):** é a diferença entre o valor máximo e o valor mínimo da amostra.

$$H = L_s - L_i$$

Pelo exemplo prático anterior, sabemos que  $L_s = 92$  e  $L_i = 6$ , logo amplitude total é  $H = 92 - 6 = 86$

Ponto médio de uma classe ( $x_i$ ): é o ponto que divide o intervalo de classe em duas partes iguais.

$$x_i = \frac{L_i + L_s}{2}$$

Pelo exemplo prático anterior na quinta classe, temos:  $L_i = 46$  e  $L_s = 56$ , logo o ponto médio dessa classe é

$$x_i = \frac{46 + 56}{2} = \frac{102}{2} = 51$$

#### Em síntese:

- Dados brutos são aqueles que não foram numericamente organizados.
- Rol é a organização dos dados brutos em ordem de grandeza crescente ou decrescente.
- Distribuição de freqüências pode ser com ou sem intervalos de classe.
- Os elementos da distribuição de freqüências são:
  - classe:** são intervalos de variação da variável
  - limites da classe:** são os extremos de cada classe
  - amplitude de um intervalo:** é a diferença entre o limite superior e inferior  
 $h = L_s - L_i$
  - amplitude total:** é a diferença entre o maior e o menor da amostra  
 $H = L_s - L_i$
  - ponto médio:** é a média aritmética dos limites da classe  $x_i = \frac{L_s + L_i}{2}$

## DISTRIBUIÇÃO DE FREQUÊNCIAS – NÚMEROS DE CLASSE – TIPOS DE FREQUÊNCIA

### Determinação do Número de Classes (K)

É importante que a distribuição conte com um número adequado de classes. Se o número de classes for excessivamente pequeno acarretará perda de detalhe e pouca informação se poderá extrair da tabela. Por outro lado, se for utilizado um número excessivo de classes, haverá alguma classe com freqüência nula ou muito pequena, não atingindo o objetivo de classificação que é tornar o conjunto de dados supervisionáveis.

Não há uma fórmula exata para determinar o número de classes. Três soluções são apresentadas abaixo:

1) Para  $n \leq 25 \rightarrow K = 5$  e para  $n > 25 \rightarrow K = \sqrt{n}$

Esse processo é menos preciso.

Exemplo:

Se a amostra tiver 23 elementos analisados o número de classes é 5, pois  $n < 25$ .

Suponha que a amostra tenha 83 elementos analisados ( $n > 25$ ) o número de classes é calculado por  $\sqrt{83} = 9,1104335 \cong 9$

2) Pode-se utilizar a regra de Sturges, que fornece o número de classes em função do total de observações:

$$K = 1 + 3,3 \cdot \log n$$

Onde: K é o número de classes;  
n é o número total de observações.

Para facilitar o cálculo do número de classes pela regra de Sturges, utilize a tabela abaixo:

n = total de observações	k = nº total de classes a usar
1	1
2	2
3 ┤ 5	3
6 ┤ 11	4
12 ┤ 22	5
23 ┤ 46	6
47 ┤ 90	7
91 ┤ 181	8
182 ┤ 362	9
363 ┤ 724	10
725 ┤ 1448	11
1449 ┤ 2896	12

observação:

- ┤ Inclui tanto o valor da direita quanto o da esquerda
- Não inclui nem o valor da direita, nem o da esquerda
- ├ Inclui o valor da direita mas não o da esquerda
- ┌ Inclui o valor da esquerda mas não o da direita

A fórmula de Sturges revela um inconveniente, propõe um número demasiado de classes para um número pequeno de observações e relativamente poucas classes, quando o total de observações for muito grande.

Exemplo:

Se a amostra tiver 94 elementos analisados, o cálculo do número de classes pela

fórmula de Sturges ficará da seguinte maneira:

$$K = 1 + 3,3 \cdot \log n \quad K = 1 + 3,3 \cdot \log 94 \quad K = 1 + 3,3 \cdot 1,97313 \quad K = 7,51 \quad K = 8$$

Também, percebemos pela tabela que 94 é um número entre 91 — 181, logo teremos 8 classes formando a tabela de freqüências.

3) Truman L. Kelley, sugere os seguintes números de classes, com base no número total de observações, para efeito de representação gráfica:

n	5	10	25	50	100	200	500
K	2	4	6	8	10	12	15

Exemplo:

Se a amostra tiver 50 elementos analisados o número de classes é 8, conforme tabela acima.



Qualquer regra para determinação do número de classes da tabela não nos leva a uma decisão final, esta vai depender, na realidade, de um julgamento pessoal, que deve estar ligado à natureza dos dados.

**Amplitude do intervalo de classe ( $A_i$ ):** é o comprimento da classe.

$$A_i = \frac{H}{K}$$

Observação: convém arredondar o número correspondente à amplitude do intervalo de classe para facilitar os cálculos (arredondamento arbitrário).

Exemplo prático:

Antes de enviar um lote de aparelhos elétricos para venda, o Departamento de Inspeção da empresa produtora selecionou uma amostra casual de 32 aparelhos avaliando o desempenho através de uma medida específica, obtendo os seguintes resultados:

154	155	156	164	165	170	172	175
175	176	178	178	180	180	180	184
190	190	190	192	195	198	200	200
202	205	205	210	211	212	215	218



Construir uma tabela de distribuição de freqüências com intervalos de classes.

1ª passo: A amplitude total será dada por:  
 $H = L_s - L_i \longrightarrow H = 218 - 154 = 64$

2ª passo: Neste caso,  $n = 32 \longrightarrow$  pela regra de Sturges (consultar Tabela)  
 $K = 6$

3ª passo: A amplitude do intervalo de cada classe será:  $Ai = \frac{H}{K} = \frac{64}{6} = 10,66... \approx 11$

4ª passo: Construir a tabela de distribuição de freqüências com intervalos de classes:

Classes	Freqüência
154  —— 165	4
165  —— 176	5
176  —— 187	7
187  —— 198	5
198  —— 209	5
209  —— 220	6
Total (Σ)	32

Perceba que utilizamos o menor valor do Rol, para iniciar a 1ª classe e a amplitude do intervalo encontrado para formar as outras classes que completam a tabela.

## Tipos de Freqüências

**Freqüências simples ou absolutas (fi):** é o número de repetições de um valor individual ou de uma classe de valores da variável. A soma das freqüências simples é igual ao número total dos dados da distribuição.

$$\sum fi = n$$

**Freqüências relativas (fri):** são os valores das razões (divisões) entre as freqüências absolutas de cada classe e a freqüência total da distribuição. A soma das freqüências relativas é igual a 1 ou 100 %

Freqüência relativa:  $fri = \frac{fi}{\sum fi}$

Freqüência relativa %:  $fri \% = \frac{fi}{\sum fi} \cdot 100$

**Freqüência simples acumulada (faci):** é o total das freqüências de todos os valores inferiores ao limite superior do intervalo de uma determinada classe.

**Freqüência relativa acumulada (fraci):** é a freqüência acumulada da classe, dividida pela freqüência total da distribuição.

$$fraci = \frac{fac}{\sum fi}$$



Vamos colocar em prática, juntos, todos esses conceitos que estudamos nas aulas 5, 6, 7 e 8 através do seguinte exercício:

O departamento comercial de uma empresa pediu o levantamento das vendas diárias (em milhares de reais) durante

os 25 dias do mês de novembro de 2007. Assim obteve os seguintes valores:

280	305	320	330	310	340	330	341	369	355
360	370	365	280	375	380	400	371	390	400
370	401	420	430	375					

Pede-se:

- 1) Montar o Rol
- 1) Calcular amplitude total.
- 2) Calcular o número de classes.
- 3) Calcular a amplitude do intervalo de cada classes.
- 4) Construir uma tabela com todos os tipos de freqüências (simples, relativa, acumulada e acumulada relativa)

Solução:

O Rol é:


A amplitude total será dada por:

Neste caso,  $n = 25 \rightarrow$  pela regra de Sturges, temos:

Amplitude do intervalo de cada classe será:

Utilizando o menor valor do Rol, para iniciar a 1ª classe e a amplitude do intervalo de cada classe, e completando a tabela com as freqüências temos:

Classes	fi	fri	fri(%)	faci	fraci	xi
Total ( $\Sigma$ )						

## Em Síntese

- Para determinar o número de classes temos três casos:

1º caso

Para  $n \leq 25 \rightarrow$  número de classes é  $K = 5$

Para  $n > 25 \rightarrow$  número de classes é  $K \cong \sqrt{n}$

2º caso

Pela regra de sturges  $\rightarrow K = 1 + 3,3 \log n$

3º caso

Pela regra de Truman L. Kelley  $\rightarrow$  conforme a tabela abaixo:

n	5	10	25	50	100	200	500
K	2	4	6	8	10	12	15

- **Amplitude do intervalo de classe:** é o comprimento da classe, calculado por:  $Ai = \frac{H}{K}$

- **Freqüência simples ou absoluta** é o número de repetições de um valor individual.

- **Freqüência relativa** são os valores das divisões entre as freqüências absolutas de cada classe e a freqüência total da distribuição.

- **Freqüência simples acumulada** é o total das freqüências de todos os valores inferiores ao limite superior do intervalo de uma determinada classe.

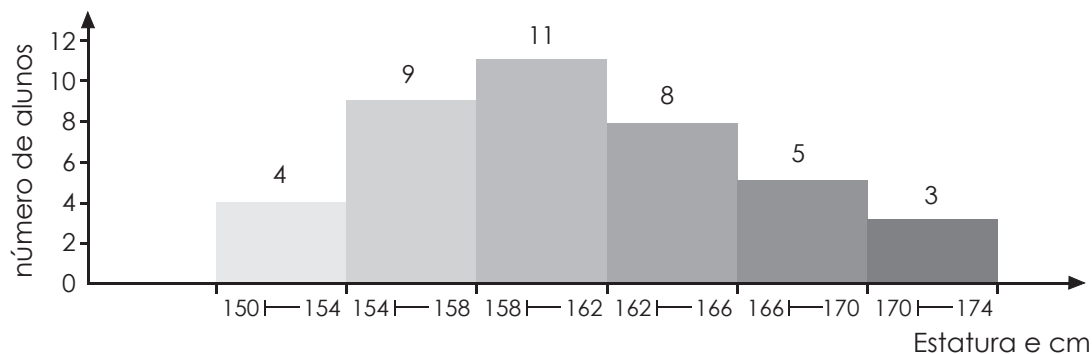
- **Freqüência relativa acumulada** é a freqüência acumulada da classe dividida pela freqüência total da distribuição.

## DISTRIBUIÇÃO DE FREQÜÊNCIAS – HISTOGRAMA E POLÍGONO DE FREQÜÊNCIAS

Histograma: é um gráfico formado por um conjunto de retângulos justapostos e é muito utilizado para representar a distribuição de freqüências cujos dados foram agrupados em classes ou intervalos de mesma amplitude. A base do retângulo é igual à amplitude do intervalo classe e altura proporcional à freqüência da classe.

Exemplo:

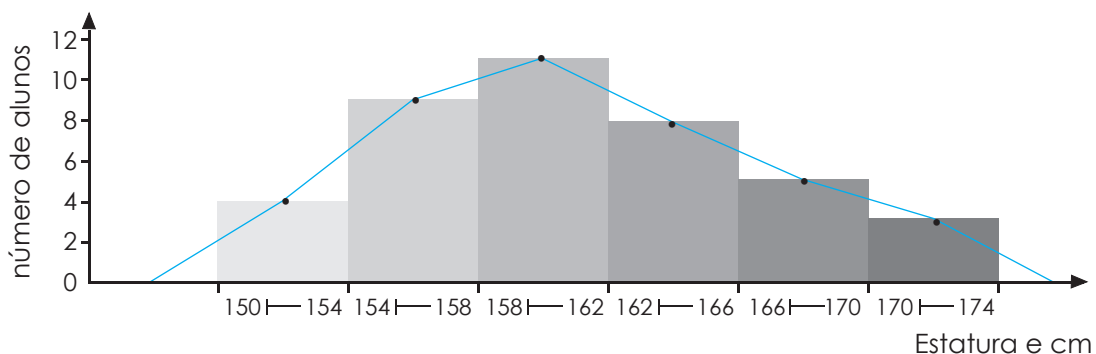
Estaturas dos alunos da Turma “A” – 2006



**Polígono de Frequências:** é obtido unindo-se por segmentos de reta os pontos médios das bases superiores dos retângulos de histograma.

Exemplo:

Estaturas dos alunos da Turma "A" – 2006



#### ATIVIDADE

Vamos colocar em prática, juntos, todos esses conceitos desta aula através dos seguintes exercícios:

1) Numa empresa, foi observada a seguinte tabela de salários semanais (em reais) de 40 operários não-especializados.

142	143	144	145	147	148	148	149	149	150
150	151	152	152	163	163	163	164	164	164
164	165	165	165	170	175	175	184	184	184
190	190	190	195	197	197	200	207	210	210

Forme com esses dados uma tabela de dados agrupados por classes e depois construa o histograma e polígono de frequência correspondente.

Solução: Antes de construirmos o histograma e o polígono de frequências, temos que montar a tabela com dados agrupados por classes, desta forma faremos inicialmente os seguintes cálculos:

A amplitude total será dada por:

Neste caso,  $n = 40$  pela regra de Sturges, temos:

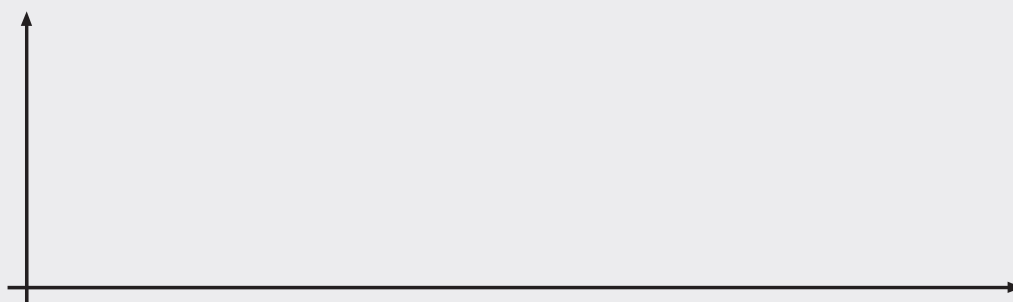
A amplitude do intervalo de cada classe será:



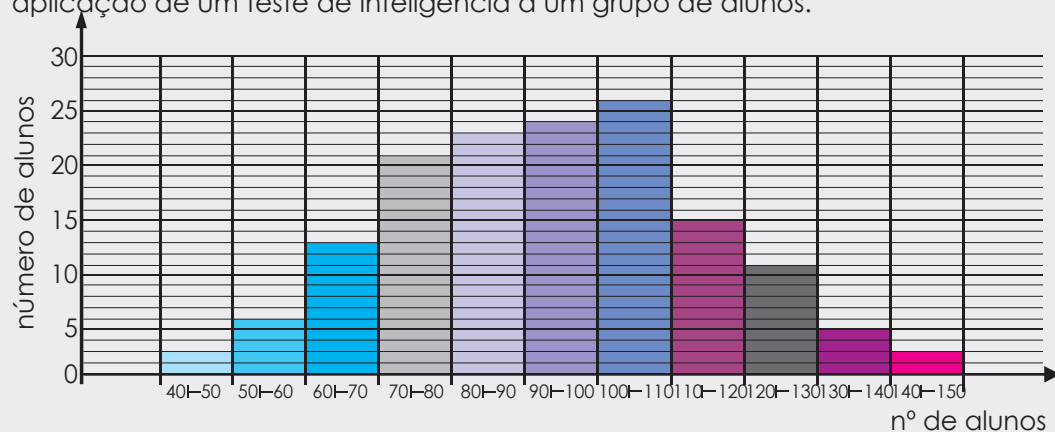
Utilizando o menor valor do Rol, para iniciar a 1ª classe e a amplitude do intervalo de cada classe, e completando a tabela com as frequências temos:

Classes	fi
Total ( $\Sigma$ )	

Depois de montarmos a tabela com dados agrupados em classes, iremos construir o histograma e o polígono de frequência:



2) Examine o histograma abaixo, que corresponde às notas relativas à aplicação de um teste de inteligência a um grupo de alunos.



Determine:

- O número de alunos: \_\_\_\_\_
- A frequência acumulada da quinta classe: \_\_\_\_\_
- A frequência relativa sétima classe: \_\_\_\_\_
- A porcentagem de alunos que atingiram pelo menos 90 pontos: \_\_\_\_\_
- O número de alunos que obtiveram pontuação menor que 110 pontos: \_\_\_\_\_

### Em síntese

- Histograma é um gráfico formado por um conjunto de retângulos justapostos.
- Polígono de freqüências é obtido unindo-se por segmentos de reta os pontos médios das bases superiores dos retângulos de um histograma.

### Responda às questões abaixo e confira as respostas, no gabarito, no final do livro

11) Quando um intervalo de classe está representado por  $1,250 \mid \text{---} 1,260$ , isto significa que:

- a) o valor 1,250 não faz parte do intervalo de classe.
- b) o valor 1,250 é o limite superior do intervalo de classe.
- c) o valor 1,260 faz parte do intervalo de classe.
- d) o valor 1,260 é o limite inferior do intervalo de classe.
- e) o ponto médio deste intervalo é o valor 1,255.

12) com relação à distribuição de freqüência dada abaixo, é possível dizer que:

Comprimento, (em cm) de certa peça produzida pela indústria "y"		
Comprimento (cm)		fi
12,32	$\mid \text{---} 12,36$	3
12,36	$\mid \text{---} 12,40$	5
12,40	$\mid \text{---} 12,44$	12
12,44	$\mid \text{---} 12,48$	15
12,48	$\mid \text{---} 12,52$	8
12,52	$\mid \text{---} 12,56$	2
		45

- a) a amplitude total da distribuição é 0,24.
- b) o limite inferior da 4a. classe é 12,44.
- c) o limite superior da 3a. classe é 12,44.
- d) a amplitude do intervalo de classe é 0,04.
- e) todas as alternativa estão corretas.

13) Dada a distribuição de freqüências a seguir,

Idades	Freqüência (f)
0 $\mid \text{---} 2$	2
2 $\mid \text{---} 4$	5
4 $\mid \text{---} 6$	18
6 $\mid \text{---} 8$	10
8 $\mid \text{---} 10$	5

responda qual o limite da quarta classe:

- a) 8
- b) 6
- c) 4
- d) 10
- e) 40

14) Na distribuição de freqüências da questão 3, qual a amplitude de cada classe ou intervalo?

- a) 10
- b) 1
- c) 2
- d) 40
- e) 8

15) Na distribuição de freqüências da questão 3, qual o ponto médio da quinta classe ou intervalo?

- a) 40
- b) 5
- c) 8
- d) 9
- e) 10

16) A tabela abaixo refere-se a uma pesquisa realizada com 20 alunos, a respeito da área da carreira universitária que cada um pretende seguir.

Área	Freqüência simples	Freqüência relativa	porcentagem
Humanas	8	0,4	40%
Biológicas		0,35	35%
Exatas	5		25%
Total	20	1,00	100%

A freqüência simples e a freqüência relativa que faltam (assinaladas por ) são respectivamente:

- a) 7 e  $\frac{5}{20}$
- b) 7 e  $\frac{7}{20}$
- c) 5 e  $\frac{7}{20}$
- d) 7 e  $\frac{8}{20}$
- e) 5 e  $\frac{5}{20}$

17. Se a distribuição de freqüência obtida através da entrevista com 35 pessoas, em relação ao número de irmãos que cada uma têm, é representada abaixo:

Números de irmãos				
Quantidade de irmãos	fi	fri	faci	fraci
0	2	0,0571	2	0,0571
1	5	0,1429	7	0,2000
2	8	1,2286	15	0,4286
3	12	0,3429	27	0,7715
4	3	0,0857	30	0,8572
5	3	0,0857	33	0,9429
6	1	0,0286	34	0,9715
7	1	0,0286	35	1,001
Total	35	1,0000		

Obs.: A última freqüência acumulada relativa resultou 1,0001 por critérios de arredondamento. É perfeitamente admissível desprezar a última casa decimal.

Em relação a esta distribuição é possível dizer que:

- a) temos 30 pessoas que possuem 3 irmãos.
- b) temos 8,57% das pessoas com, pelo menos, 3 irmãos.
- c) temos 94,29% das pessoas entrevistadas com 5 ou menos irmãos.
- d) temos 34 pessoas com mais de 6 irmãos.

18) dada a distribuição de freqüências a seguir,

Idade	Frequência
19 — 21	8
21 — 23	12
23 — 25	15
25 — 27	13
27 — 29	7
29 — 31	5

responda qual a frequência acumulada total:

- a) 31
- b) 55
- c) 20
- d) 60
- e) 12

As questões de número 19 e 20 devem ser respondidas com base na tabela primitiva de dados obtidos em uma pesquisa.

Massa, em Kg, de pacotes de certo produto produzido por uma indústria "X"

1,026	1,015	1,084	1,095	1,067	1,084	1,054	1,018	1,004	1,048
1,035	1,101	1,072	1,006	1,045	1,065	1,053	1,058	1,026	1,043
1,080	1,035	1,060	1,054	1,039	1,022	1,090	1,049	1,018	1,057
1,058	1,031	1,029	1,062	1,048	1,038	1,045	1,034	1,062	1,068
1,005	1,063	1,038	1,081	1,052	1,064	1,024	1,048	1,034	1,050

Fonte: Dados Fictícios

**19) O limite superior da quinta classe é:**

- a) 1,060
- b) 1,046
- c) 1,088
- d) 1,074
- e) 1,052

**20) As frequências da terceira e da quinta classe são, respectivamente:**

- a) 8 e 9
- b) 10 e 12
- c) 8 e 12
- d) 10 e 4
- e) 10 e 9

Análise os dados abaixo que se referem a uma amostra de óbitos/mês nos hospitais de uma grande capital brasileira, nos últimos 25 meses e faça o que se pede:

410, 440, 440, 584, 579, 577, 577, 454, 460, 460, 491, 561, 556, 491, 490, 491, 490, 553, 553, 540, 540, 520, 491, 520, 491.

**21) colocar os dados acima em forma de rol e calcular a amplitude total.**


**22) Calcular o número de classes.**

**23) calcular a amplitude das classes.**

**24) construir a tabela com os dados agrupados por classes.**

**25) construir o histograma e o polígono de frequência.**



# UNIDADE

MEDIDAS DE POSIÇÃO –  
MÉDIAS, MODA E MEDIANA

4





Na maior parte das vezes em que os dados estatísticos são analisados, procuramos obter um valor para representar um conjunto de dados. Este valor deve sintetizar, da melhor maneira possível, o comportamento do conjunto do qual ele é originário. Nem sempre os dados estudados têm um bom comportamento, isto pode fazer com que um único valor possa representá-lo ou não perante o grupo.

As medidas de posição mais importantes são as medidas de tendência central, que recebem tal denominação pelo fato de os dados observados tenderem, em geral, a se agrupar em torno dos valores centrais. Dentre as medidas de tendência central, destacam-se as seguintes: Médias, Moda e Mediana. Cada uma com um significado diferenciado, porém tendo como serventia representar um conjunto de dados.

## MÉDIAS

**Média Aritmética Simples:** para se obter a média aritmética simples de um conjunto de dados, devemos dividir a soma dos valores de todos os dados do conjunto pela quantidade deles.

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{N} \quad \text{OU} \quad \mu = \frac{\sum x_i}{N}$$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad \text{OU} \quad \bar{x} = \frac{\sum x_i}{n}$$

Onde:  $x_i$  = são os valores que a variável  $x$  assume  
 $n$  = o número de valores  
 $\bar{x}$  = é a medida aritmética da amostra  
 $\mu$  = é a média aritmética da população

Exemplo:

Sabendo-se que as vendas diárias da empresa A, durante uma semana, foram de 10, 14, 13, 15, 16, 18 e 12 unidades. Determinar a média de vendas nesta semana feitas pela empresa A:

Para obter a média aritmética simples das vendas, faremos o seguinte cálculo:

$x_1 = 10$ ,  $x_2 = 14$ ,  $x_3 = 13$ ,  $x_4 = 15$ ,  $x_5 = 16$ ,  $x_6 = 18$  e  $x_7 = 12$  e  $n = 7$ , logo:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{10+14+13+15+16+18+12}{7} = \frac{98}{7} = 14$$

**Média Aritmética Ponderada:** média ponderada é uma média aritmética na qual será atribuído um peso a cada valor da série.

$$\bar{x}_p = \frac{x_i \cdot p_i}{p_i}$$

Exemplo:

O capital da empresa está sendo formado pelos acionistas, por financiamentos e por debêntures. Cada tipo tem um custo diferente para a empresa, definido pela sua taxa de juros anual. Calcule a taxa de juros média do capital da empresa, considerando os dados apresentados na tabela seguinte:

Capital da Empresa	Participação	Taxa de Juros
Acionista	R\$ 1.000.000,00	12 %
Financiamento	R\$ 600.000,00	8 %
Debêntures	R\$ 400.000,00	14 %

A taxa de juros média é calculada pela seguinte relação:  $\bar{x} = \frac{x_1 \cdot p_1 + x_2 \cdot p_2 + x_3 \cdot p_3}{p_1 + p_2 + p_3}$

$$\bar{x} = \frac{12\% \cdot 1.000.000 + 8\% \cdot 600.000 + 14\% \cdot 400.000}{1.000.000 + 600.000 + 400.000} = 11,20\%$$

## Média aritmética para dados agrupados sem intervalos de classes

As frequências são as quantidades de vezes que a variável ocorre na coleta de dados, elas funcionam como fatores de ponderação, o que nos leva a calcular uma média aritmética ponderada.

$$\mu = \frac{\sum x_i f_i}{N} \quad (\text{População})$$

$$\bar{x} = \frac{\sum x_i f_i}{n} \quad (\text{amostra})$$

Exemplo:

Após ter sido realizado trabalho bimestral, numa turma de Estatística, o professor efetuou levantamento das notas obtidas pelos alunos, e observou a seguinte distribuição e calculou a média de sua turma:

Notas dos alunos - $x_i$	Números de alunos - $f_i$	$x_i f_i$
1	1	1
2	3	6
3	5	15
4	1	4
Total $\Sigma$	N= 10	26

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{26}{10} = 2,6$$

## Média aritmética para dados agrupados com intervalos de classes

Neste caso, convencionamos que todos os valores incluídos em um determinado intervalo de classe coincidem com o seu ponto médio, e determinamos a média aritmética ponderada por meio das seguintes fórmulas:

$$\mu = \frac{\sum x_i f_i}{N}, \text{ onde } x_i = \frac{l_i + l_s}{2} \quad (\text{População})$$

$$\bar{x} = \frac{\sum x_i f_i}{n}, \text{ onde } x_i = \frac{l_i + l_s}{2} \quad (\text{amostra})$$

Exemplo:

Determine a renda média familiar, de acordo com os dados da tabela:

Classes – Renda familiar	$x_i$	$f_i$ – Número de famílias	$x_i f_i$
2  — 4	3	5	15
4  — 6	5	10	50
6  — 8	7	14	98
8  — 10	9	8	72
10  — 12	11	3	33
Total $\Sigma$		$n = 40$	268

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{268}{40} = 6,7$$

## MODA (MO)

Define-se a moda como o valor que ocorre com maior frequência em um conjunto de dados.

### Moda – para dados não agrupados

Primeiramente os dados devem ser ordenados para, em seguida, observar o valor que tem maior frequência.

Exemplo: Calcular a moda dos seguintes conjuntos de dados:

$X = (4, 5, 5, 6, 6, 6, 7, 7, 8, 8) \rightarrow Mo = 6$  (o valor mais freqüente)

Esse conjunto é unimodal, pois apresenta apenas uma moda.

$Y = (1, 2, 2, 2, 3, 4, 4, 4, 5, 5, 6) \rightarrow Mo = 2$  e  $Mo = 4$  (valores mais freqüentes) Esse conjunto é bimodal, pois apresenta duas modas.

$Z = (1, 2, 2, 2, 3, 3, 3, 4, 4, 4, 5) \rightarrow Mo = 2, Mo = 3$  e  $Mo = 4$  (valores mais freqüentes)

Esse conjunto é plurimodal, pois apresenta mais de duas modas.

$W = (1, 2, 3, 4, 5, 6) \rightarrow$  Esse conjunto é amodal porque não apresenta um valor predominante.



A moda é utilizada quando desejamos obter uma medida rápida e aproximada de posição ou quando a medida de posição deva ser o valor mais típico da distribuição.

Já a média aritmética é a medida de posição que possui a maior estabilidade.

## MEDIANA (MD)

É uma medida de posição cujo número divide um conjunto de dados em duas

partes iguais. Portanto, a mediana se localiza no centro de um conjunto de números ordenados segundo uma ordem de grandeza.

Para se obter o elemento mediano de uma série deveremos seguir os seguintes passos:

- Se N for ímpar a mediana é o termo de ordem:

$$P = \frac{N+1}{2}$$

- Se o N for par a mediana é a média aritmética dos termos de ordem:

$$P_1 = \frac{N}{2} \quad \text{e} \quad P_2 = \frac{N}{2} + 1$$

Exemplos:

1) Determine o valor da mediana da série que é composta dos seguintes elementos: 56, 58, 62, 65 e 90.

$$N = 5 \text{ (ímpar)} \longrightarrow P = \frac{N+1}{2} = \frac{5+1}{2} = 3 \longrightarrow 3^\circ \text{ elemento} \longrightarrow Md = 62$$

2) Em uma pesquisa realizada a respeito de erros por folha, cometidos por digitadores, revelou as seguintes quantidades: 12, 12, 13, 13, 15, 16, 18 e 20. Determinar a quantidade mediana de falhas.

$$N = 8 \text{ (par)} \longrightarrow P_1 = \frac{N}{2} = \frac{8}{2} = 4 \longrightarrow 4^\circ \text{ elemento} \longrightarrow Md = 13$$

$$P_2 = \frac{N}{2} + 1 = \frac{8}{2} + 1 = 5 \longrightarrow 5^\circ \text{ elemento} \longrightarrow Md = 15$$

$$\text{Logo a mediana será } Md = \frac{13 + 15}{2} = \frac{28}{2} = 14$$

## Comparação entre Médio e Moda

Medida	Definição	Vantagens	Desvantagens
Média	Centro da distribuição de	1. reflete cada valor	1. é afetada por valores extremos
Mediana	Metade dos valores são maiores, metade menores	1. possui propriedades matemáticas atraentes 2. menos sensível a valores extremos do que a média	1. difícil de determinar para grande quantidade de dados
Moda	Valor mais freqüente	1. valor "típico": maior quantidade de valores concentrados neste ponto	1. não se presta a análise matemática 2. pode não ter moda para certos conjuntos de dados

## Em síntese

Média aritmética simples:  $\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{N}$  ou  $\mu = \frac{\sum x_i}{N}$  da população

Média aritmética simples:  $\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$  ou  $\bar{x} = \frac{\sum x_i}{n}$  da amostra

Média aritmética ponderada:  $\bar{x}_p = \frac{x_i \cdot p_i}{p_i}$

Média aritmética para dados agrupados sem intervalos de classes:

$$\mu = \frac{\sum x_i f_i}{N} \quad (\text{População})$$

$$\bar{x} = \frac{\sum x_i f_i}{n} \quad (\text{amostra})$$

Média aritmética para dados agrupados com intervalos de classes:

$$\mu = \frac{\sum x_i f_i}{N}, \text{ onde } x_i = \frac{l_i + l_s}{2} \quad (\text{População})$$

$$\bar{x} = \frac{\sum x_i f_i}{n}, \text{ onde } x_i = \frac{l_i + l_s}{2} \quad (\text{amostra})$$

**Moda** é o valor que se repete o maior número de vezes, entre os dados obtidos.

**Mediana** (md) representa o valor central entre os dados obtidos, estando esses dados em ordem crescente ou decrescente.

Para obter o elemento mediano devemos considerar:

1º caso: se o N for ímpar a mediana é o termo de ordem:  $P = \frac{N+1}{2}$

2º caso: se N for par a mediana é a medida aritmética dos termos de ordem:

$$P_1 = \frac{N}{2} \text{ e } P_2 = \frac{N}{2} + 1$$

## POSIÇÃO – QUARTIS, DECIS E PERCENTIS

As medidas de posição denominadas quartis, decis e percentis têm o mesmo princípio da mediana. Enquanto a mediana separa a distribuição em duas partes iguais, a característica principal de cada uma dessas medidas é que:

**Quartis:** dividem a distribuição em quatro partes iguais;

**Decis:** dividem em dez partes iguais;

**Percentis:** dividem em cem partes iguais.

Notações:

$Q_i$  = quartil de ordem  $i$ ;

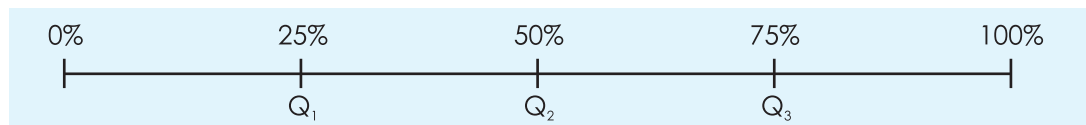
$D_i$  = decil de ordem  $i$ ;

$P_i$  = percentil de ordem  $i$ .

Essas medidas - os quartis, os decis e os percentis - são, juntamente com a mediana, conhecidas pelo nome genérico de separatrizes.

## Quartis para dados não agrupados em classes

Os quartis dividem um conjunto de dados em 4 partes iguais. Assim:



Para o cálculo das posições usaremos:

$$Q_1 \rightarrow P_1 = \frac{n+1}{4}$$

$$Q_2 \rightarrow P_2 = \frac{2 \cdot (n+1)}{4}$$

$$Q_3 \rightarrow P_3 = \frac{3 \cdot (n+1)}{4}$$

Onde  $n$  = número de dados (valores)

Exemplo:

Considere as seguintes notas de uma turma e calcule os quartis dessa amostra:

52,0	55,9	56,7	59,4	60,2	54,4	55,9	56,8	59,4	60,3
54,5	56,2	57,2	59,5	60,5	55,7	56,4	57,6	59,8	60,6
55,8	56,4	58,9	60,0	60,8					

Utilizando as relações para a posição dos quartis, temos:

$$Q_1 \rightarrow P_1 = \frac{n+1}{4} \rightarrow Q_1 \rightarrow P_1 = \frac{25+1}{4} = 6,5 \rightarrow P_1 = 7$$

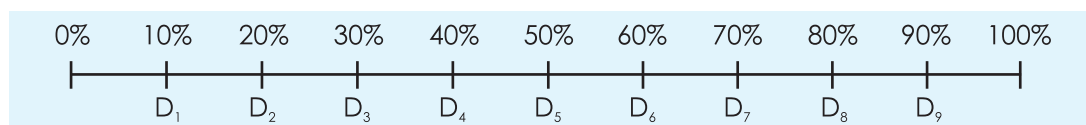
$$Q_2 \rightarrow P_2 = \frac{2 \cdot (n+1)}{4} \rightarrow P_2 = \frac{2 \cdot (25+1)}{4} = 13 \rightarrow P_2 = 13$$

$$Q_3 \rightarrow P_3 = \frac{3 \cdot (n+1)}{4} \rightarrow P_3 = \frac{3 \cdot (25+1)}{4} = 19,5 \rightarrow P_3 = 20$$

O primeiro quartil ( $Q_1$ ) abrange 25% dos valores da série, o segundo quartil 50% ( $Q_2$ ) e o terceiro ( $Q_3$ ) 75%.

## Decis para dados não agrupados

São valores que dividem a série em 10 partes iguais. Assim:



Para o cálculo das posições usaremos:

$$D_1 \rightarrow P_1 = \frac{n+1}{10} \quad D_2 \rightarrow P_2 = \frac{2 \cdot (n+1)}{10} \quad D_5 \rightarrow P_5 = \frac{5 \cdot (n+1)}{10} \quad D_9 \rightarrow P_9 = \frac{9 \cdot (n+1)}{10}$$

Onde n = número de dados (valores)

Exemplo:

Considere as seguintes notas de uma turma e calcule os decis dessa amostra:

52,0	55,9	56,7	59,4	60,2	54,4	55,9	56,8	59,4	60,3
54,5	56,2	57,2	59,5	60,5	55,7	56,4	57,6	59,8	60,6
55,8	56,4	58,9	60,0	60,8					

Utilizando as relações para a posição dos decis, temos:

$$D_1 \rightarrow P_1 = \frac{n+1}{10} \rightarrow P_1 = \frac{25+1}{10} = 2,6 \rightarrow P_1 = 3$$

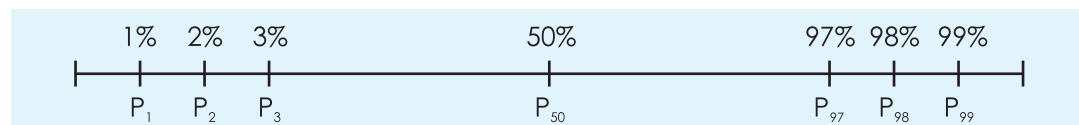
$$D_2 \rightarrow P_2 = \frac{2 \cdot (n+1)}{10} \rightarrow P_2 = \frac{2 \cdot (25+1)}{10} = 5,2 \rightarrow P_2 = 5$$

$$D_9 \rightarrow P_9 = \frac{9 \cdot (n+1)}{10} \rightarrow P_9 = \frac{9 \cdot (25+1)}{10} = 23,4 \rightarrow P_9 = 20$$

O primeiro decil ( $D_1$ ) abrange 10% dos termos da série, o segundo decil ( $D_2$ ) 20% e assim por diante, o nono ( $D_9$ ) 90%.

## Percentis para dados não agrupados

São as medidas que dividem a amostra em 100 partes iguais. Assim



Para o cálculo das posições usaremos:

$$P_1 \rightarrow P_1 = \frac{n+1}{100} \quad P_2 \rightarrow P_2 = \frac{2 \cdot (n+1)}{100} \quad P_{50} \rightarrow P_{50} = \frac{50 \cdot (n+1)}{100} \quad P_{99} \rightarrow P_{99} = \frac{99 \cdot (n+1)}{100}$$

Onde n = número de dados (valores)

Exemplo:

Considere as seguintes notas de uma turma e calcule os percentis dessa amostra:

52,0	55,9	56,7	59,4	60,2	54,4	55,9	56,8	59,4	60,3
54,5	56,2	57,2	59,5	60,5	55,7	56,4	57,6	59,8	60,6
55,8	56,4	58,9	60,0	60,8					

Utilizando as relações para a posição dos Percentis, temos:

$$P_{10} \rightarrow P_{10} = \frac{10.(n+1)}{100} \rightarrow P_{10} = \frac{10.(25+1)}{100} = 2,6 \rightarrow P_{10} = 3$$

$$P_{20} \rightarrow P_{20} = \frac{20.(n+1)}{100} \rightarrow P_{20} = \frac{20.(25+1)}{100} = 5,2 \rightarrow P_{20} = 5$$

$$P_{90} \rightarrow P_{90} = \frac{90.(n+1)}{100} \rightarrow P_{90} = \frac{90.(25+1)}{100} = 23,4 \rightarrow P_{90} = 23$$

O décimo percentil ( $P_{10}$ ) abrange 10% dos termos da série, o vigésimo percentil ( $P_{20}$ ) 20% e assim por diante, o nonagésimo ( $P_{90}$ ) 90%

### EM SÍNTESE

**Quartis** dividem a distribuição em quatro partes iguais. Para o cálculo das posições usaremos:

$$Q_1 \rightarrow P_1 = \frac{n+1}{4}$$

$$Q_2 \rightarrow P_2 = \frac{2.(n+1)}{4}$$

$$Q_3 \rightarrow P_3 = \frac{3.(n+1)}{4}$$

**Decis** dividem a distribuição em dez partes iguais. Para o cálculo das posições usaremos:

$$D_1 \rightarrow P_1 = \frac{n+1}{10}$$

$$D_2 \rightarrow P_2 = \frac{2.(n+1)}{10}$$

$$D_5 \rightarrow P_5 = \frac{5.(n+1)}{10}$$

$$D_9 \rightarrow P_9 = \frac{9.(n+1)}{10}$$

**Percentis** dividem a distribuição em cem partes iguais. Para o cálculo das posições usaremos:

$$P_1 \rightarrow P_1 = \frac{n+1}{100}$$

$$P_2 \rightarrow P_2 = \frac{2.(n+1)}{100}$$

$$P_{50} \rightarrow P_{50} = \frac{50.(n+1)}{100}$$

$$P_{99} \rightarrow P_{99} = \frac{99.(n+1)}{100}$$

## MEDIDAS DE DISPERSÃO (MEDIDAS DE VARIABILIDADE)

São medidas utilizadas para medir o grau de variabilidade, ou dispersão dos valores observados em torno da média aritmética. Servem para medir a representatividade da média e proporcionam conhecer o nível de homogeneidade ou heterogeneidade dentro de cada grupo analisado. Para compreender esse conceito, considere a seguinte situação:

Um empresário deseja comparar a performance de dois empregados, com base na produção diária de determinada peça, durante cinco dias:

Empregado A: 70, 71, 69, 70, 70  $\rightarrow \bar{x} = 70$

Empregado B: 60, 80, 70, 62, 83  $\rightarrow \bar{x} = 71$



A performance média do empregado A é de 70 peças produzidas diariamente, enquanto que a do empregado B é de 71 peças. Com base na média aritmética, verifica-se que a performance de B é melhor do que a de A. Porém, observando bem os dados, percebe-se que a produção de A varia apenas de 69 a 71 peças, ao passo que a de B varia de 60 a 83 peças, o que revela que a performance de A é bem mais uniforme do que de B.

## Tipos de medidas de dispersão absoluta

Amplitude total ( $A_t$ ): é a diferença entre o maior e o menor valor observado.

$$A_t = x_{\max} - x_{\min}$$

Exemplo:

Pela situação sugerida na introdução, temos para a amplitude total os seguintes cálculos para os empregados:

$$\begin{aligned}\text{Empregado A} &\longrightarrow A_t = 71 - 69 = 2 \\ \text{Empregado B} &\longrightarrow A_t = 83 - 60 = 23\end{aligned}$$

Observações:

- **A amplitude total** é a medida mais simples de dispersão.
- A desvantagem desta medida de dispersão é que leva em conta apenas os valores mínimo e máximo do conjunto. Se ocorrer qualquer variação no interior do conjunto de dados, a amplitude total não nos dá qualquer indicação dessa mudança.
- A amplitude total também sofre a influência de um valor "atípico" na distribuição (um valor muito elevado ou muito baixo em relação ao conjunto).

## Variância( $\sigma^2$ ou $S^2$ ) e Desvio Padrão( $\sigma$ ou $S$ )

São as medidas de dispersão mais empregadas, pois levam em consideração a totalidade dos valores da variável em estudo. É um indicador de variabilidade bastante estável. Para medir a dispersão dos dados em torno da média, os estatísticos usam a soma dos quadrados dos desvios dividida pelo tamanho da população ou da amostra. Definindo assim, variância como média aritmética dos quadrados dos desvios.

	Dados não agrupados	Dados agrupados
População	$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{N}$	$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{N}$
Amostra	$S^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$	$S^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{n - 1}$

Sendo a variância calculada a partir dos quadrados dos desvios, ela é um número em unidade quadrada em relação à variável em questão, o que, sob o ponto de vista prático, é um inconveniente.

Por isso mesmo, imaginou-se uma nova medida que tem utilidade e interpretação práticas, denominada desvio padrão, definido como a raiz quadrada da variância.

$$= \sqrt{s^2} \quad (\text{População})$$

$$s = \sqrt{s^2} \quad (\text{Amostra})$$

Observações:

- Se os valores dos dados se repetirem em todas as amostras, então a variância da amostra será zero.
- Se os dados estiverem muito espalhados, então a variância da amostra acusará um número positivo elevado. Assim, uma grande variância significará uma grande dispersão dos dados em relação à média.
- A variância é uma medida que tem pouca utilidade na estatística descritiva, porém é extremamente importante na inferência estatística e em combinações de amostras.
- Quanto menor o desvio padrão, mais os valores da variável se aproximam de sua média.
- Quanto maior o desvio padrão, mais significativo à heterogeneidade entre os elementos de um conjunto

Exemplo:

Pela situação sugerida na introdução, temos a variâncias e o desvio padrão apresentando os seguintes cálculos para os empregados:

- Como a média do empregado A é 70, a variância será:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{N} = \frac{(70-70)^2 + (71-70)^2 + (69-70)^2 + (70-70)^2 + (70-70)^2}{5} = \frac{2}{5} = 0,4$$

E o desvio padrão será:  $s = \sqrt{s^2} = \sqrt{0,4} \cong 0,64$

Como a média do empregado B é 71 a variância será:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{N} = \frac{(60-71)^2 + (80-71)^2 + (70-71)^2 + (62-71)^2 + (83-71)^2}{5} = \frac{428}{5} = 85,6$$

E o desvio padrão será:  $s = \sqrt{s^2} = \sqrt{85,6} \cong 9,25$



Vamos colocar em prática, juntos, todos esses conceitos desta aula através do seguinte exercício:

1) Calcule a variância e o desvio padrão da tabela seguinte.

Salários semanais (R\$)	$f_i$	$x_i = \frac{l_i + l_s}{2}$	$x_i \cdot f_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2 \cdot f_i$
150  — 154	4				
154  — 158	9				
158  — 162	11				
162  — 166	8				
166  — 170	5				
170  — 174	3				
Total $\Sigma$	40				

Fonte: pesquisa do setor de pessoal

Calcular a média dos salários:  $\bar{x} = \frac{\sum x_i f_i}{n} =$

Calcular a variância dos salários:  $S^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{n - 1} =$




Calcular o desvio padrão:  $S = \sqrt{S^2} =$

## Construção de tabelas e cálculos estatísticos usando o excel


Partindo da seguinte série de dados, vamos construir: uma distribuição de freqüência sem intervalo de classe e com intervalo de classe, um histograma e cálculos de estatística descritiva.

14	12	13	11	12	13	16	14	14	15	17	14	11
13	14	15	13	12	14	13	14	13	15	16	12	12

## Entrada de Dados

- Para entrar no Microsoft Excel, clicar no botão .
- Entrar com os dados na coluna A.
- Na célula C1, digitar variáveis.
- Na célula D1, digitar fi.
- Na célula E1, digitar fri.
- Na célula F1, digitar Faci.
- Na célula G1, digitar Fraci.- Marcar as células C1:G1, na barra de formatos, clicar no botão Negrito **N** e no botão Centralizar .
- Nas células C2:C8, digitar as variáveis correspondentes a cada classe para uma distribuição de frequência sem intervalo de classe. (11, 12, 13, 14, 15, 16, 17)
- Marcar as células C2:C9, na barra de formatos, clicar no botão Negrito **N** e no botão Centralizar .

### Frequência Simples



- Marcar as células D2:D8, na barra de ferramentas, clicar no botão Colar Função .
- Na janela da esquerda, clicar em Estatística e na janela da direita, procurar Frequência, clicar em Ok.
- Na Matriz\_dados, digitar A1:A26.
- Na Matriz\_bin, digitar C2:C8.
- Como os dados foram preparados para obter os dados da função frequência como matriz coluna, no lugar de pressionar Ok ou a tecla Enter, devemos pressionar ao mesmo tempo as três teclas Ctrl + Shift + Enter.

Se não funcionar, começamos por:

Marcar as células D2:D8

Digitar a fórmula = FREQUÊNCIA(A1:A26;C2:C8)

Sem pressionar a tecla Enter, devemos pressionar ao mesmo tempo as três teclas Ctrl+ Shift + Enter.

- Na célula C9, digitar total.
- Clicar na célula D9 e na barra de ferramentas, clicar no botão Auto Soma . (Vai aparecer os valores acima marcados e a fórmula.) Pressionar a tecla Enter.
- Marcar D2:G9 e clicar no botão Centralizar .

### Frequência Simples relativa

- Na célula E2, digitar a fórmula = D2/\$D\$9 e pressionar a tecla Enter.
- Retornar para a célula E2 e no canto inferior direito, pressionar o botão

esquerdo do mouse e arrastar até a célula E8.

- Na barra de ferramentas, clicar no botão Diminuir Casas Decimais, até 4 casas decimais.
- Na célula E9, clicar no botão Auto Soma  $\Sigma$ .

#### Frequência Acumulada

- Na célula F2, digitar a fórmula = FREQÜÊNCIA (\$A\$1:\$A\$26,C2) e pressionar a tecla Enter.
- Retornar para a célula F2 e no canto inferior direito, pressionar o botão esquerdo do mouse e arrastar até a célula F8.

#### Frequência Acumulada relativa

- Na célula G2, digitar a fórmula = F2/\$D\$9 e pressionar a tecla Enter.
- Retornar para a célula G2 e no canto inferior direito, pressionar o botão esquerdo do mouse e arrastar até a célula G8.
- Na barra de ferramentas, clicar no botão Diminuir Casas Decimais, até 4 casas decimais.

	A	B	C	D	E	F	G
1	11		<b>Variáveis</b>	<b>fi</b>	<b>fri</b>	<b>Fi</b>	<b>Fri</b>
2	11		<b>11</b>	2	0.0769	2	0.0769
3	12		<b>12</b>	5	0.1923	7	0.2692
4	12		<b>13</b>	6	0.2308	13	0.5000
5	12		<b>14</b>	7	0.2692	20	0.7692
6	12		<b>15</b>	3	0.1154	23	0.8846
7	12		<b>16</b>	2	0.0769	25	0.9615
8	13		<b>17</b>	1	0.0385	26	1
9	13		<b>Total</b>	<b>26</b>	<b>1</b>		
10	13						
11	13		<i>Bloco</i>	<i>Frequência</i>	<i>% cumulativo</i>		
12	13		11	2	7.69%		
13	13		12	5	26.92%		
14	14		13	6	50.00%		
15	14		14	7	76.92%		
16	14		15	3	88.46%		
17	14		16	2	96.15%		
18	14		17	1	100.00%		

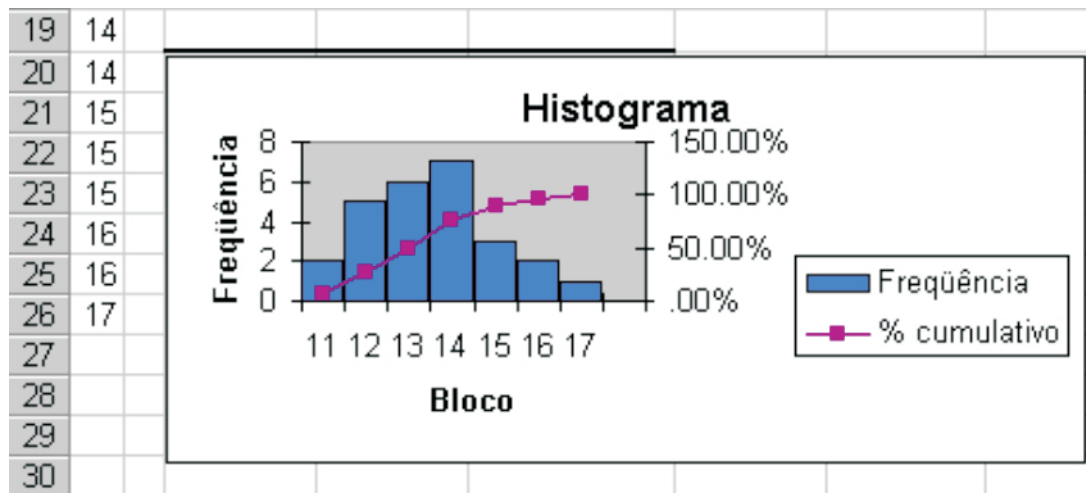
## Histograma e Frequência Simples

Na barra de títulos, abrimos o menu Ferramentas, escolher Análise de Dados, recebe-se a caixa de diálogo com todas as ferramentas de análise disponíveis. Nessa caixa de diálogo escolher Histograma e pressionar o botão Ok.

Observação:

Se ao abrirmos o menu Ferramentas, não estiver disponível Análise de Dados, devemos clicar em Suplementos. Ao abrir uma janela, devemos ativar Ferramentas de Análise, pressionar o botão Ok e repetir o passo anterior.

- No Intervalo de Entrada, digitar \$A\$1:\$A\$26.
- No Intervalo do Bloco, digitar \$C\$2:\$C\$8.
- No Intervalo de Saída, digitar \$C\$11. (O resultado gráfico será apresentado a partir desta célula.)
- Assinalar, através de um click, Porcentagem Cumulativa e Resultado Gráfico e pressionar o botão Ok.
- Marcar e deletar as células C19:E19.
- Dar um duplo click dentro de um dos retângulos do histograma.
- Dentro da caixa Formatar Ponto de Dados, selecionar Opções.
- A Largura do Espaçamento, reduzir para zero, e pressionar o botão Ok.



## Distribuição de Frequência com intervalos de classe

- Selecionar uma nova planilha, copiando para essa os dados da coluna A.
- Na célula C1, digitar Classes.
- Na célula D1, digitar Li.
- Na célula E1, digitar fi.
- Marcar as células C1:E1, centralizar e colocar em negrito.
- Marcar as células C2:E5 e centralizar.
- Digitar as classes na coluna C e os limites superiores na coluna D.
- Marcar as células E2:E5, na barra de ferramentas, clicar no botão Colar Função  $f_x$ . Na janela da esquerda, clicar em Estatística e na janela da direita, procurar Frequência, clicar em Ok.
- Na Matriz\_dados, digitar A1:A26.
- Na Matriz\_bin, digitar D2:D5.
- Como os dados foram preparados para obter os dados da função frequência como matriz coluna, no lugar de pressionar Ok ou a tecla Enter, devemos pressionar ao mesmo tempo as três teclas Ctrl + Shift + Enter.

## Análise dos Dados

Na barra de títulos, abrimos o menu Ferramentas, escolher Análise de Dados, recebe-se a caixa de diálogo com todas as ferramentas de análise disponíveis. Nessa caixa de diálogo escolher Estatística Descritiva e pressionar o botão Ok.

- No Intervalo de Entrada, digitar \$A\$1:\$A\$26.
- No Intervalo de Saída, digitar \$B\$7.
- Assinalar em Resumo Estatístico, Enésimo Maior e Enésimo Menor, pressionar o botão Ok.

Abaixo, encontram-se os resultados das atividades descritas acima, confira:

	A	B	C	D	E
1	11		<b>Classes</b>	<b>Li</b>	<b>fi</b>
2	11		10 ---  12	12	7
3	12		12 ---  14	14	13
4	12		14 ---  16	16	5
5	12		16 ---  18	18	1
6	12				
7	12	Coluna1			
8	13				
9	13	Média	13.53846		
10	13				
11	13	Mediana	13.5		
12	13	Modo	14		
13	13	Desvio padrão	1.529203		
14	14	Variância da amostra	2.338462		

## EM SÍNTESE

Amplitude total é a diferença entre o maior e o menor valor observado.

$$A_t = x_{\max} - x_{\min}$$

Variância de dados não agrupados em classe:

$$\text{População: } \sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

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

Variância de dados agrupados em classe:

$$\text{População: } \sigma^2 = \frac{\sum (x_i - \mu)^2 \cdot f_i}{N}$$

$$\text{Amostra: } s^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{n - 1}$$

Desvio padrão:

$$\text{População: } \sigma = \sqrt{s^2}$$

$$\text{Amostra: } s = \sqrt{s^2}$$

## MEDIDAS DE DISPERSÃO (MEDIDAS DE VARIABILIDADE) – COEFICIENTE DE VARIAÇÃO DE PEARSON

### Coeficiente de variação de Pearson (medida de dispersão relativa)

Quando se deseja comparar a variabilidade de duas ou mais distribuições, mesmo quando essas se referem a diferentes fenômenos e sejam expressas em unidades de medida distintas, podemos utilizar o Coeficiente de Variação de Pearson.

O coeficiente de variação para um conjunto de  $n$  observações é definido como o quociente entre o desvio padrão e a média aritmética da distribuição.

$$\text{População: } cv = \frac{\sigma}{\mu} \times 100$$

$$\text{Amostra: } cv = \frac{s}{\bar{x}} \times 100$$

Observação:

CV = coeficiente de variação de Pearson ou apenas coeficiente de variação, geralmente é expresso em porcentagem. Alguns analistas consideram:

- Baixa dispersão:  $cv \leq 15\%$
- Média dispersão:  $15\% < cv < 30\%$
- Alta dispersão:  $cv \geq 30\%$

Um coeficiente de variação maior ou igual a 30% revela que a série é heterogênea e a média tem pouco significado. Se o coeficiente de variação for menor que 30%, portanto a série é homogênea e a média tem grande significado.

Exemplo:

Considerando uma amostra que apresenta uma distribuição de frequências com  $\bar{x} = 161$  e  $s = 4,22$ , logo temos coeficiente de variação valendo:

$$cv = \frac{s}{\bar{x}} \times 100 = \frac{4,2}{161} \times 100 = 2,62111 \cong 2,6$$

### Em síntese

coeficiente de variação de Pearson:

$$\text{População: } cv = \frac{\sigma}{\mu} \times 100$$

$$\text{Amostra: } cv = \frac{s}{\bar{x}} \times 100$$

Pelos resultados do coeficiente de variação de Pearson podemos classificar:

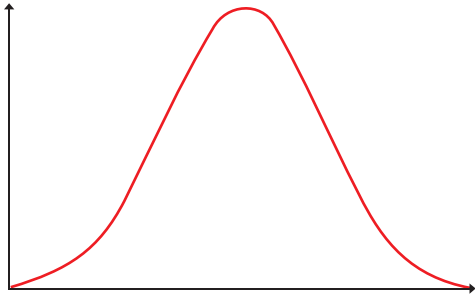
- baixa dispersão:  $cv \leq 15\%$



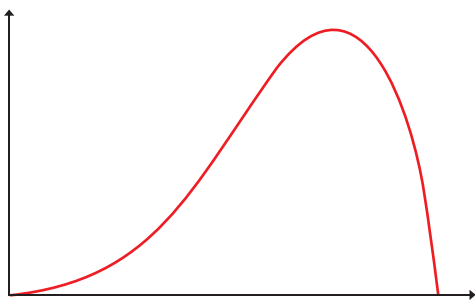
- média dispersão:  $15\% < cv < 30\%$
- alta dispersão:  $cv \geq 30\%$

## MEDIDAS DE ASSIMETRIA

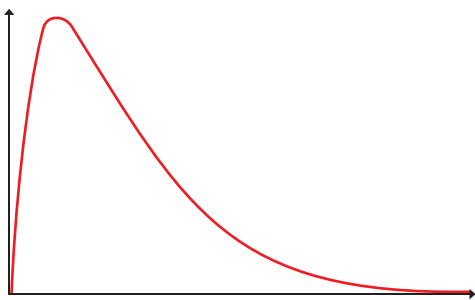
Uma distribuição de frequências com classes é **simétrica** quando:  
Média = Mediana = Moda



Uma distribuição de frequências com classes é:  
**Assimétrica à esquerda ou negativa** quando: Média < Mediana < moda



**Assimétrica à direita ou positiva** quando: Média > Mediana > Moda



Para fazer a classificação do tipo de assimetria através de cálculo, temos a seguinte relação:

$$\bar{x} = Mo$$

- Se:
- $\bar{x} - Mo = 0 \longrightarrow$  assimetria nula ou distribuição simétrica.
  - $\bar{x} - Mo < 0 \longrightarrow$  assimetria negativa ou à esquerda.
  - $\bar{x} - Mo > 0 \longrightarrow$  assimetria positiva ou à direita.

Exemplo:

Determinando os tipos de assimetria das distribuições abaixo:

Distribuição A	
Classes	fi
2  — 6	6
6  — 10	12
10  — 14	24
14  — 18	12
18  — 22	6
Total $\Sigma$	60

Distribuição B	
Classes	fi
2  — 6	6
6  — 10	12
10  — 14	24
14  — 18	30
18  — 22	6
Total $\Sigma$	78

Distribuição C	
Classes	fi
2  — 6	6
6  — 10	30
10  — 14	24
14  — 18	12
18  — 22	6
Total $\Sigma$	78

Calculando a média, mediana, moda e o desvio padrão para cada distribuição, temos:

Distribuição A  $\rightarrow \bar{x} = 12$ , Md = 12, Mo = 12 kg e s = 4,42

Distribuição B  $\rightarrow \bar{x} = 12,9$ , Md = 13,5, Mo = 16 kg e s = 4,20

Distribuição C  $\rightarrow \bar{x} = 11,1$ , Md = 10,5, Mo = 8 kg e s = 4,20

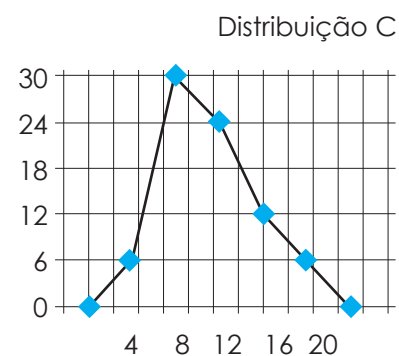
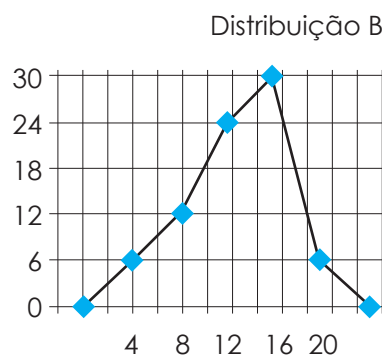
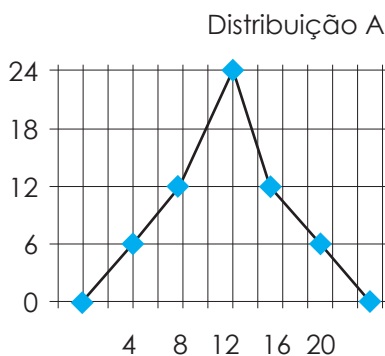
Logo, calculando o tipo de assimetria:

Distribuição A  $\rightarrow \bar{x} - Mo = 12 - 12 = 0 \rightarrow$  a distribuição é simétrica.

Distribuição B  $\rightarrow \bar{x} - Mo = 12,9 - 16 = -3,1 \rightarrow$  a distribuição é assimétrica negativa.

Distribuição C  $\rightarrow \bar{x} - Mo = 11,1 - 8 = 3,1 \rightarrow$  a distribuição é assimétrica positiva.

Construindo os gráficos das distribuições anteriores, temos:



## Coeficiente de assimetria (coeficiente de Pearson)

A medida anterior, por ser absoluta, apresenta a mesma deficiência do desvio padrão, isto é, não permite a possibilidade de comparação entre as medidas de

duas distribuições. Por esse motivo, daremos preferência ao coeficiente de assimetria de Pearson:

$$AS = \frac{3(\bar{x} - Md)}{s}$$

Escalas de assimetria:

$0,15 < | AS | < 1 \longrightarrow$  assimetria moderada

$| AS | \geq 1 \longrightarrow$  assimetria elevada

Observação:

Coeficiente de assimetria de Pearson (AS) é considerado em módulo para se fazer a classificação acima.

Exemplo:

Considerando as distribuições anteriores A , B e C , calcule os coeficientes de assimetria e faça a análise dos resultados obtidos:

$$AS_A = \frac{3(\bar{x} - Md)}{s} = \frac{3(12 - 12)}{4,42} = 0 \longrightarrow \text{Simetria}$$

$$AS_B = \frac{3(\bar{x} - Md)}{s} = \frac{3(12,9 - 13,5)}{4,20} = -0,429 \longrightarrow \text{Assimetria é considerada moderada e negativa}$$

$$AS_C = \frac{3(\bar{x} - Md)}{s} = \frac{3(11,1 - 10,5)}{4,20} = 0,429 \longrightarrow \text{Assimetria é considerada moderada e positiva}$$

### Em síntese

Uma distribuição de frequências com classes é simétrica quando:

$$\text{Média} = \text{Mediana} = \text{Moda}$$

Uma distribuição de frequências com classes é assimétrica à esquerda ou negativa quando:

$$\text{Média} < \text{Mediana} < \text{Moda}$$

Uma distribuição de frequências com classes é assimétrica à direita ou positiva quando:

$$\text{Média} > \text{Mediana} > \text{Moda} \quad \text{Coeficiente de assimetria: } AS = \frac{3(\bar{x} - Md)}{s}$$

Pelos resultados do coeficiente de assimetria podemos classificar:

Simétrico:  $| AS | = 0$

Assimetria moderada:  $0,15 < | AS | < 1$

Assimetria elevada:  $| AS | \geq 1$

Responda às questões abaixo e confira as respostas, no gabarito, no final do livro

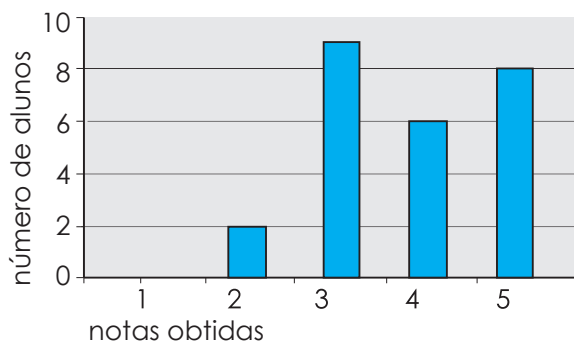
26) A tabela a seguir representa a distribuição de frequências dos salários de um grupo de 50 empregados de uma empresa, num certo mês.

Número de classe	Salário do mês em R\$	Número de empregados
1	1000 — 2000	20
2	2000 — 3000	18
3	3000 — 4000	9
4	4000 — 5000	3

o salário médio desses empregados, nesse mês, foi de:

- a) R\$ 2 637,00
- b) R\$ 2 520,00
- c) R\$ 2 500,00
- d) R\$ 2 420,00
- e) R\$ 2 400,00

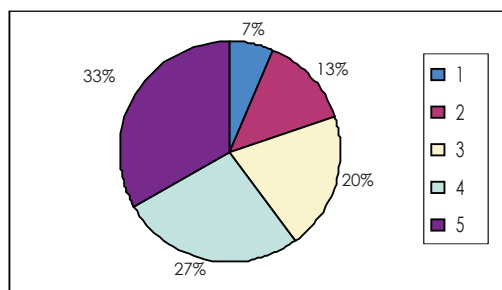
27) O professor Javert aplicou uma prova de matemática a 25 alunos, contendo 5 questões, valendo 1 ponto cada uma. Após fazer a correção, o professor construiu o gráfico abaixo, que relaciona o número de alunos às notas obtidas por eles.



Observando o gráfico, conclui-se que a mediana das notas obtidas pelos 25 alunos correspondem, a:

- a) 1
- b) 2
- c) 3
- d) 4
- e) 5

28) O gráfico de setores abaixo mostra os resultados obtidos em uma pesquisa com clientes de um supermercado sobre um de seus produtos.



De acordo com a satisfação, cada cliente indicava uma nota de 1 a 5 para o mesmo produto. Esse gráfico, por exemplo, mostra que 7% dos consumidores deram nota 1 para o produto consultado. considerando o mesmo gráfico, é correto afirmar que a moda do conjunto de todas as notas dadas para o produto foi:

- a) 1
- b) 2
- c) 3
- d) 4
- e) 5

**29) As notas de um candidato nas provas de um concurso foram: 8,4; 9,1; 7,2; 6,8; 8,7 e 7,2. A nota média, a nota mediana e a nota modal desse aluno são respectivamente:**

- a) 7,9; 7,8; 7,2
- b) 7,1; 7,8; 7,9
- c) 7,8; 7,8; 7,9
- d) 7,2; 7,8; 7,9
- e) 7,8; 7,9; 7,2

**30) Analise os dados abaixo que se referem a uma amostra de óbitos/mês nos hospitais de uma grande capital brasileira, nos últimos 25 meses e faça o que se pede:**

410, 440, 440, 584, 579, 577, 577, 454, 460, 460, 491, 561, 556, 491, 490, 491, 490, 553, 553, 540, 540, 520, 491, 520, 491.

Calcular a média aritmética, mediana e a moda.

Considere as seguintes notas de uma amostra de uma turma e faça o que se pede:

35, 45, 40, 40, 52, 56, 59, 60, 54, 55, 56, 59, 60, 54, 56, 57, 59, 60, 55, 56, 57, 59, 60, 55, 56, 58, 60, 60, 70, 75, 80, 85, 60, 65, 63.

**31) calcular o terceiro quartil da amostra.**

**32) calcular o segundo decil da amostra.**

**33) calcular o nono decil da amostra.**

**34) calcular o vigésimo percentil da amostra.**

**35) calcular o nonagésimo percentil da amostra.**

**36) Dado o conjunto de números 8, 4, 6, 9, 10, 5, determine a variância desse conjunto, supondo que esses valores correspondem a uma amostra.**

- a) 28
- b) 2,3664
- c) 7
- d) 2,8
- e) 5,6

**37) Determine o desvio padrão do conjunto de números da questão anterior, supondo que esses valores correspondem a uma amostra.**

- a) 28
- b) 2,3664
- c) 7
- d) 2,8
- e) 5,6

Considere as seguintes notas de uma amostra de uma turma e responda as perguntas a seguir:

35, 45, 40, 40, 52, 56, 59, 60, 54, 55, 56, 59, 60, 54, 56, 57, 59, 60, 55, 56, 57, 59, 60, 55, 56, 58, 60, 60, 70, 75, 80, 85, 60, 65, 63.

**38) calcule a amplitude total do rol.**

- a) 50
- b) 49
- c) 48
- d) 47
- e) 45

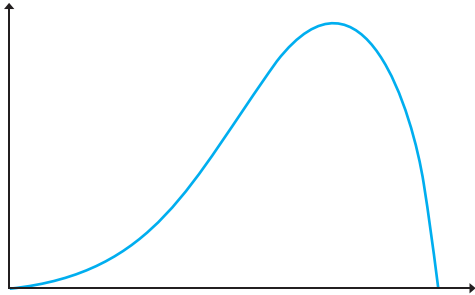
**39) calcule a variância do rol.**

- a) 94,56
- b) 105,83
- c) 110,83
- d) 109,83
- e) 120,00

**40) calcule o desvio padrão do rol.**

- a) 10,29
- b) 9,72
- c) 10,53
- d) 10,48
- e) 10,99

41) Dada a figura a seguir, podemos afirmar que:



- a) a moda é maior do que a mediana e menor do que a média;
- b) a moda é menor do que a mediana e maior do que a média;
- c) a moda é menor do que a mediana e está maior do que a média;
- d) a mediana é maior do que a média e menor do que a moda;
- e) n.d.a.

42) os coeficientes de variação dos resultados a seguir são:

Estatística:  $\bar{x} = 80$  e  $s = 16$

História:  $\bar{x} = 20$  e  $s = 5$

- a) 16% e 40%
- b) 20% e 25%
- c) 50% e 40%
- d) 80% e 40%
- e) n.d.a.

43) Considere os seguintes resultados relativos a três distribuições de frequência:

Distribuição	$\bar{x}$	Mo
A	52	52
B	45	50
C	48	46

Determine o tipo de assimetria de cada uma delas.

**44) uma distribuição de freqüência apresenta as seguintes medidas:**

$$\bar{x} = 48,1, \text{ md} = 47,9 \text{ e } s = 2,12.$$

Calcule o coeficiente de assimetria.

**45) Em uma distribuição de freqüência foram encontradas as seguintes medidas:**

$$\bar{x} = 33,18, \text{ mo} = 27,50, \text{ md} = 31,67 \text{ e } s = 12,45.$$

a) classifique o tipo de assimetria.

b) Calcule o coeficiente de assimetria.



# UNIDADE

PROBABILIDADE

5



Já sabemos que para se obter informações sobre alguma característica da população, o tamanho amostral é de fundamental importância. Estudaremos agora a probabilidade, que é uma ferramenta usada e necessária para se fazerem ligações entre a amostra e a população, de modo que a partir de informações da amostra se possam fazer afirmações sobre características da população.

Assim, pode-se dizer que a probabilidade é a ferramenta básica da Estatística Inferencial.

## Experimento Aleatório

São aqueles que não podem ser previamente determinados. Essa impossibilidade de prever-se os resultados, chamamos de acaso .

Exemplo: Lançar um dado e anotar o número que ocorrerá na face voltada para cima.

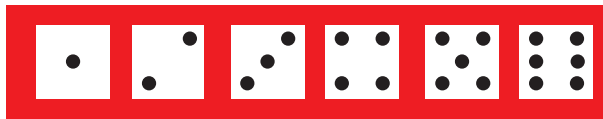


## Espaço Amostral (S)

É o conjunto formado por todos os resultados possíveis de um experimento aleatório.

Exemplo 1: Ao se lançar um dado e observar a face superior, têm-se o espaço amostral:

$$S = \{1, 2, 3, 4, 5, 6\}$$



Exemplo 2: Numa partida de futebol, uma das equipes pode obter resultados tais como: vitória (v), empate (e) ou derrota (d). Tem-se então:

$$S = \{v, e, d\}$$



## Evento (E)

É um conjunto qualquer de resultados de um experimento aleatório. Pode-se dizer que um evento é um subconjunto do espaço amostral.

Exemplo: no lançamento de 2 moedas apareçam faces iguais. Os elementos do evento são  $E = \{(K, K), (C, C)\}$ .



## tipos de eventos

- Evento certo – é o próprio espaço amostral.

Exemplo: Lançamento de um dado e ocorrência de um número menor ou igual a 6 na face superior.

- Evento impossível – é o subconjunto vazio do espaço amostral.

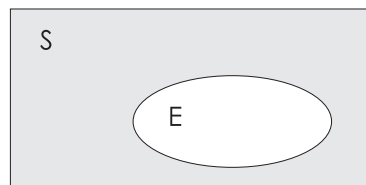
Exemplo: Lançamento de um dado e ocorrência de um número maior do que 6 na face superior.

- Eventos elementares – são aqueles que têm um só elemento.

Exemplo: Lançamento de um dado e ocorrência de um número ímpar maior do que 4 na face superior.

Cálculo da probabilidade de um evento ocorrer

Podemos definir o cálculo da probabilidade de um evento como a razão (divisão) entre o número de elementos do evento e o número de elementos do espaço amostral.



$$P(E) = \frac{n(E)}{n(S)}$$

Onde:  $n(E)$  = o número de elementos do evento  
 $n(S)$  = o número de elementos do espaço amostral  
 $P(E)$  = a probabilidade de ocorrer o evento

Na prática, calcular a probabilidade é dividir:

$$P(E) = \frac{\text{número de resultados favoráveis}}{\text{número de resultados possíveis}}$$

Observação: Percentualmente, a probabilidade varia de 0% a 100%, ou seja,  $0\% \leq P(E) \leq 100\%$  ou  $0 \leq P(E) \leq 1$ .

## Eventos complementares $P(\bar{A})$

A probabilidade de não ocorrer o evento A é igual a 1 menos a probabilidade de ocorrer A, que pode ser representada por:

$$P(\bar{A}) = 1 - P(A)$$

Exemplo:

Analizando um lote de 360 peças para computador, o departamento de controle

de qualidade de uma fábrica constatou que 40 peças estavam com defeito. Retirando-se uma das 360 peças, ao acaso, a probabilidade de esta peça NÃO ser defeituosa é:

- **Sendo S** = conjunto dos elementos do espaço amostral, casos possíveis, e  $n(s)$  o número de elementos deste conjunto.
- **Sendo E** = conjunto de elementos das peças defeituosas, e  $n(E)$  o número de elementos deste conjunto.
- **Use  $\sim E$**  = conjunto dos elementos das peças não defeituosas, e  $n(\sim E)$  o número de elementos deste conjunto. Neste caso, é o conjunto dos casos favoráveis.
- $n(S) = 360$  ,  $n(E) = 40$  e  $n(\sim E) = 320$

Para calcular a probabilidade de, retirada uma peça que seja não defeituosa, faça assim:

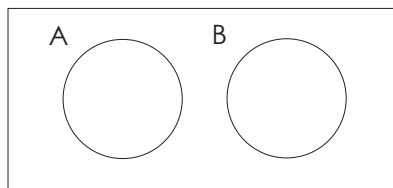
$$P(\sim E) = \frac{n(E)}{n(S)} = \frac{320}{360} = \frac{8}{9}$$

### Probabilidade da união $P(A \cup B) = P(A \text{ ou } B)$

Nesse caso, existem dois tipos possíveis de situação

- Eventos mutuamente exclusivos:

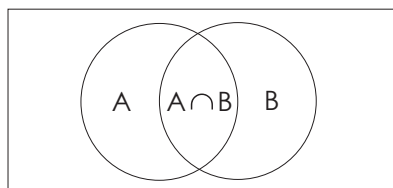
Dois eventos são mutuamente exclusivos se  $A \cap B = \emptyset$ , neste caso:



$$P(A \cup B) = P(A) + P(B)$$

- Eventos não mutuamente exclusivos:

Dois eventos não são mutuamente exclusivos, se  $A \cap B \neq \emptyset$ , neste caso:



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Exemplo:

No lançamento de um dado, determine a probabilidade de se obter um número par ou maior que 3.

- Espaço amostral:  $S = \{1, 2, 3, 4, 5, 6\} \longrightarrow n(S) = 6$

- Evento A (números pares):  $A = \{2, 4, 6\} \longrightarrow n(A) = 3$
- Evento B (números maiores que 3):  $B = \{4, 5, 6\} \longrightarrow n(B) = 3$
- Evento de A e B:  $A \cap B = \{4, 6\} \longrightarrow n(A \cap B) = 2$

Calculando a probabilidade, temos:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \longrightarrow P(A \cup B) = \frac{3}{6} + \frac{3}{6} - \frac{2}{6} = \frac{4}{6} = \frac{2}{3}$$

## Probabilidade condicional

Sejam dois eventos A e B associados a um espaço amostral S. A probabilidade de A ocorrer dado que o evento B ocorreu é definida por:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad \text{onde } P(B) \neq 0$$

Portanto, quando calculamos  $P(A/B)$ , tudo se passa como se o evento B fosse um novo espaço amostral reduzido dentro do qual queremos calcular a probabilidade do evento.

Exemplo:

Considere o conjunto de números inteiros  $\{1, 2, 3, 4, 5, \dots, 18, 19, 20\}$ , e, por meio de um sorteio aleatório, retire um número. Se o número sorteado for ímpar, qual a probabilidade de o número sorteado ser o 13?

- Espaço amostral  $S = \{1, 2, 3, \dots, 19, 20\} \longrightarrow n(S) = 20$
- Evento  $A = \{13\} \longrightarrow n(A) = 1$
- Evento B: Condição para ocorrência do evento A  $= \{1, 3, 5, 7, 9, 11, 13, 15, 17, 19\} \longrightarrow n(B) = 10$
- $(A \cap B) = \{13\} \longrightarrow n(A \cap B) = 1$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{10}$$

## Eventos independentes

Diz-se que dois ou mais eventos são independentes, quando a ocorrência de um não depende (ou não é condicionada, ou não se vincula) da ocorrência do outro, isto é, a informação adicional de que um dos eventos já ocorreu em nada altera a probabilidade de ocorrência do outro.

Dados dois eventos independentes A e B, a probabilidade de que ocorram os eventos A e B é dado por:

$$P(A \cap B) = P(A) \cdot P(B)$$

Exemplo:

Num grupo de jovens estudantes a probabilidade de que um jovem, escolhido ao acaso, tenha média acima de 7,0 é  $\frac{1}{5}$

- Nesse mesmo grupo, probabilidade de que um jovem saiba jogar futebol é  $\frac{5}{6}$
- Qual a probabilidade de escolhermos um jovem (ao acaso) que tenha média maior que 7,0 e saiba jogar futebol?

O fato de ter média maior que 7,0, não depende do fato de saber jogar futebol, e vice-versa. Quando isso ocorre, dizemos que os eventos são independentes.

A: ter média acima de 7,0.

B: saber jogar futebol.

A e B: ter média acima de 7,0 e saber jogar futebol.

Num grupo de jovens estudantes a probabilidade de que um jovem, escolhido ao acaso, tenha média acima de 7,0 é  $\frac{1}{5}$

Num grupo de jovens estudantes a probabilidade de que um jovem, escolhido ao acaso, tenha média acima de 7,0 é  $\frac{1}{5.6}$

## Em Síntese

- Experimentos aleatórios são aqueles quando repetidos várias vezes, em situações idênticas, apresentam resultados diferentes, não sendo possível a determinação do resultado antes de sua realização.
- Espaço amostral é o conjunto formado por todos os resultados possíveis de um experimento. Notação: S.
- Evento é qualquer subconjunto do espaço amostral. Notação: E.
- Cálculo da probabilidade de um evento:  $P(E) = \frac{n(E)}{n(S)}$

## Número de resultados favoráveis

número de resultados possíveis com  $0\% < P(E) < 100\%$  ou  $0 < P(E) < 1$ .

- Eventos complementares:  $P(\bar{A}) = 1 - P(A)$

- Probabilidade da união de eventos mutuamente exclusivos:

$$P(A \cup B) = P(A) + P(B)$$

- Probabilidade da união de eventos não mutuamente exclusivos:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Probabilidade condicional:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad \text{onde } P(B) \neq 0$$

- Probabilidade de eventos independentes:

$$P(A \cup B) = P(A) \cdot P(B)$$

**Responda às questões abaixo e confira as respostas, no gabarito, no final do livro**

- 46)** Dois jogadores A e B vão lançar um par de dados. Eles combinam que se a soma dos números dos dados for 5, A ganha e se a soma for 8, B é quem ganha. Os dados são lançados. Sabe-se que A não ganhou. Qual a probabilidade de B ter ganho?
- a)  $10/36$
  - b)  $5/32$
  - c)  $5/36$
  - d)  $5/35$
  - e) Não se pode calcular sem saber os números sorteados.
- 47) (UNIRIO)** As probabilidades de três jogadores marcarem um gol cobrando um pênalti são, respectivamente,  $1/2$ ,  $2/5$  e  $5/6$ . Se cada um bater um único pênalti, a probabilidade de todos errarem é igual a:
- a) 3%
  - b) 5%
  - c) 17%
  - d) 20%
  - e) 25%
- 48) (MACK)** Numa competição de tiro ao alvo, a probabilidade de um atirador A errar é 8% e a de um atirador B errar é o dobro. Ocorridos 200 tiros, 100 para cada atirador, e tendo havido erro num dos tiros, a probabilidade de o mesmo ter sido dado por A é:
- a)  $1/5$
  - b)  $1/3$
  - c)  $3/4$
  - d)  $1/2$
  - e)  $1/6$
- 49)** O resultado de uma pesquisa realizada pelo IPESPE sobre o perfil dos fumantes e publicada pela revista veja de 03/06/07 mostra que num grupo de 1000 pessoas 17% fumam e, dentre os fumantes, 44% são mulheres. Se nesse grupo de 1000 pessoas uma é escolhida ao acaso, a probabilidade de ela ser fumante e mulher é aproximadamente:
- a) 0,044
  - b) 0,075
  - c) 0,44
  - d) 0,0075
  - e) 0,0044



## RETOMANDO A CONVERSA INICIAL

Inicialmente nos propusemos a fundamentar as possibilidades de coletar, organizar, descrever e interpretar dados.

Apresentamos, na primeira unidade, conceitos que embasaram as atividades das unidades seqüentes.

Para cada unidade, adotamos construir passo a passo, com você, tabelas, gráficos, distribuições, e medidas e, ainda, projetar as probabilidades.

As atividades propostas tiveram por objetivo aplicar o conhecimento construído.

Espero ter despertado o gosto pela estatística e ter colaborado para seu sucesso profissional.



## REFERÊNCIAS BIBLIOGRÁFICAS

- COSTA NETO, P. L. de O. *Probabilidades*. São Paulo: Editora Edgard Blucher Ltda, 1985.
- COSTA NETO, P. L. de O. *Estatística*. São Paulo: Editora Edgard Blucher Ltda, 17ª ed. 1999.
- CRESPO, A. A. *Estatística Fácil*. São Paulo: Editora Saraiva, 17ª ed. 1999.
- DANTE, L. R. *Matemática: Contexto de Aplicações*. São Paulo: Editora Makron Books Ltda., 1982.
- DOWNING, D., CLARK, J. *Estatística Aplicada*. São Paulo: Editora Saraiva, 2000.
- KAZMIER, L. J. *Estatística Aplicada à Economia e Administração*. São Paulo: Editora Makron Books Ltda., 1982.
- LAPPONI, J. C. *Estatística Usando Excel*. São Paulo: Editora Lapponi, 2000.
- LEVIN, J. *Estatística Aplicada a Ciências Humanas*. 2ª edição. São Paulo: Editora Harper & Row do Brasil Ltda, 1978.
- NICK, E., KELLNER, S. R. O. *Fundamentos de Estatística para as Ciências do Comportamento*. Rio de Janeiro: Editora Renes, 1971.
- SIEGEL, S. *Estatística Não Paramétrica*. São Paulo: Editora McGraw-Hill do Brasil, 1975.
- STEVENSON, W. J. *Estatística Aplicada à Administração*. São Paulo: Editora Harper & Row do Brasil Ltda., 1981.
- TRIOLA, M. F. *Introdução à Estatística*. Rio de Janeiro: Livros Técnicos e Científico Editora S.A., 7ª ed., 1999.

## GABARITO

- |                |                   |                |                    |
|----------------|-------------------|----------------|--------------------|
| 1. C           | 16. A             | 31.            | 44. 0,283          |
| 2. B           | 17. C             | 32.            | 45. A) ASSIMÉTRICA |
| 3. D           | 18. D             | 33.            | POSITIVA,          |
| 4. E           | 19. D             | 34.            | B)0364             |
| 5. D           | 20. E             | 35.            | 46. B              |
| 6. A           | 21.               | 36. E          | 47. B              |
| 7. C           | 22.               | 37. B          | 48. B              |
| 8. D           | 23.               | 38. A          | 49. C              |
| 9. B           | 24.               | 39.            |                    |
| 10. CONSTRUÇÃO | 25.               | 40.            |                    |
| COM O USO      | 26. E             | 41. D          |                    |
| DO EXCEL       | 27. D             | 42. B          |                    |
| 11. E          | 28. E             | 43. SIMÉTRICA, |                    |
| 12. E          | 29. A             | ASSIMÉTRICA    |                    |
| 13. A          | 30. MEDIA 510,36, | NEGATIVA,      |                    |
| 14. C          | MEDIANA: 491,     | ASSIMÉTRICA    |                    |
| 15. D          | MODA 491          | POSITIVA       |                    |

Autora

**JAVERT GUIMARÃES FALCO** é tecnólogo em Gestão Financeira pela Faculdade OPET e licenciado em matemática pela UTFPr (Universidade Tecnológica Federal do Pr antigo \CEFET- Pr); cursa MBA em Gestão de organizações Educacionais pela Faculdade OPET.

Atualmente é professor convidado do Departamento de Educação a Distância do EtuFPr, onde ministra as disciplinas de matemática Financeira, matemática Aplicada e Estatística, lecionando também em outras instituições de Educação a Distância nas mesmas disciplinas.

Atua como professor de Pré-vestibulares dos estados do Paraná e Santa Catarina, ministrando aulas de matemática. Também é professor universitário do Grupo UNINtEr, lecionando Estatística e matemática Aplicada para cursos de tecnologia na FATEc internacional.

Trabalha com consultoria Pedagógica na Educação a Distância e Presencial, desenvolvendo trabalhos de seleção e formação de equipe docente, organização do trabalho pedagógico, desenvolvimento de planos estratégicos de gestão empresarial.