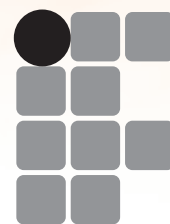




Estatística

Javert Guimarães Falco
Roberto José Medeiros Junior



INSTITUTO FEDERAL
PARANÁ
Educação à Distância

Curitiba-PR
2012

Presidência da República Federativa do Brasil

Ministério da Educação

Secretaria de Educação a Distância

2012 © INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA – PARANÁ
– EDUCAÇÃO A DISTÂNCIA

Este Caderno foi elaborado pelo Instituto Federal do Paraná para o Sistema Escola
Técnica Aberta do Brasil – e-Tec Brasil.

Prof. Irineu Mario Colombo
Reitor

Prof.^a Cristina Maria Ayroza
**Assessora de Ensino, Pesquisa e Extensão –
DEPE/EaD**

Prof.^a Mara Christina Vilas Boas
Chefe de Gabinete

Prof. Ezequiel Westphal
Pró-Reitoria de Ensino - PROENS

Prof. Gilmar José Ferreira dos Santos
Pró-Reitoria de Administração - PROAD

Prof. Silvestre Labiak
**Pró-Reitoria de Extensão, Pesquisa e Inovação
- PROEPI**

Neide Alves
**Pró-Reitoria de Gestão de Pessoas e Assuntos
Estudantis - PROGEPE**

Bruno Pereira Faraco
**Pró-Reitoria de Planejamento e
Desenvolvimento Institucional - PROPLAN**

Prof. José Carlos Ciccarino
Diretor Geral do Câmpus EaD

Prof. Ricardo Herrera
**Diretor de Planejamento e Administração do
Câmpus EaD**

Prof.^a Mércia Freire Rocha Cordeiro Machado
**Diretora de Ensino, Pesquisa e Extensão –
DEPE/EaD**

Prof.^a Cristina Maria Ayroza
**Assessora de Ensino, Pesquisa e Extensão –
DEPE/EaD**

Prof.^a Márcia Denise Gomes Machado Carlini
**Coordenadora de Ensino Médio e Técnico do
Câmpus EaD**

Prof. Roberto José Medeiros Junior
Coordenador do Curso

Prof.^a Ediane Santos Silva
Vice-coordenadora do Curso

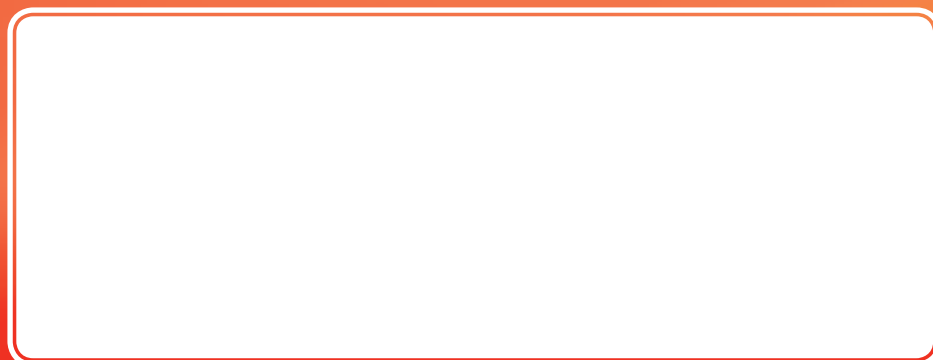
Adriana Valore de Sousa Bello
Cassiano Luiz Gonzaga da Silva
Jéssica Brisola Stori
Denise Glovaski Souto
Assistência Pedagógica

Prof.^a Ester dos Santos Oliveira
Prof.^a Sheila Cristina Mocellin
Idamara Lobo Dias
Prof.^a Maria Angela Motta
Revisão Editorial

Eduardo Artigas Antoniacomi
Flávia Terezinha Vianna da Silva
Diagramação

e-Tec/MEC
Projeto Gráfico

Catálogo na fonte pela Biblioteca do Instituto Federal do Paraná



Apresentação e-Tec Brasil

Prezado estudante,

Bem-vindo ao e-Tec Brasil!

Você faz parte de uma rede nacional pública de ensino, a Escola Técnica Aberta do Brasil, instituída pelo Decreto nº 6.301, de 12 de dezembro 2007, com o objetivo de democratizar o acesso ao ensino técnico público, na modalidade a distância. O programa é resultado de uma parceria entre o Ministério da Educação, por meio das Secretarias de Educação a Distância (SEED) e de Educação Profissional e Tecnológica (SETEC), as universidades e escolas técnicas estaduais e federais.

A educação a distância no nosso país, de dimensões continentais e grande diversidade regional e cultural, longe de distanciar, aproxima as pessoas ao garantir acesso à educação de qualidade, e promover o fortalecimento da formação de jovens moradores de regiões distantes, geograficamente ou economicamente, dos grandes centros.

O e-Tec Brasil leva os cursos técnicos a locais distantes das instituições de ensino e para a periferia das grandes cidades, incentivando os jovens a concluir o ensino médio. Os cursos são ofertados pelas instituições públicas de ensino e o atendimento ao estudante é realizado em escolas-polo integrantes das redes públicas municipais e estaduais.

O Ministério da Educação, as instituições públicas de ensino técnico, seus servidores técnicos e professores acreditam que uma educação profissional qualificada – integradora do ensino médio e educação técnica, – é capaz de promover o cidadão com capacidades para produzir, mas também com autonomia diante das diferentes dimensões da realidade: cultural, social, familiar, esportiva, política e ética.

Nós acreditamos em você!

Desejamos sucesso na sua formação profissional!

Ministério da Educação
Janeiro de 2010

Nosso contato
etecbrasil@mec.gov.br



Indicação de ícones

Os ícones são elementos gráficos utilizados para ampliar as formas de linguagem e facilitar a organização e a leitura hipertextual.



Atenção: indica pontos de maior relevância no texto.



Saiba mais: oferece novas informações que enriquecem o assunto ou “curiosidades” e notícias recentes relacionadas ao tema estudado.



Glossário: indica a definição de um termo, palavra ou expressão utilizada no texto.



Mídias integradas: sempre que se desejar que os estudantes desenvolvam atividades empregando diferentes mídias: vídeos, filmes, jornais, ambiente AVEA e outras.



Atividades de aprendizagem: apresenta atividades em diferentes níveis de aprendizagem para que o estudante possa realizá-las e conferir o seu domínio do tema estudado.



Sumário

Palavra dos professores-autores	9
Aula 1 - Estatística aplicada	11
1.1 Panorama histórico da estatística	12
1.2 O que é estatística?	13
1.3 Conceitos estatísticos	13
Aula 2 – Divisão da estatística	19
2.1 Estatística Descritiva	19
2.2 Estatística Indutiva ou Inferência	19
2.3 Fases do método estatístico	20
Aula 3 – Tabelas e gráficos	23
3.1 Tabelas	24
3.2 Séries estatísticas	24
AULA 4 – Gráficos estatísticos	27
4.1 Principais tipos de gráficos	27
Aula 5 – Distribuição de frequências	33
5.1 Representação dos dados (amostrais ou populacionais)	33
Aula 6 – Elementos principais da tabela de distribuição de frequências	159
6.1 Elementos de uma distribuição de frequência	159
Aula 7 - Distribuição de frequências	161
7.1 Determinação do número de classes (K)	161
Aula 8 - Tipos de frequências	165
8.1 Frequências simples ou absolutas (f_i)	165
8.2 Frequências relativas (fr_i)	165
8.3 Frequência simples acumulada (fac_i)	165
8.4 Frequência relativa acumulada ($frac_i$)	165
Aula 9 - Resolução de exercício prático	167
Aula 10 - Distribuição de frequências graficamente	169
10.1 Histograma e polígono de frequências	169

Aula 11 - Resolução de exercícios práticos	171
Aula 12 - Medidas de posição	173
12.1 Médias	173
Aula 13 - Medidas de posição - moda e mediana	179
13.1 Moda (mo) – para dados não agrupados	179
13.2 Mediana (Md)	180
Aula 14 - Comparação entre média, mediana e moda	183
Aula 15 - Medidas de dispersão (medidas de variabilidade)	185
15.1 Tipos de medidas de dispersão absoluta	186
Aula 16 - Variância (σ^2 ou s^2) e desvio padrão (σ ou s)	187
Aula 17 - Exercício prático sobre variância e desvio padrão	189
Aula 18 - Probabilidade	191
18.1 Experimento aleatório	191
18.2 Espaço Amostral (S) ou Universo	191
18.3 Evento (E)	192
Aula 19 - Cálculo da probabilidade de um evento ocorrer – princípio da contagem	195
19.1 Eventos complementares $P(A)$	195
19.2 Probabilidade da união $P(A \cup B) = P(A \text{ ou } B)$	196
Aula 20 - Probabilidade condicional	199
20.1 Eventos independentes	199
Referências	203
Atividades autoinstrutiva	205
Currículo dos professores-autores	227

Palavra dos professores-autores

Este material foi elaborado a partir de nossa 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.

Esperamos que através dos conteúdos contemplados neste material didático, 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!

Javert Guimarães Falco

Roberto José Medeiros Junior



Aula 1 - Estatística aplicada

Nesta primeira aula você irá compreender a importância da Estatística no nosso cotidiano.

O homem é um ser 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. Fique ligado quanto aos conceitos passados nessa aula.



Figura 1.1: Gráfico

Fonte: Acervo do ator

O relacionamento da Estatística com as demais ciências é cada vez mais intenso e mais importante. Veja-se, 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, etc; é básica para as Ciências Sociais, nas pesquisas socioeconômicas; é de aplicação intensa na Engenharia Industrial, no controle de qualidade, na comparação de fabricações, etc; é indispensável à Administração, à Programação, à Medicina, à Psicologia, à História, e, de forma direta ou indireta, às demais atividades.

No decorrer das aulas veremos exemplos práticos das aplicações citadas.

1.1 Panorama histórico da estatística



Figura 1.2: Censo Romano

Fonte: Acervo do autor

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.



Figura 1.3: Impostos Idade Média

Fonte: Acervo do autor

Atualmente, informações numéricas são necessárias para cidadãos e organizações de qualquer natureza, e de qualquer parte do mundo globalizado, para a tomada de decisões que envolvem a coletividade.



Figura 1.4: Multa e Semáforo

Fonte: Acervo do autor

1.2 O que é estatística?



Figura: 1.5: Gráfico

Fonte: Acervo do autor



Figura: 1.6: Estudos

Fonte: www.sxc.hu



Figura: 1.7: População

Fonte: www.seplan.go.gov.br

“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

1.3 Conceitos estatísticos

1.3.1 População

Conjunto de indivíduos, objetos ou informações que apresentam pelo menos uma característica comum, cujo comportamento interessa-nos analisar.

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.

Deseja-se saber se nas indústrias situadas no Estado do Paraná, em 2010, existia algum tipo de controle ambiental.

- [illegible]

Deseja-se conhecer o consumo total de energia elétrica em MWh nas residências da cidade de Curitiba no ano de 2010.

- ## Estatística



Figura: 1.9: Cidade iluminada

Fonte: Acervo do autor

1.3.2 Divisão da população

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

Exemplo: Idade dos alunos do curso de Administração na modalidade a Distância (EAD) no Estado do Paraná.

- b) **População ou Universo Estatístico:** Todos os alunos de Administração na modalidade a Distância (EAD) no Estado do Paraná.

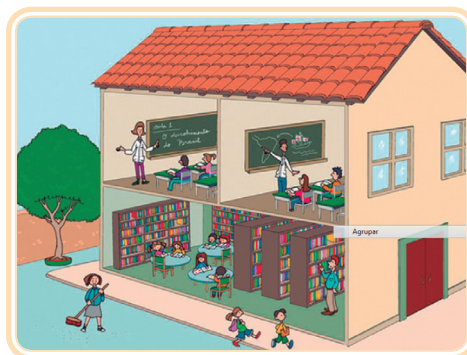


Figura 1.10: Colégio

Fonte: <http://2.bp.blogspot.com>

- c) **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.

Exemplos: Tipos de bactérias no corpo humano

d) População: Todas as bactérias existentes no corpo humano.



Figura 1.11: Bactérias

Fonte: Acervo do autor

Em geral, como os universos são grandes, investigar todos os elementos populacionais para determinarmos a característica necessita 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 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.

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 levantamento de dados; melhor investigação dos elementos observados.

A Estatística ocupa-se fundamentalmente das propriedades das populações cujas características são passíveis de representação numérica, tais 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**.



Para ter acesso a alguns dados do Brasil, basta acessar o site do Instituto Brasileiro de Geografia e Estatística (IBGE).
www.ibge.gov.br/agencias_pnuma.php

a) **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).

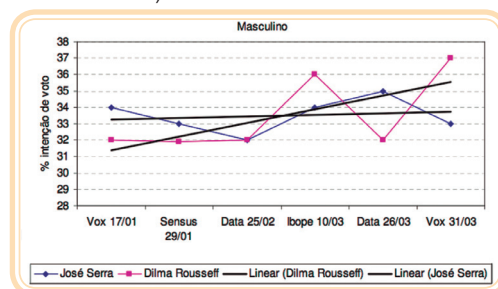


Figura 1.12: Gráfico

Fonte: <http://uolpolitica.blog.uol.com.br>

b) **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.)



Figura 1.13: Peso e altura

Fonte: Acervo do autor

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

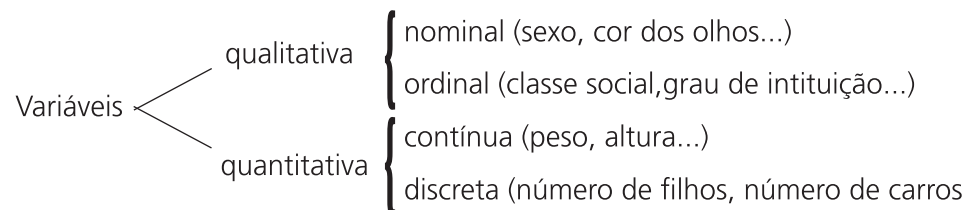


Figura 1.14: Conjunto de números reais

Fonte: <http://alfaconnection.net>

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

As variáveis podem ser resumidas da seguinte maneira:



Aula 2 – Divisão da estatística

Nesta aula conheceremos os dois ramos de estudo mais comuns em Estatística: a Descritiva e a Inferencial.

A Teoria Estatística moderna se divide em dois grandes campos:

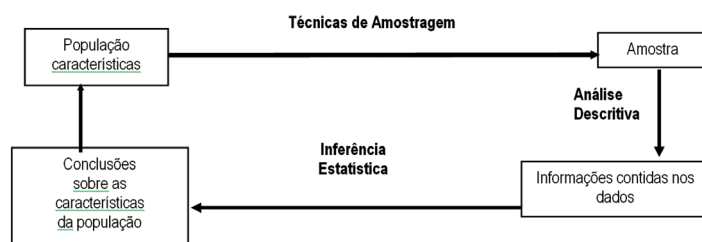
2.1 Estatística Descritiva

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 permita descrever o fenômeno.

2.2 Estatística Indutiva ou Inferência

É a 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 analisar 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:



Fluxograma 2.1: técnicas de amostragem

Fonte: Elaborado pelo autor

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

2.3 Fases do método estatístico

- 1. Fase – definição do problema:** Saber exatamente aquilo que se pretende pesquisar é o mesmo que definir corretamente o problema.
- 2. 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.
- 3. fase – coleta 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 qual 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 pode ser de dois tipos:

- I. 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.
- **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.

II. 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.

Observação:

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

- 4. Fase – apuração dos dados:** Resumo dos dados através de sua contagem e agrupamento. É a condensação e tabulação de dados.
- 5. 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.
- 6. 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.

Resumo

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 em 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 ou Inferência: se preocupa com as hipóteses e conclusões sobre a população.

Anotações

Aula 3 – Tabelas e gráficos

Na aula três serão apresentados os tipos mais usuais de representação dos dados estatísticos: os gráficos e as tabelas.

Fique atento aos conceitos dessas aulas, pois por meio deles poderemos compreender melhor a apresentação das tabelas e gráficos encontrados em jornais, revistas, e analisar e interpretar as informações transmitidas.



Figura 3.1 Gráficos

Fonte: <http://4.bp.blogspot.com>

A apresentação tabular 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, que normalmente é uma forma mais útil e elegante de demonstrar as características que serão analisadas.

3.1 Tabelas

É um quadro (sem que se fechem por completo as linhas e colunas, pois do contrário seria uma grade), 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 elaboração.



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

- 1) um traço horizontal (—) quando o valor é zero, não só quanto a natureza das coisas, como quanto ao 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; 0,000; ...).

Exemplo:

TABELA 11 - EVOLUÇÃO DA RENDA DOS EMPREENDEDORES BRASILEIROS 2000-2003				
FAIXA DE RENDA	2000 (%)	2001 (%)	2002 (%)	2003 (%)
Menos de 3 SM	30	35	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	8	9	8
Mais de 15 SM	9	11	3	6
Não Sabo	1	3	1	1
Recusou	4	6	2	2
Total	100	100	100	100

Fonte: Pesquisas GEM 2000 - 2003

Figura 3.2 Planilha

Fonte: Acervo do autor

3.2 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**.

3.2.1 Série temporal, histórica ou cronológica

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

Exemplo:

Tabela 3.1 Preço do artigo “Y” no atacado na cidade “X”

ANOS	PREÇO MÉDIO EM REAIS
2003	2,43
2004	2,54
2005	3,01
2006	2,99
2007	2,83

Fonte: Dados Fictícios

3.2.2 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:

Tabela 3.2 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

3.2.3 Série específica ou categórica

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

Exemplo:

Tabela 3.3 Número de Candidatos ao vestibular da Universidade “X”

ÁREA OFERTADA	NÚMERO DE CANDIDATOS
Ciências Sociais Aplicadas	2086
Ciências Exatas	1065
Ciências Humanas	1874
Ciências Biológicas	1102
Ciências Tecnológicas	1902

Fonte: Dados Fictícios

3.2.4 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:

Tabela 3.4 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

Resumo

Ao final desta aula você adquiriu conhecimentos suficientes para melhor interpretar e compreender qualquer tabela ou gráfico que lhe seja apresentado no decorrer de sua carreira profissional.

Portanto, mãos à obra! Este foi mais um passo muito importante à sua formação!

AULA 4 – Gráficos estatísticos

Na aula quatro veremos os tipos mais específicos de gráficos utilizados em pesquisas estatísticas, suas funções, vantagens e desvantagens.

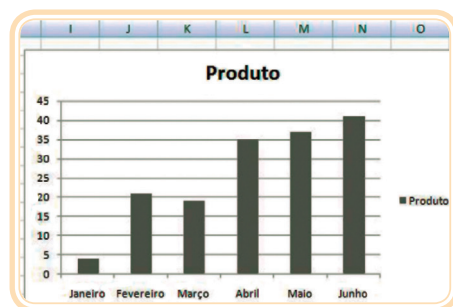


Figura 4.1: Gráfico em planilha

Fonte: Acervo do autor

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 frequências observadas. Os gráficos propiciam uma ideia 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:** o gráfico deve ser destituído de detalhes e traços desnecessários;
- **clareza:** o gráfico deve possuir uma correta interpretação dos valores representativos do fenômeno em estudo;
- **veracidade:** o gráfico deve expressar a verdade sobre o fenômeno em estudo.

4.1 Principais tipos de gráficos

4.1.1 Gráficos em curvas ou em linhas

São usados para representar séries temporais, principalmente quando a série cobrir um grande número de períodos de tempo e o mais relevante é

A-Z

Inclusão social: é um conjunto de meios e ações que combatem a exclusão aos benefícios da vida em sociedade, provocada pela diferença de classe social, origem geográfica, educação, idade, existência de deficiência ou preconceitos raciais

Hidrelétricas: usinas de água que geram energia elétrica

Costa: a borda do mar, litoral.

Análise Comparativa: examinar um conjunto, estabelecer algo semelhante ou diferente.

Socioeconômico: envolve uma combinação de fatores sociais e econômicos.

Reordenação institucional: ordenar, organizar novamente as questões referentes a instituições como associações, governos, colônias de pescadores, entre outros.

Precária: condição não garantida, nem estável.

Infraestrutura: divisão da sociedade em forças produtivas (homem e material necessário à produção) e relações de produção (classes, instituições etc.)

Beneficiamento: melhorar por processos técnicos.

o “sobe e desce” dos índices, taxas, coeficientes representados no eixo das ordenadas (eixo vertical).

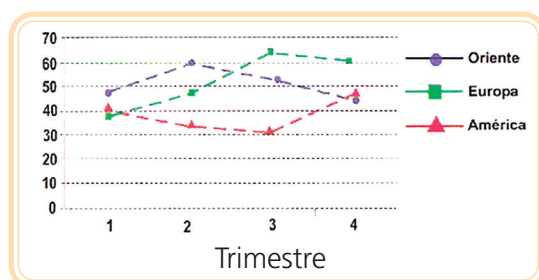


Figura 4.2: Gráfico em curvas e linhas

Fonte: imagem do autor

4.1.2 Gráficos em colunas

É a representação de uma série estatística através de retângulos, dispostos em colunas (na vertical). Este tipo de gráfico representa praticamente qualquer série estatística quando se deseja relacionar grandes quantidades de dados e as respectivas variações de crescimento e/ou decrescimento dos dados apresentados.

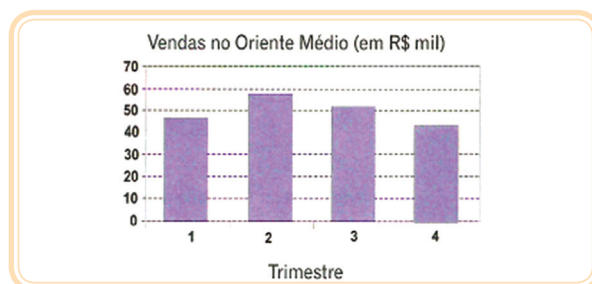


Figura 4.3: Gráficos com colunas

Fonte: imagem do autor.

Importante

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, mas depende do tipo de dado coletado e do autor adotado. Por hora adotaremos a seguinte regra: análise do tipo de variável coletada.

Por exemplo, se os dados forem variáveis do tipo contínua, não faz sentido apresentar as colunas separadas, pois os dados serão contínuos e, por consequência as colunas serão apresentadas de modo contínuo.

No caso do gráfico anterior com o cabeçalho: “**Vendas no Oriente Médio (em R\$ mil)**” as colunas estão separadas. A justificativa para isso é que a variável preço “em R\$ mil” não foi apresentada de modo contínuo na pesquisa, ou seja, no 1º Trimestre as Vendas foram de mais de R\$ 40 mil e ponto! No 2º Trimestre (não importa o dia exato em que se encerrou a coleta de dados do 1º Trimestre) iniciou-se nova coleta, que é **independente** da primeira.

4.1.3 Gráficos em barras

É representado por retângulos dispostos horizontalmente, prevalecendo os mesmos critérios adotados na elaboração de gráfico em coluna, porém é mais adequado quando se deseja destacar a variação (máximos e mínimos) de duas, até três variáveis do eixo vertical.

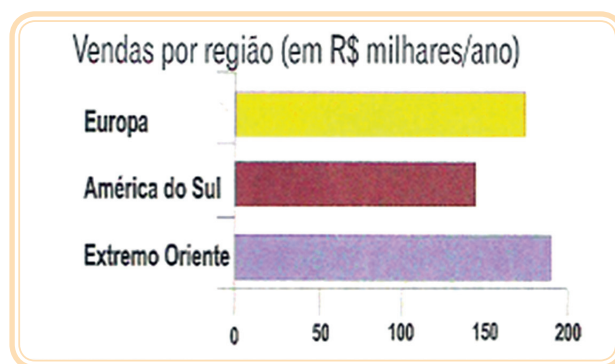


Figura 4.4: Gráficos em barras

Fonte: imagem do autor

4.1.4 Gráfico em setores

É a representação gráfica de uma série estatística em um círculo de raio qualquer, pôr 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° (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 proporção simples (a famosa “regra de três”):

Total → 360°

Parte → x°

O símbolo “→” pode ser lido como:

“está para”.

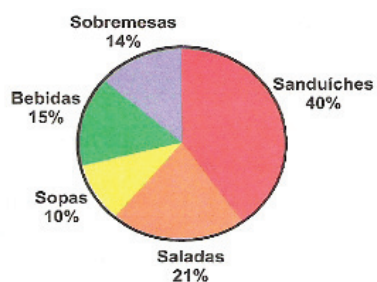


Figura 4.5: Gráfico em setores

Fonte: Acervo do autor

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

Pede-se:

Representar os dados obtidos graficamente:

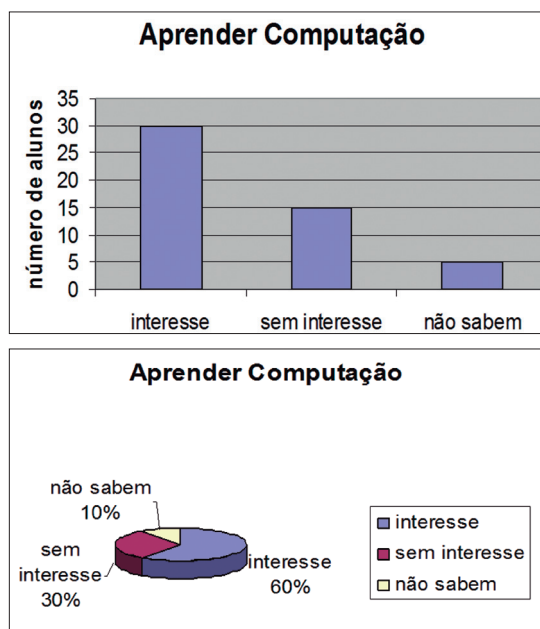
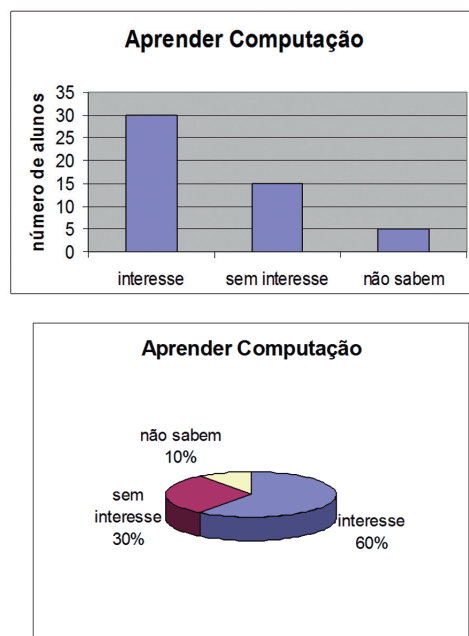


Figura 4.6: Gráficos em setores

Fonte: Acervo do autor

a) Comparação dos Gráficos:



Observe que comparamos lado a lado os dois gráficos, essa prática, é feita quando desejamos obter a melhor representação dos dados, seja para fins acadêmicos ou de apresentação profissional. Em síntese, ter várias opções de gráficos é desejável para podermos estabelecer melhor a comparação de um fenômeno e decidir sobre a melhor e mais impactante apresentação gráfica.

Para refletir

Qual dos dois representou melhor os dados obtidos na pesquisa?

Resumo

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
- geométrica, territorial ou de localidade
- específica ou categórica

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

Os gráficos podem ser classificados em:

- curvas ou linhas
- colunas
- barras
- setores (ou gráfico de pizza)

Anotações

[illegible]

Aula 5 – Distribuição de frequências

Na aula 5 você estudará que 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. Portanto, pode-se dizer que as distribuições de frequências são de extrema importância para a visualização dos dados e posterior análise das informações.

5.1 Representação dos dados (amostrais ou populacionais)

5.1.1 Dados brutos:

São aqueles que não foram numericamente organizados, ou seja, estão na forma com que foram coletados.

Por exemplo:

Número 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

5.1.2 Rol

É a organização dos dados brutos em ordem de grandeza crescente ou decrescente.

Por exemplo:

Número 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

Fonte: Acervo do autor

5.1.3 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:

Tabela 5.1 Número de pontos obtidos pelos alunos da disciplina-X, colégio-Y, em 2006.

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

Fonte: Acervo do autor

5.1.4 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:

Tabela 5.2 Peso de 80 estudantes da escola-x, 2007

Peso (kg)	Número de alunos
40 50	30
50 60	25
60 70	13
70 80	12
80 90	08

Fonte: Acervo do autor

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

a) 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						

Fonte: Dados fictícios

b) 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
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						

Fonte: Dados fictícios

c) Distribuição de frequências para dados agrupados em classes:

Taxas (em %)	Número de Municípios (f.)
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 (Σ)	94

Fonte: Acervo do autor

Representaremos o somatório pela letra grega maiúscula sigma (Σ).

Resumo

Vimos que através das distribuições em frequências poderemos organizar todas as informações obtidas em uma pesquisa, facilitando a compreensão e transmitindo adequadamente todas as informações ao leitor e ao pesquisador.

Não esqueça que esta "ferramenta" deverá ser utilizada, quando oportuna, pois enriquecerá seu trabalho.

Anotações

Aula 6 – Elementos principais da tabela de distribuição de frequências

6.1 Elementos de uma distribuição de frequência

6.1.1 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 $16 \mid 26$ define a segunda classe ($i = 2$). A distribuição é formada por nove classes, podemos afirmar que $i = 9$.

Observação: O símbolo representa um intervalo fechado à esquerda e aberto à direita: significa que o número a esquerda pertence à classe e o número a direita não pertence. É o mesmo que a representação matemática de conjuntos: lê-se “fechado à esquerda e aberto à direita”.

6.1.2 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 ls .

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

6.1.3 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.

6.1.4 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 $LS = 92$ e $Li = 6$, logo amplitude total é $H = 92 - 6 = 86$

6.1.4 Ponto Médio de uma Classe (xi):

É o ponto que divide o intervalo de classe em duas partes iguais.

$$xi = \frac{li + ls}{2}$$

Pelo exemplo prático anterior na quinta classe, temos: $li_3 = 46$ e $ls_3 = 56$, logo o ponto médio dessa classe é

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

Resumo

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 frequências: pode ser com ou sem intervalos de classe.

Os elementos da distribuição de frequência 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

$$h = ls - li$$

- amplitude total: é a diferença entre o maior e o menor da amostra

$$H = Ls - Li$$

- ponto médio: é a média aritmética dos limites da classe

$$xi = \frac{li + ls}{2}$$

Aula 7 - Distribuição de frequências

Nesta aula trataremos da segunda parte da distribuição de frequências em tabelas de estatística.

7.1 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 frequência nula ou muito pequena, não atingindo o objetivo de classificação que é tornar o conjunto de dados supervisionáveis.

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

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

Exemplo:

Se a amostra tiver 23 elementos analisados, o número de classes é 5, pois $n \leq 25$. Suponha que a amostra tenha 83 elementos analisados ($n \geq 25$) o número de classes é calculado por $\sqrt{83} = 9,1104335 \approx 9$.

b) 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;

Log é a abreviação de logaritmo e o seu valor pode ser obtido com uma calculadora científica;

n é o número total de observações.

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

nº total de observação	k = nº 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
362 ┤ 724	10
725 ┤ 1448	11
1.449 ┤ 2.896	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

Importante: 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:

$$K \approx 1 + 3,3 \cdot \log n \rightarrow K \approx 1 + 3,3 \cdot \log 94 \rightarrow K \approx 1 + 3,3 \cdot 1,97313 \rightarrow K 7,51 \rightarrow K \approx 8$$

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

- c) 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.

Importante:

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.

7.1.1 Amplitude do intervalo de classe (Ai):

É o comprimento da classe.

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

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

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 frequências com intervalos de classes.

1º.passo: A amplitude total será dada por:

$$H = L_s - L_i \rightarrow H = 218 - 154 = 64$$

2º.passo: Neste caso, $n = 32$ ► pela regra de Sturges (consultar Tabela)

$$K = 6$$

3º.passo: A amplitude do intervalo de cada classe será:

4º.passo: Construir a tabela de distribuição de frequências com intervalos de classes

$$A_i = \frac{H}{K} = \frac{64}{6} = 10,66... \approx 11$$

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.

Classes	Frequências
154 165	4
165 176	5
176 187	7
187 198	5
198 209	6
209 220	5
TOTAL (Σ)	32

Resumo

Independente do método escolhido para determinar o número de classes, o importante é utilizar o bom senso na tabulação e exposição dos dados.

Cada regra ou método tem características próprias e são adequados a este ou aquele número de elementos da população. Sendo assim, no momento da confecção da tabela é importante ter em mente a quantidade de dados e fazer um esboço da tabela.

Caso a última linha não contemple a plenitude dos dados, deve-se optar por outro método.

Aula 8 - Tipos de frequências

8.1 Frequências simples ou absolutas (f_i)

É o número de repetições de um valor individual ou de uma classe de valores da variável. A soma das frequências simples é igual ao número total dos dados da distribuição.

$$\sum F_i = n$$

8.2 Frequências relativas (fr_i)

São os valores das razões (divisões) entre as frequências absolutas de cada classe e a frequência total da distribuição. A soma das frequências relativas é igual a 1 ou 100 %

$$\text{Frequência relativa simples: } fr_i = \frac{F_i}{\sum F_i}$$

$$\text{Frequência relativa simples \% } fr_i = \frac{F_i}{\sum F_i} - 100$$

8.3 Frequência simples acumulada (fac_i)

É o total das frequências de todos os valores inferiores ao limite superior do intervalo de uma determinada classe.

8.4 Frequência relativa acumulada ($frac_i$)

É a frequência acumulada da classe, dividida pela frequência total da distribuição.

$$frac_i = \frac{Fac_i}{\sum F_i}$$

Resumo

Os tipos de frequências determinam com agilidade e objetividade a relação dos dados com o universo da pesquisa.

Em suma as frequências relativas (expressa em porcentagem) são úteis para a apresentação dos dados em linguagem comum e de amplo entendimento, que é a porcentagem, e as frequências acumuladas servem de gabarito para que se tenha certeza de que a totalidade dos dados coletados foram apresentados.

Anotações

[illegible]

Aula 9 - Resolução de exercício prático

Vamos colocar em prática, juntos, todos esses conceitos que estudamos nas aulas 5, 6, 7 e 8 através de um exercício de aplicação.

Atividades de aprendizagem

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 valores:



7, 9, 1, 3, 3, 3, 4, 5, 5, 7, 8, 9, 3, 2, 4, 5, 6, 7, 8, 9, 2, 3, 4, 2, 1.

Pede-se:

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

Solução:

O Rol é:

A amplitude total será dada por:

Neste caso, $n = 25$ ► 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 da cada classe e completando a tabela com as frequências, temos:

Classe	f_i	fr_i	$fr_i (\%)$	fac_i	$faci(\%)$	X_i
Total (Σ)						

Resumo

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 $K \approx$

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:

Amplitude do intervalo de classe: é o conjunto da classe, calculado por:

Frequência simples ou absoluta: é o número de repetições de um valor individual.

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

Frequência relativa: são os valores das divisões entre as frequências absolutas de cada classe e a frequência total da distribuição.

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

Frequência simples acumulada: é o total das frequências de todos os valores inferiores ao limite superior do intervalo de uma determinada classe.

Frequência relativa acumulada: é a frequência acumulada da classe dividida pela frequência total da distribuição.

Aula 10 - Distribuição de frequências graficamente

Nesta aula estudaremos a distribuição de frequências no modo gráfico: o Histograma e o Polígono de Frequência.

10.1 Histograma e polígono de frequências

10.1.1 Histograma

É um tipo de gráfico (semelhante ao gráfico de colunas) formado por um conjunto de retângulos justapostos e é muito utilizado para representar a distribuição de frequências cujos dados foram agrupados em classes ou intervalos de mesma amplitude. A base do retângulo é igual à amplitude do intervalo classe e sua altura é proporcional à frequência da classe.

Exemplo:

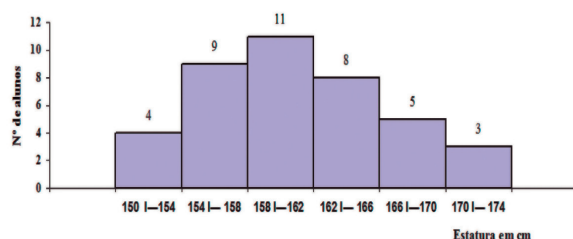


Figura 10.1: Estaturas dos alunos da Turma "A" – 2006

Fonte: Fictícia – imagem do autor

Polígono de Frequências: é obtido unindo-se por segmentos de reta os pontos médios das bases superiores dos retângulos de um histograma desenhando então um polígono de área bem definida.

Importante: A área do polígono e o número de total de observações é diretamente proporcional ao número de observações do Universo ou Espaço Amostral da pesquisa.

Exemplo:

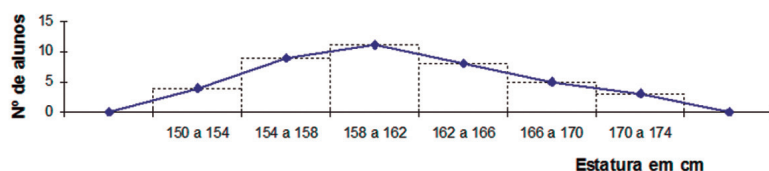


Figura 10.2: Estaturas dos alunos da Turma "A" – 2006 (polígono de frequências)

Fonte: Fictícia – imagem do autor



Aula 11 - Resolução de exercícios práticos

Vamos colocar em prática, juntos, todos os conceitos da aula anterior através de exercícios.

Atividades de aprendizagem



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	162	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:

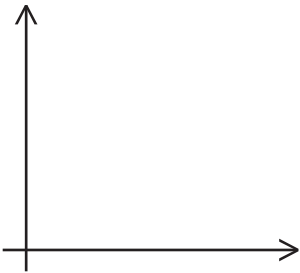
Antes de construirmos o histograma e o polígono de frequências, temos que montar a tabela com dados agrupados por classes, desta forma faremos inicialmente os seguintes cálculos:

- 1°. A amplitude total será dada por:
- 2°. Neste caso, $n = 40$ ► pela regra de Sturges, temos:
- 3°. A amplitude do intervalo de cada classe será:

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

Classes	fi
TOTAL (Σ)	

4º. Depois de montarmos a tabela com dados agrupados em classes, iremos construir o histograma e o polígono de frequência:



Resumo

Histograma: é um gráfico formado por um conjunto de retângulos justapostos.

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

Anotações

Aula 12 - Medidas de posição

Nesta aula veremos as medidas de posição, em especial as medidas de tendência central: média, mediana e moda.

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.

12.1 Médias

12.1.1 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.

$$\mu = \frac{X_1 + X_2 + X_3 + \dots + X_n}{N} \quad \text{ou} \quad \mu = \frac{\sum X_i}{N}$$

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n} \quad \text{ou} \quad \overline{X} = \frac{\sum X_i}{n}$$

Onde: x_i = são os valores que a variável x assume

n = o número de valores

\bar{X} = é a média aritmética da amostra

μ = 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:

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

12.1.2 Média Aritmética Ponderada

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

$$X_p = \frac{\sum X_i \cdot p_i}{\sum p_i}$$

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	Taxas 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:

$$\overline{x_p} = \frac{x_1.p_1 + x_2.p_2 + x_3.p_3}{p_1 + p_2 + p_3}$$

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

12.1.3 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.

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

$$\overline{X} = \frac{\sum X_i.f_i}{n} \quad (\text{Amostra})$$

Exemplo:

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

Notas dos alunos - x_i	Número de alunos - f_i	$x_i f_i$
1	1	1
2	3	6
3	5	15
4	1	4
Total (Σ)	$n = 10$	26

$$\overline{X} = \frac{\sum X_i.f_i}{n} = \frac{26}{10} = 2,6$$

12.1.4 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:

$$\mu = \frac{\sum X_i \cdot f_i}{N}, \text{ onde } X_i = \frac{l_i + l_s}{2}$$

(População)

$$\overline{X} = \frac{\sum X_i \cdot f_i}{n}, \text{ onde } x_i = \frac{l_i + l_s}{2}$$

(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 Σ		n = 40	268

$$\overline{X} = \frac{\sum X_i \cdot f_i}{n} = \frac{268}{40} = 6,7$$

Resumo

Média aritmética simples:

(Referente à população)

$$\mu = \frac{X_1 + X_2 + X_3 + \dots + X_n}{N} \quad \text{ou} \quad \mu = \frac{\sum X_i}{N}$$

Média aritmética simples:

(Referente à amostra)

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n} \quad \text{ou} \quad \overline{X} = \frac{\sum X_i}{n}$$

Média aritmética ponderada

$$\overline{X}_p = \frac{\sum X_i \cdot p_i}{\sum p_i}$$

Média aritmética para dados agrupados sem intervalos de classes:

$$\mu = \frac{\sum X_i \cdot f_i}{N}$$

(População)

$$\overline{X} = \frac{\sum X_i \cdot f_i}{n}$$

(Amostra)

Média aritmética para dados agrupados com intervalos de classes:

$$\mu = \frac{\sum X_i \cdot f_i}{N}, \text{ onde } X_i = \frac{l_i + l_s}{2}$$

(População)

$$\overline{X} = \frac{\sum X_i \cdot f_i}{n}, \text{ onde } x_i = \frac{l_i + l_s}{2}$$

(Amostra)

Anotações



Aula 13 - Medidas de posição - moda e mediana

Nesta aula abordaremos dois tipos de medida de tendência central: a Moda e a Mediana, considerando que esta tem caráter mais geométrico, de simetria, em comparação às demais medidas.

Define-se a moda como o valor que ocorre com maior frequência em conjunto de dados. Tal e qual aparece no senso comum, a moda “vai e volta”, e várias vezes se repete.



Figura 13.1: Moda

Fonte: Acervo do autor

13.1 Moda (mo) – para dados não agrupados

Primeiramente os dados devem ser ordenados (colocados em Rol) para, em seguida, podermos 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 frequente)

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 frequentes)

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 frequentes)

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, ou seja, não tem moda.

Observação:

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. É uma medida pouco utilizada.

Já a **média aritmética** é a medida de posição que possui a maior confiabilidade numérica, além de ser a mais intuitiva, do ponto de vista matemático.

13.2 Mediana (Md)

Nesta aula abordaremos outro tipo de medida de tendência central: a Mediana. Como já dissemos, esta tem caráter mais geométrico, de simetria, em comparação as demais medidas.

É uma medida de posição (Medida de Tendência Central) cujo valor 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 procedimentos:

- Se **N** (número de elementos do conjunto) for **ímpar** a mediana é o termo de ordem P dado pela razão:

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

- Se **N** (número de elementos do conjunto) for **par**, a mediana é a média aritmética dos termos de ordem, em um primeiro passo de P1 (média aritmética simples) e, em seguida, pela razão P2 (termo subsequente da ordem P):

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

Exemplos:

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

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

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.

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

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

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

Ou ainda, de modo simplista, dividimos os dados em limites à esquerda e à direita, destacando os valores centrais.

Destaque: O procedimento seguinte é para uma quantidade par de elementos.

1º. passo: organizamos o Rol:

Rol: 12, 12, 13, 13, 15, 16, 18, 20.

2º. passo: destacamos os valores centrais (metade à esquerda, metade à direita):

12, 12, 13, 13, 15, 16, 18, 20

3º. passo: efetuamos a média aritmética:

$$\frac{13+15}{2} = \frac{28}{2} = 14$$

Resumo

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 dois casos:

1º. caso: se **N** for **ímpar**, a mediana é o termo de ordem **P** dado pela razão:

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

2º. caso: se **N** for **par**, a mediana é a média aritmética dos termos de ordem **P** dado pela razão:

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

Anotações

Aula 14 - Comparação entre média, mediana e moda

Nesta aula apresentaremos um quadro comparativo das vantagens e desvantagens dos diferentes tipos de medidas de posição.

MEDIDA	DEFINIÇÃO	VANTAGENS	DESVANTAGENS
Média	Centro de distribuição de frequências.	<ul style="list-style-type: none">- reflete cada valor.- possui propriedades matemáticas atraentes.	<ul style="list-style-type: none">- é afetada por valores extremos.
Mediana	Metade dos valores são maiores, metade menores.	<ul style="list-style-type: none">- menos sensíveis a valores extremos do que a média.	<ul style="list-style-type: none">- difícil de determinar quando há uma grande quantidade de dados.
Modo	Valor mais frequente.	<ul style="list-style-type: none">- valor "típico": maior quantidade de valores concentrados neste ponto.	<ul style="list-style-type: none">- não se presta à análise matemática.- pode não ter modo para certos conjuntos de dados.

Figura 14.1: Comparação entre média, mediana e moda

Fonte: Acervo do autor

Memorize as seguintes frases (três frases mágicas, garantimos!):



- 1ª. A seta puxa a Média!
- 2ª. A Moda está no topo!
- 3ª. A Mediana está no meio!

Assim, nestas pequenas frases teremos as características das três Medidas de Posição.

Resumo

A Média é sempre influenciada por valores extremos, os quais são "atraídos pela seta".

A Moda é o elemento de maior frequência, e a maior frequência está no topo (no ponto mais alto da curva).

A Mediana está sempre no meio do conjunto, dividindo-o em duas partes iguais!



Aula 15 - Medidas de dispersão (medidas de variabilidade)

Nesta aula veremos mais um tipo de medida estatística: a dispersão ou variabilidade.



Figura 15.1: Linha de produção

Fonte: Acervo do autor

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

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

Empregado A: 70, 71, 69, 70, 70 → = 70

Empregado B: 60, 80, 70, 62, 83 → = 71

O desempenho médio 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 o desempenho 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 o desempenho de A é bem mais uniforme do que de B.

15.1 Tipos de medidas de dispersão absoluta

15.1.1 Amplitude total (AT):

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:

- Empregado A $\rightarrow A_T = 71 - 69 = 2$
- Empregado B $\rightarrow A_T = 83 - 60 = 23$

Resumo

A amplitude total é a medida mais simples de dispersão.

- A desvantagem desta medida de dispersão é que ela 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).

Anotações

Aula 16 - Variância (σ^2 ou s^2) e desvio padrão (σ ou s)

Nesta aula veremos mais um tipo de medida de dispersão ou variabilidade: o cálculo da Variância e do Desvio Padrão.

A **Variância** e o **Desvio Padrão** são as medidas de dispersão mais empregadas nos procedimentos estatísticos de constatação da variação de um dado e a consequente validação de uma pesquisa, pois leva em consideração a totalidade dos valores da variável em estudo. É um indicador de variabilidade bastante estável e de alto grau de confiabilidade. 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	$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$	$\sigma^2 = \frac{\sum (x_i - \mu)^2 f_i}{N}$
Amostra	$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$	$s^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{n-1}$

Figura 16.1: Dados

Fonte: Acervo do autor

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.

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

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 as 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:

$$S = \sqrt{s^2} = \sqrt{0,4} \cong 0,63$$

- 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:

$$S = \sqrt{s^2} = \sqrt{85,6} \cong 9,25$$

Aula 17 - Exercício prático sobre variância e desvio padrão

Na aula 17 veremos uma aplicação do cálculo de Variância e Desvio Padrão em uma amostra de funcionários da “Empresa Z”.

Atividades de aprendizagem



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 Σ	40				

Figura 17.1: Salários semanais em reais de uma amostra dos funcionários da empresa

Fonte: Acervo do autor

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:

$$s = \sqrt{s^2}$$

Resumo

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:

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

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

Variância de dados agrupados em classe:

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

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

Desvio padrão:

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

Anotações

Aula 18 - Probabilidade

Nesta aula trataremos da Probabilidade ou Cálculo Probabilístico. Trata-se de assunto complexo no campo da Teoria da Probabilidade, porém abordaremos somente a relação da Estatística com o princípio da contagem.

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 fazer 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.

18.1 Experimento aleatório

São aqueles que não podem ser previamente determinados. A esta impossibilidade de preverem-se os resultados chamamos de acaso.

Exemplo: Lançamento de dados. Em seguida anotar o número que ocorrerá na face voltada para cima.



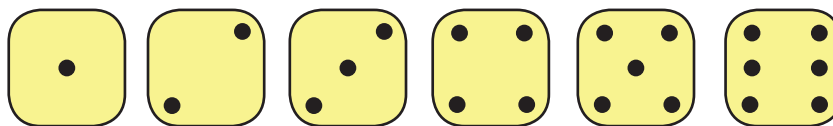
Figura 18.1: Dados
Fonte: www.sxc.hu

18.2 Espaço Amostral (S) ou Universo

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:

$S = \{1, 2, 3, 4, 5, 6\}$, ou seja, o Espaço Amostral (ou Universo) é formado por seis elementos.



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

$S = \{v, e, d\}$, ou seja, o Espaço Amostral (ou Universo) é formado por três elementos.



Figura - 18.2 Futebol

Fonte: Acervo do autor

18.3 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 duas moedas pode ser que apareçam faces iguais. Os elementos do evento são $E = \{(K, K), (C, C)\}$.



Figura 18.3: Moeda cara e coroa

Fonte: Acervo do autor

18.3.1 Tipos de eventos

- Evento certo – é o próprio espaço amostral.

Exemplo: Lançamento de um dado e ocorrência de um número menor ou igual a seis 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.

Resumo

O problema fundamental da probabilidade consiste em atribuir um número a cada evento (E), o qual avaliará as chances de ocorrência de (E).

O quadro seguinte resume a simbologia utilizada na teoria da probabilidade:

U ► representa a soma, sendo representado pelo sinal "+", ou pela palavra "ou".

\cap ► representa a multiplicação, sendo representada pelo sinal "x", ou pela palavra "e".

$A \cup B$ ► é o evento que ocorre se A ocorrer ou B ocorrer ou ambos ocorrerem.

$A \cap B$ ► É evento que ocorre se A e B ocorrerem.
 A^c ► É evento que ocorre se A não ocorrer.

N ► Fornece o número total de eventos de S.
 N ► É o número de casos favoráveis.
 n ► É o número de elementos.

Figura 18.4: Quadro resumo de probabilidade – simbologia

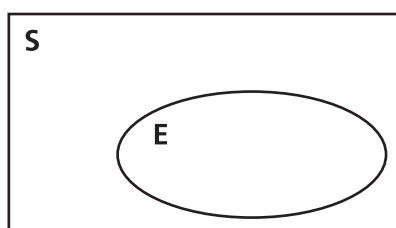
Fonte: <http://w3.ufsm.br/>



Aula 19 - Cálculo da probabilidade de um evento ocorrer – princípio da contagem

Nesta aula veremos um tipo bastante simples de cálculo de probabilidade, utilizando o princípio da contagem, e a representação das razões em termos de porcentagem.

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 (preferencialmente em por cento)

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

Na prática, calcular a probabilidade é dividir.

Observação:

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

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

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

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

Exemplo prático:

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 a nomenclatura “ $\sim E$ ” como sendo o 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, proceda assim:

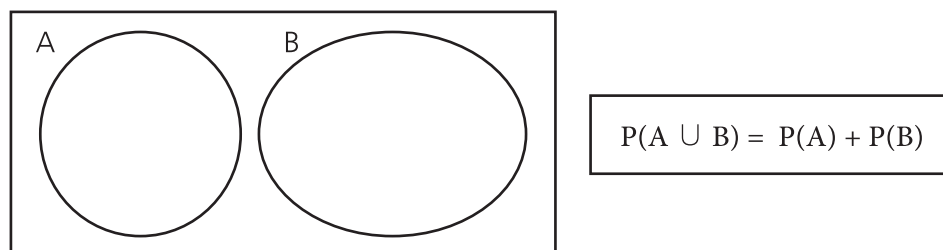
$$P(\sim E) = \frac{n(\sim E)}{n(S)} = \frac{320}{360} = \frac{8}{9} \cong 88,9\%$$

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

Nesse caso, existem dois tipos possíveis de situação:

- **Eventos mutuamente exclusivos:**

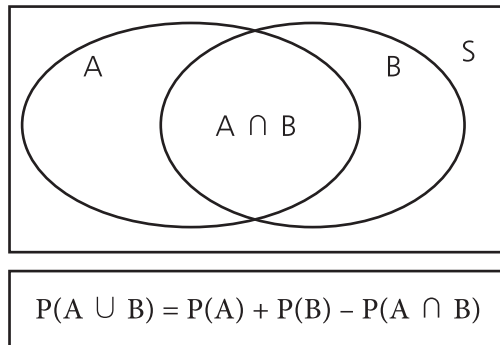
Dois eventos são mutuamente exclusivos se $A \cap B = \emptyset$, neste caso:



- **Eventos não mutuamente exclusivos:**

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

a probabilidade de se obter um número par ou maior que três.



- Espaço amostral: $S = \{1, 2, 3, 4, 5, 6\} \rightarrow n(S) = 6$
- Evento A (números pares): $A = \{2, 4, 6\} \rightarrow n(A) = 3$
- Evento B (números maiores que 3): $B = \{4, 5, 6\} \rightarrow n(B) = 3$
- Evento de $A \cap B$: $A \cap B = \{4, 6\} \rightarrow n(A \cap B) = 2$
- Calculando a probabilidade, temos:

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

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

Anotações



Aula 20 - Probabilidade condicional

Nesta aula veremos a Probabilidade Condicional, que envolve conceitos da Teoria dos Conjuntos, como a União (\cup) e a Interseção (\cap) de conjuntos.

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:

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

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

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

20.1 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 pelo produto de P (A) e P (B):

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

Exemplo prático:

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, a 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**.

Considere então os eventos:

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

Como queremos calcular P(A e B), pense assim: de todos os jovens, $\frac{1}{5}$ têm

média acima de 7,0 e $\frac{5}{6}$ sabem jogar futebol. Ora, $\frac{5}{6}$ de $\frac{1}{5}$, ou seja, $\frac{5}{6} \cdot \frac{1}{5} = \frac{1}{6}$

sabem jogar futebol e têm média acima de 7,0. Portanto, $P(A \text{ e } B) = \frac{1}{6}$.

Resumo

Experimentos aleatórios: são aqueles que, 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:

$P(E) = \frac{\text{número de resultados favoráveis}}{\text{número de resultados possíveis}}$ com $0\% \leq P(E) \leq 100\%$ ou $0 \leq P(E) \leq 1$.

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:

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

Probabilidade de eventos independentes:

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

Anotações



Referências

COSTA NETO, P.L. de O. **Probabilidades**. São Paulo: Edgard Blucher, 1985.

COSTA NETO, P.L. de O. **Estatística**. 17. ed. São Paulo: Edgard Blucher, 1999.

CRESPO, A.A. **Estatística Fácil**. 17. ed. São Paulo: Saraiva, 1999.

DANTE, L.R. **Matemática**: Contexto de Aplicações. São Paulo: Ática, 1999.

DOWNING, D., CLARK, J. **Estatística Aplicada**. São Paulo: Saraiva, 2000.

KAZMIER, L.J. **Estatística Aplicada à Economia e Administração**. São Paulo: Makron Books, 1982.

LAPPONI, J.C. **Estatística usando o Excel**. São Paulo: Lapponi, 2000.

LEVIN, J. **Estatística Aplicada a Ciências Humanas, 2ª edição**. São Paulo: Harper & Row do Brasil, 1978.

NICK, E., KELLNER, S. R. O. **Fundamentos de Estatística para as Ciências do Comportamento**. Rio de Janeiro: Renes, 1971.

SIEGEL, S. **Estatística Não Paramétrica**. São Paulo: McGraw-Hill do Brasil, 1975.

STEVENSON, W. J. **Estatística Aplicada à Administração**. São Paulo: Harper & Row do Brasil, 1981.

TRIOLA, M. F. **Introdução à Estatística**. 7. ed. Rio de Janeiro: Livros Técnicos e Científicos, 1999.

Referências das Figuras

Figuras 1.1 a 1.5
Acervo do professor-autor

Figura 1.6 Estudos
www.sxc.hu

Figura 1.7 População

[http:// www.seplan.go.gov.br/sepin/pub/GoDados/2003/imagens/demografia.jpg](http://www.seplan.go.gov.br/sepin/pub/GoDados/2003/imagens/demografia.jpg)

Figura 1.10 Colégio

http://2.bp.blogspot.com/_A5IDaz1qLcg/S7C9R3eon_I/AAAAAAAAIU/gH68-YYCBZw/s1600/2007_school_escola.jpg

Figura 1.12 Gráfico

<http://uolpolitica.blog.uol.com.br/images/Mulher3.jpg>

Figura 1.14 Conjunto de números reais

<http://alfaconnection.net/images/NUM010106a.gif>

Figura 3.1 Gráficos

http://4.bp.blogspot.com/_mzA7VqOPJc4/S_xnLUnwBWI/AAAAAAAAAeM/8bZ88_rxxEk/s1600/modelos-graficos1.jpg

Figura 3.2 até 18.3

Acervo do professor-autor

Atividades autoinstrutiva

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, as variáveis “cor dos cabelos” e número de irmãos” 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.

6. Classifique a série abaixo:

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

APLICAÇÕES EM MILHÕES DE REAIS, DO ÚLTIMO QUADRIMESTRE DE 2006, EM UM DETERMINADO BANCO

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

7. Classifique a série abaixo:

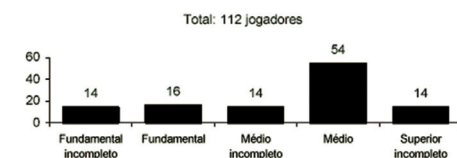
Porte	Exportações
Pequeno	100
Médio	150
Grande	200

EXPORTAÇÕES EM MILHÕES DE DÓLARES POR PORTE DAS EMPRESAS, NO ANO DE 2006

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

- 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.

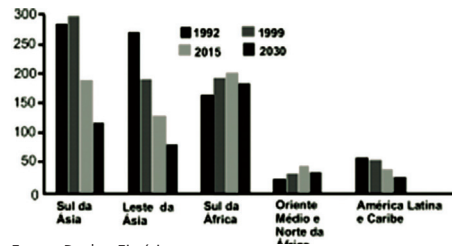
De acordo com esses dados, o percentual dos jogadores dos quatro clubes que concluíram o Ensino Médio é de aproximadamente:



Fonte: Dados Fictícios (O Globo, 24/07/2005.)

- 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.** Segundo a Organização das Nações Unidas (ONU), o mundo não conseguirá atingir a meta de reduzir a fome pela metade em 2015. Nem mesmo em 2030 esse objetivo poderá ser alcançado. O gráfico a seguir mostra o número, em milhões, de pessoas com fome em cinco regiões do mundo, em diferentes anos (1992, 1999, 2015 e 2030), segundo dados e estimativas da ONU.

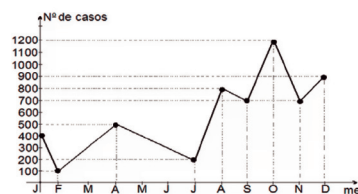


Fonte: Dados Fictícios

Com base nos dados fornecidos pelo gráfico, pode-se **AFIRMAR** que:

- a) em 2030, haverá mais de 700 milhões de pessoas com fome nas regiões destacadas no gráfico;
- b) em cada região destacada no gráfico, o número de pessoas com fome em 2030 será menor do que em 1992;
- c) em cada região destacada no gráfico, o número de pessoas com fome em 2030 será menor do que em 2015;
- d) em cada região destacada no gráfico, o número de pessoas com fome em 2015 será menor do que em 1999;
- e) em 2030, o número de pessoas com fome no Sul da África será maior do que três vezes o número de pessoas com fome no Sul da Ásia.

- 11.** Em um município foi pesquisado, durante um ano, o número de casos de certa doença, encontrando-se os dados representados no gráfico abaixo:



Pela análise do gráfico, é **CORRETO** afirmar que:

- a) O número de casos foi máximo no mês de dezembro.
- b) O número total de registros no 2º semestre é de 1500 casos.
- c) A maior diferença do número de casos registrados ocorreu entre os meses de julho e agosto.
- d) O número de casos foi mínimo no mês de janeiro.
- e) O número total de registros no 3º trimestre é de 1700 casos.

12. Quando um intervalo de classe está representado por $1,250 \mid 1,260$, 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.

13. Com relação à distribuição de frequência dada abaixo, é possível dizer que:

Comprimento, (em cm) de certa peça produzida pela indústria "y"

Comprimento (cm)	fi
$12,32 \mid 12,36$	3
$12,36 \mid 12,40$	5
$12,40 \mid 12,44$	12
$12,40 \mid 12,44$	15
$12,48 \mid 12,52$	8
$12,52 \mid 12,56$	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.

14. Dada a distribuição de frequências a seguir, responda qual o limite superior da quarta classe:

idades	Frequência (f)
$0 \mid 2$	2
$2 \mid 4$	5
$4 \mid 6$	18
$6 \mid 8$	10
$8 \mid 10$	5

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

15. Na distribuição de frequências da questão 14, qual a amplitude de cada classe ou intervalo?

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

16. Na distribuição de frequências da questão 14, qual o ponto médio da quinta classe ou intervalo?

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

17. 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	Frequência simples	Frequência relativa	Porcentagem
Humanas	8	$\frac{8}{20}=0,4$	40%
Biológicas	□	$\frac{7}{20}=0,35$	36%
Exatas	5	□	25%
Total	20	1,00	100%

A frequência simples e a frequência relativa que faltam (assinaladas por \square) 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 $\frac{5}{20}$

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

Obs.: A última frequência acumulada relativa resultou 1,0001. Por critérios de arredondamento, é perfeitamente admissível desprezar a última casa decimal.

Número de irmãos

Quantidade de irmãos	f_i	f_{ri}	f_{ai}	f_{ari}
0	2	0,0571	2	0,0571
1	5	0,1429	7	0,2000
2	8	0,2286	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	1,0001
	35	1,0000		

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.
- e) temos 15 pessoas com apenas 2 irmãos.

19. Dada a distribuição de frequências a seguir, responda qual a frequência acumulada total:

Idades	Frequências (f)
19 21	8
21 23	12
23 25	15
25 27	13
27 29	5
29 31	7

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

As questões de número 20 e 21 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

20. O limite superior da quinta classe é:

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

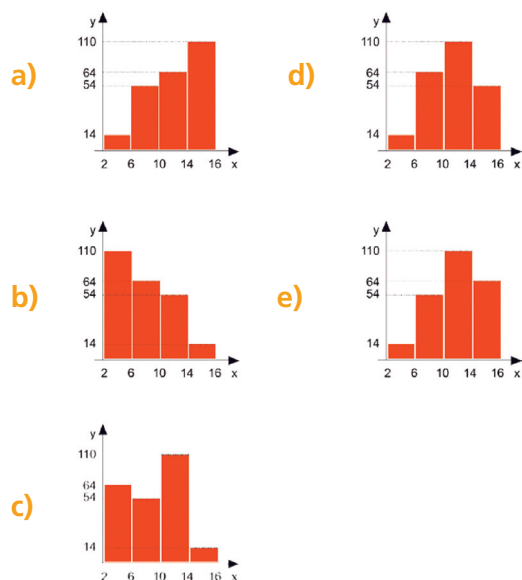
21. As frequências da terceira e da quinta classe são, respectivamente:

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

22. Os dados da tabela abaixo, gerados pelo departamento de qualidade de uma indústria, são agrupados em classes, mostrando a distribuição das peças defeituosas em agosto de 2008.

	X: classe	y: número de peças com defeito
I	2-6	14
II	6-10	54
III	10-14	110
IV	14-16	64

O histograma que melhor se adapta aos dados da tabela é:



As questões de número 23, 24, 25 e 26 devem ser respondidas com base na tabela de dados obtidos em uma pesquisa.

Os seguintes dados se referem à média salarial (em R\$) dos funcionários de uma determinada empresa.

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

23. Qual será a amplitude total da distribuição de frequência?

- a) 164
- b) 154
- c) 174
- d) 184
- e) 176

24. Qual será a amplitude de cada uma das distribuições classes?

- a) 28
- b) 30
- c) 27
- d) 29
- e) 26

25. As frequências simples da terceira e da quinta classe são, respectivamente:

- a) 4 e 7
- b) 5 e 7
- c) 2 e 4
- d) 6 e 7
- e) 7 e 4

26. As frequências acumulada relativa da segunda e da quarta classe são, respectivamente:

- a) 0,24 e 0,54
- b) 0,6 e 0,76
- c) 0,04 e 0,6
- d) 0,6 e 1
- e) 0,24 e 0,6

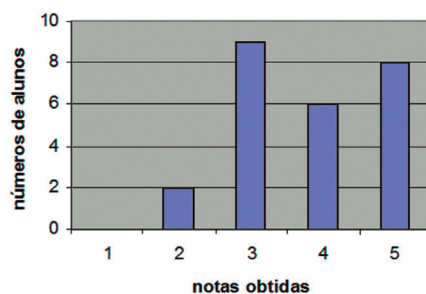
27. A tabela a seguir representa a distribuição de frequências dos salários de um grupo de 50 empregados de uma empresa, num certo mês.

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

28. O professor Javert e o professor Roberto aplicaram uma prova de matemática a 25 alunos, contendo 5 questões, valendo 1 ponto cada uma. Após fazer a correção, o professor Roberto construiu o gráfico seguinte, que relaciona o número de alunos às notas obtidas por eles.

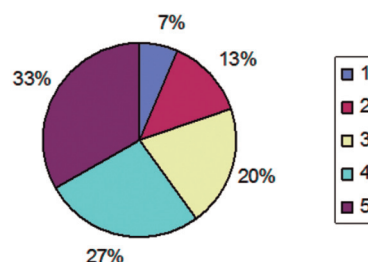


Fonte: Dados Fictícios

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

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

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



Fonte: Dados Fictícios

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:

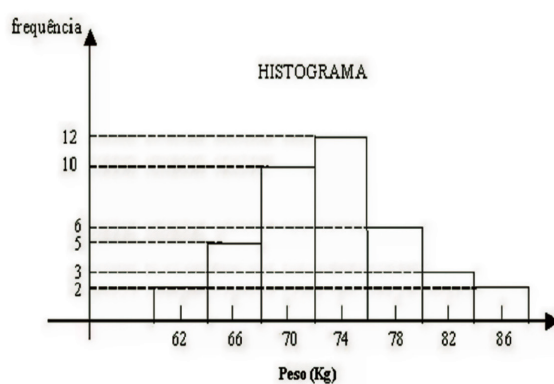
- a) 1
- b) 2
- c) 3
- d) 4
- e) 5

30. 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

31. Uma equipe de futebol realizou um levantamento dos pesos dos seus 40 atletas e chegou à distribuição de frequência dada pela tabela a seguir, cujo histograma correspondente é visto abaixo.

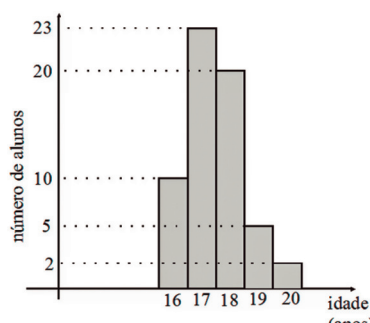
Peso (kg)	Frequência
60 64	2
64 68	5
68 72	10
72 76	12
76 80	6
80 84	3
84 88	2
Total de atletas	40



Com base nestes dados, pode-se afirmar que o valor da mediana dos pesos é igual a:

- a) 75
- b) 72
- c) 74
- d) 73
- e) 71

32.A distribuição das idades dos alunos de uma classe é dada pelo gráfico abaixo.



Qual das alternativas representa melhor a média de idades dos alunos?

- a) 16 anos e 10 meses.
- b) 17 anos e 1 mês.
- c) 17 anos e 5 meses.
- d) 18 anos e 6 meses.
- e) 19 anos e 2 meses.

33.Na busca de solução para o problema da gravidez na adolescência, uma equipe de orientadores educacionais de uma instituição de ensino pesquisou um grupo de adolescentes de uma comunidade próxima a essa escola e obteve os seguintes dados:

Idade (em anos)	Frequência absoluta de Adolescentes Grávidas
13	4
14	3
15	2
16	5
17	6

Com base nos textos e em seus conhecimentos, é CORRETO AFIRMAR, em relação às idades das adolescentes grávidas, que:

- a) a média é 15 anos.
- b) a mediana é 15,3 anos.
- c) a mediana é 16,1 anos.
- d) a moda é 16 anos.
- e) a média é 15,3 anos.

34. 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

35. 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.

36. Calcule a amplitude total do rol.

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

37. Calcule a variância do rol.

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

38. Calcule o desvio padrão do rol.

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

39. O serviço de atendimento ao consumidor de uma concessionária de veículos recebe as reclamações dos clientes via telefone. Tendo em vista a melhoria nesse serviço, foram anotados os números de chamadas durante um período de sete dias consecutivos. Os resultados obtidos foram os seguintes:

Dia	Domingo.	Segunda	Terça	Quarta	Quinta	Sexta	Sábado
Número de chamadas	3	4	6	9	5	7	8

Sobre as informações contidas nesse quadro, considere as seguintes afirmativas:

- I. O número médio de chamadas dos últimos sete dias foi 6.
- II. A variância dos dados é 4.
- III. O desvio padrão dos dados é $\sqrt{2}$.

Assinale a alternativa CORRETA.

- a) Somente as afirmativas I e II são verdadeiras.
- b) Somente as afirmativas I e III são verdadeiras.
- c) Somente as afirmativas II e III são verdadeiras.
- d) Somente a afirmativa I é verdadeira.

40. Na tabela abaixo encontra-se a distribuição de frequência dos salários das três funções existentes em uma empresa de médio porte.

Função	Salário (R\$)	Número de Funcionários
Operário	500,00	40
Inpetor	2.500,00	8
Diretor	5.000,00	2

Com base nesses dados, assinale a alternativa **CORRETA**.

- a) O salário médio dessa empresa é de R\$ 1.000,00.
- b) O salário médio dessa empresa é obtido somando-se os salários das três funções e dividindo-se esse resultado por três.
- c) A mediana é o salário de R\$ 2.500,00, recebido pelo inspetor dessa empresa.
- d) O desvio padrão é zero, pois 80% dos funcionários recebem o mesmo valor.
- e) Existe uma grande homogeneidade de salários, pois a maioria dos funcionários recebe o mesmo valor de R\$ 500,00.

41. Uma determinada região apresentou, nos últimos cinco meses, os seguintes valores (fornecidos em mm) para a precipitação pluviométrica média:

A média, a mediana e a variância do conjunto de valores acima são, respectivamente:

Jun	Jul	Ago	Set	Out
32	34	27	29	28

- a) 30, 27 e 6,8.
- b) 27, 30 e 2,4.
- c) 30, 29 e 6,8.
- d) 29, 30 e 7,0.
- e) 30, 29 e 7,0.

42.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 ganhado?

- a) $10/36$
- b) $5/32$
- c) $5/36$
- d) $5/35$
- e) Não se pode calcular sem saber os números sorteados.

43.(UNIRIO) As probabilidades de três jogadores marcarem um gol cobrando um pênalti são, respectivamente, $\frac{1}{2}$, $\frac{2}{5}$ e $\frac{5}{6}$. 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%

44.(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) $1/6$

45. 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 ser fumante e mulher é de aproximadamente:

- a) 0,044
- b) 0,075
- c) 0,44
- d) 0,0075
- e) 0,0044

46. Um casal planeja ter 3 filhos. Sabendo que a probabilidade de cada um dos filhos nascer do sexo masculino ou feminino é a mesma, considere as seguintes **AFIRMATIVAS**:

- I. A probabilidade de que sejam todos do sexo masculino é de 12,5%.
- II. A probabilidade de o casal ter pelo menos dois filhos do sexo feminino é de 25%.
- III. A probabilidade de que os dois primeiros filhos sejam de sexos diferentes é de 50%.
- IV. A probabilidade de o segundo filho ser do sexo masculino é de 25%.

Assinale a alternativa CORRETA.

- a) Somente as afirmativas I e III são verdadeiras.
- b) Somente as afirmativas I e II são verdadeiras.
- c) Somente as afirmativas II e III são verdadeiras.
- d) Somente as afirmativas I, III e IV são verdadeiras.
- e) Somente as afirmativas II e IV são verdadeiras.

47. As cartas de um baralho são amontoadas aleatoriamente. Qual é a probabilidade de a carta de cima ser de copas e a de baixo também? (O baralho é formado por 52 cartas de 4 naipes diferentes 13 cartas por naipe).

- a) $\frac{1}{17}$ b) $\frac{1}{25}$ c) $\frac{1}{27}$ d) $\frac{1}{35}$ e) $\frac{1}{45}$

48. Um jogo de memória é formado por seis cartas, conforme as figuras que seguem:



Após embaralhar as cartas e virar as suas faces para baixo, o jogador deve buscar as cartas iguais, virando exatamente duas. A probabilidade de ele retirar, ao acaso, duas cartas iguais na primeira tentativa é de

- a) $\frac{1}{2}$ b) $\frac{1}{3}$ c) $\frac{1}{4}$ d) $\frac{1}{5}$ e) $\frac{1}{6}$

49.(UNIRIO) As probabilidades de três jogadores marcarem um gol cobrando um pênalti são, respectivamente, $\frac{1}{2}$, $\frac{2}{5}$ e $\frac{5}{6}$. 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%

50. Para ter acesso a um determinado programa de computador o usuário deve digitar uma senha composta por 5 letras distintas. Supondo que o usuário saiba quais são essas quatro letras, mas não saiba a ordem correta em que deve ser digitada, qual a probabilidade de esse usuário conseguir acesso ao programa numa única tentativa?

- a) $\frac{1}{4}$ b) $\frac{1}{12}$
c) $\frac{1}{16}$ d) $\frac{1}{32}$ e) $\frac{1}{256}$

Anotações

[illegible]



Currículo dos professores-autores

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 IFPR, 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, lecionando Estatística e Matemática Aplicada. 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.

Roberto José Medeiros Junior

É Licenciado e Bacharel em Matemática pela Universidade Tuiuti do Paraná (1999), Especialista em Educação Matemática com ênfase em Tecnologias pela Universidade Tuiuti do Paraná (2001), Especialista em Educação à Distância (Tutoria a Distância) – EaD/FACINTER (2007) tem Mestrado em Educação Matemática pela Universidade Federal do Paraná (2007). Entre os anos de 1996 e 2008, atuou como professor de Matemática do Ensino Fundamental ao Médio da rede pública e privada, e desde 2003 vem atuando como professor no Ensino Superior, nos cursos de Licenciatura em Matemática, Física e Pedagogia, na modalidade presencial e a distância em instituições públicas e privadas com as disciplinas de Cálculo, Estruturas Algébricas, Estatística e Matemática Financeira, e entre os anos de 2003 e 2005 atuou como professor de Metodologia, Prática de Ensino e Estágio Supervisionado em Matemática na Universidade Federal do Paraná, nos cursos de Licenciatura em Matemática, Física e Pedagogia. Atualmente é professor de Matemática em regime de Dedicação Exclusiva do Instituto Federal do Paraná na modalidade presencial e a distância. É um dos autores do Livro Didático Público de Matemática para o Ensino Médio do Estado do Paraná, e é também autor de livros para a formação continuada do Centro Interdisciplinar de Formação Continuada de Professores (CINFOP), da Universidade Federal do Paraná. Prestador de serviços como assessor pedagógico em Educação Matemática para as escolas públicas (municipal e estadual) e privadas.

