UNIVERSIDADE FEDERAL DO TOCANTINS
CÂMPUS UNIVERSITÁRIO DE PALMAS
PROGRAMA DE MESTRADO PROFISSIONAL EM MATEMÁTICA
EM REDE NACIONAL – PROFMAT

Jairo Rodrigues Barros

Estatística no Ensino Médio: uma proposta metodológica de ensino

Palmas-TO
2015

Jairo Rodrigues Barros

Estatística no Ensino Médio: uma proposta metodológica de ensino

Dissertação apresentada ao Programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT), coordenado pela Sociedade Brasileira de Matemática e ofertado pela Universidade Federal do Tocantins, sob a orientação do Prof. Msc. Gilmar Pires Novaes, como requisito parcial para a obtenção do título de Mestre em Matemática.

Palmas-TO
2015

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal do Tocantins

B277e Barros, Jairo Rodrigues .

Estatística no Ensino Médio: uma proposta metodológica de ensino. / Jairo Rodrigues Barros. – Palmas, TO, 2015.

79 f.

Dissertação (Mestrado Profissional) - Universidade Federal do Tocantins – Câmpus Universitário de Palmas - Curso de Pós-Graduação (Mestrado) Profissional em Matemática, 2015.

Orientador: Gilmar Pires Novaes

1. Livros didáticos. 2. Estatística. 3. Enem. 4. Proposta de ensino. I. Título

CDD 510

JAIRO RODRIGUES BARROS

Estatística no Ensino Médio: uma proposta metodológica de ensino

Dissertação apresentada ao programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT), coordenado pela Sociedade Brasileira de Matemática e ofertado pela Universidade Federal do Tocantins, sob a orientação do prof. Msc. Gilmar Pires Novaes, como requisito parcial para a obtenção do título de Mestre em Matemática.

Aprovada em 05 de junho de 2015.

BANCA EXAMINADORA:

Prof. Dr. Pedro Alexandre da Cruz (Presidente – UFT)

Prof. Dr. Chrystian de Assis Siqueira (UFT)

Profª. Drª. Lady Sakay (UNIRG)

Dedico este trabalho a toda a minha família que sempre esteve, e estará ao meu lado, e a todos os professores que, em seu dia-a-dia, lutam por uma educação de qualidade.

Agradecimentos

Meu grandessíssimo agradecimento a Deus que assume todo o senhorio em minha vida. Eternamente minha gratidão por sempre estar ao meu lado, com cuidado, carinho e amor a todo instante. A minha família, em especial a minha esposa Jéssica Kássia, que se orgulha de mim, que me acompanhou em toda essa trajetória tão árdua em minha vida. As meus queridíssimos filhos Bianca Nicole e Bernardo, por fazerem parte da minha vida (um beijo no coração de cada um!). Aos meus pais Gilberto e Tereza, que sempre me deram atenção especial por toda a vida. Aos Mestres, aqueles a quem tenho como referencial, de professor, de educador, podendo aqui com muito aprazimento lembrar-me de Gilmar Pires Novaes por tal exemplo e muito mais. Foi muito honroso para mim tê-lo como orientador deste trabalho. Meu muito obrigado!

Resumo

O presente trabalho se constitui de uma avaliação do uso da estatística no ensino médio, nas diferentes áreas do conhecimento, bem como a maneira como este conteúdo é tratado nos livros didáticos de matemática do PNLD atual, e aplicado no Exame nacional do Ensino Médio (ENEM). O mesmo apresenta uma proposta metodológica de ensino, baseada nas concepções de Lima[9], estruturada em três componentes básicas: conceituação, manipulação e aplicação da estatística no contexto social. Foi possível perceber que a estatística é de fundamental importância no cotidiano. No que diz respeito aos livros didáticos, percebemos muitos problemas relacionados, em especial, na conceituação de alguns termos estatísticos, e que de certa forma interferem negativamente na manipulação de várias atividades propostas. A aplicação está bastante presente nos materiais analisados, justificando a estatística como um dos ramos da matemática mais aplicáveis. Em relação a sua aplicação no Exame Nacional do Ensino Médio verificamos que os itens apresentados pelo mesmo são de fácil interpretação, e consequentemente, de fácil solução. Esperamos que com este trabalho o professor tenha uma ferramenta norteadora, dandolhe assim mais dinamicidade e objetividade na elaboração e execução de suas aulas.

Palavras-chave: Estatística. Proposta de ensino. Livros didáticos. ENEM.

Abstract

This study is an evaluation of the use of statistics in school, in different areas of knowledge and how that content is treated in textbooks of mathematics of the current PNLD and applied in the National Secondary Education Examination (ENEM). The same presents a methodological teaching proposal, based on the concepts of Elon Lages Lima cite lima, divided into three basic components: definition, manipulation and application of statistics in the social context. It was revealed that the statistic is of fundamental importance in everyday life. Regarding the textbooks, we see many problems, especially in the conceptualization of some statistical analysis, and that somehow interfere negatively in handling various activities proposed. The application is very present in the analyzed materials, justifying statistics as one of the relevant branches of mathematics. In relation to its application in the National Secondary Education Examination found that the items presented by the same are easy to interpret, and therefore an easy solution. We hope that with this work the teacher has a policy instrument, giving you more dynamism and objectivity in the development and implementation of their classes.

Keywords: Statistics. Proposal for education. Textbooks. ENEM

Lista de Figuras

Lista de Tabelas

Sumário

1 INTRODUÇÃO

Com as reais necessidades do mundo moderno, o novo Parâmetro Curricular Nacional para o Ensino Médio (PCNEM) [17], orientado pela Lei de Diretrizes e Bases da Educação (LDB), estabelece o Ensino Médio como uma etapa conclusiva da Educação Básica, e não só mais como uma etapa de preparação para outra fase escolar ou exercício profissional. E isso desafia os educadores a por em prática novas propostas que superem o antigo Ensino Médio, organizadas em duas etapas: a pré-universitária e a profissionalizante. O novo Ensino Médio deixa de ser estritamente preparatório para o Ensino Superior e Profissionalizante e passa a assumir necessariamente a responsabilidade de completar a Educação Básica. Isso significa preparar o aluno para a vida, qualificar para a cidadania e capacitar para o aprendizado permanente.

No novo Ensino Médio, a Matemática deve ser vista como uma parcela do conhecimento humano essencial para a formação do estudante, que contribua para uma visão de mundo, para ler e interpretar a realidade e para desenvolver capacidades que lhe serão exigidas ao longo da sua vida social e profissional.

A Matemática, segundo o PCNEM [17], é dividida em três eixos ou temas estruturadores: Álgebra, Geometria e Medidas e Análise de Dados. É precisamente no terceiro eixo em que se localiza o nosso objeto de estudo: a Estatística. A Estatística vem sendo usada com mais frequência em nosso dia a dia. Analisar a intenção de voto em uma eleição, antes da eleição, ou o possível êxito do lançamento de um produto no mercado, antes da sua fabricação, por exemplo, expõem as necessidades dos estudantes entenderem as reais aplicabilidades e utilidades da Estatística.

Conforme Nunes [14], a Estatística é uma ferramenta (ou método) que nos ajuda a interpretar e analisar grandes conjuntos de números. É, pois, a ciência da análise de dados. Ensina-nos como os dados podem ser coletados, organizados e analisados, e como podemos tirar conclusões a partir desses dados.

Para Nunes [14], sem a Estatística seria impossível efetuar sondagens políticas, apresentar os números de desempregos, efetuar o controle de qualidade dos bens de consumo, medir os níveis de audiência dos programas de televisão ou efetuar o planejamento de campanhas de marketing, dentre outros. Por outras palavras, o termo Estatística pode ser apresentado como um conjunto de instrumentos que podem ser usados para coletar, classificar, apresentar, interpretar conjuntos de dados numéricos e predizer informações sobre eles, no mundo da Matemática. Assim é que, associando as tecnologias de comunicação e informação ao componente curricular Estatística, proporcionou-se a criação de uma metodologia voltada ao Ensino Médio.

A sociedade moderna exige cada vez mais domínio da Estatística, para que o indivíduo possa desenvolver suas capacidades e orientar-se em seu mundo. A Estatística deve ser concebida como uma maneira de pensar, uma maneira de proporcionar uma consciência quantitativa dos fenômenos socioeconômicos. [1]

Uma das grandes competências do ensino da Matemática diz respeito à contextualização sociocultural, e a Estatística pode ser vista como um conjunto de ideias e procedimentos que podem ser aplicados a várias questões do mundo real, facilitando, assim, a interação aluno-Matemática-cotidiano.

A proposta de ensino e aprendizagem sugerida neste trabalho, é fundamentada em três componentes básicas, conforme Lima [9]: conceituação, manipulação e aplicação. A mesma visa subsidiar o trabalho do professor em sala, no que diz respeito a Estatística, propondo a ele mais objetividade em suas aulas, fornecendo-lhes para isso a estrutura metodológica necessária. Através dos conceitos estatísticos aplicados à situações cotidiana, fazer com que o aluno perceba importância da Estatística dentro do contexto social e justifiquem a relaçao aluno-Matemática-cotidiano.

O presente trabalho se constitui de uma avaliação do uso da Estatística no ensino médio, nas diferentes áreas do conhecimento, bem como a maneira como este conteúdo é tratado nos livros didáticos de matemática do PNLD atual, e aplicado no Exame nacional do Ensino Médio (ENEM).

Partindo de um breve contexto histórico, mostraremos a evolução do conhecimento estatístico, com o objetivo de situar o leitor cronologicamente, e assim dá sentido ao conteúdo abordado. Originado deste contexto histórico apresentaremos um embasamento teórico básico, fundamento em três componentes proposta por Lima [9]: conceituação, manipulação e aplicação, visando que o leitor entenda a importância de cada termo estatístico. Procurando inserir o trabalho no âmbito no que é sugerido na PNLD[1] atual, realizaremos uma análise qualitativa de algumas obras adotadas pelo plano, fundamentado pelo livro "Exame de textos: Análise de livros de Matemática para o Ensino Médio"[9]. Por fim, vendo o ENEM[2] como umas principais formas de democratização de acesso ao ensino superior, através da unificação dos vestibulares de intituições públicas e do programa ProUni[3] onde o governo distribui bolsas parciais ou integrais em instuições de ensino superior privada. Faremos uma análise do eixo de conhecimento" matemática e suas tecnologias"e em seguida uma relação entre o conteúdo abordado e questões retiradas do exame, por meio de soluções práticas e explicativas.

[1]Plano Nacional do Livro Didático
[2]Exame Nacional do Ensino Médio
[3]Programa Universidade para todos

2 BREVE HISTÓRICO

A Estatística é um ramo da Matemática aplicada que surgiu dos estudos de algumas das questões de estado e governo. Situações ocasionais tais como número de habitantes, quantidade de óbitos e nascimentos, quantidades produzidas e quantitativos das riquezas formaram os primórdios dos problemas que deram início ao pensamento estatístico. Daí o nome Estatística ter se originado do termo latino status.

Algumas referências bíblicas nos remetem à utilização da Estatística cerca de 1700 a.C. Por exemplo, o livro de Números, em seu Capítulo 1, Versículo 2, relata: "Tomai a soma de toda a congregação dos filhos de Israel, segundo as suas famílias, segundo as casas de seus pais, conforme o número dos nomes de todo homem, cabeça por cabeça".

As tentativas de enumeração de indivíduos ou de bens seguiram com os grandes impérios da Antiguidade, cujas estruturas administrativas eram fortes: preocupados em gerir e administrar seu império do melhor modo, os poderes centrais procuraram conhecer melhor sua extensão territorial e o número de seus súditos. Foi assim que as civilizações egípcia, mesopotâmica e chinesa, bem como antes delas, a civilização dos sumérios (5000 a 2000 a.C.), realizavam pesquisas censitárias, das quais alguns traços chegaram até nós. Esses recenseamentos tinham como objetivo comum, antes de tudo, responder à necessidade da administração do império: responder às necessidades de mão-de-obra em vista da construção das grandes pirâmides; responder às preocupações fiscais.

Para Nogueira [13], na idade Média, no ano de 1086, Guilherme I de Inglaterra (1028-1087) (chamado de Guilherme, o Conquistador), invasor normando da Inglaterra, mandou realizar um levantamento detalhado das propriedades rurais dos conquistados anglo-saxões, para se inteirar de suas riquezas. Desse trabalho resultou o Doomesday

Figura 1: Escribas recolhendo impostos

Book (livro do dia do juízo), uma importante fonte de dados sobre Estatística descritiva da Inglaterra Medieval. No século XVI, mais precisamente em 1583, foram realizados levantamentos de dados estatísticos para o governo italiano por Francesco Sansovino (1521-1586) e Giovanni Botero (1544-1617), denominado "Del Governo et Administrazoni de

Diversi Regni e Republiche”, o que evidencia o uso da Estatística pelo Estado.

Inicialmente, no século XVI, pensada pelos ingleses como uma ciência política, a Estatística destinava-se a descrever características de um país, tais como população, área, riquezas e recursos naturais. Desse papel histórico, origina-se a sua função de caracterização numérica de uma série de informações populacionais. Com essa abordagem, o termo Estatística é utilizado no plural, como as estatísticas de saúde, as estatísticas de mortalidade, as estatísticas do registro civil, dentre outras.

A Estatística, até então ligada à Filosofia, desvincula-se desta e passa a ser uma disciplina autônoma. John Graunt (1620-1674) publicou o livro “Natural and Political Observations Mentioned in a Following Index, and made upon the Bills of Mortality”(“Observações Naturais e Políticas Mencionado em um Índice de Seguidores, e feito em cima das Letras de Mortalidade”), que foi baseado nas razões e proporções de fatos vitais, nos quais se observam uma realidade estatística em um grande números de dados.

A Estatística, vista enquanto ciência, só ocorreu a partir do século XVIII, nos registros do alemão Gottfried Achenwall (1719-1772), ainda como catalogação não regular de dados. Os modelos estatísticos, enquanto modelos matemáticos aplicados, reúnem características de precisão na linguagem, adequadas ao ambiente de informações rápidas.

A contribuição de Sir Ronald Aylmer Fisher (1890-1962) para a Estatística Moderna é, sem dúvida, a mais importante e decisiva de todas. Formado em Astronomia pela Universidade de Cambridge, em 1912, foi o fundador do célebre Statistical Laboratory da prestigiosa Estação Agronômica de Rothamsted, contribuindo enormemente, tanto para o desenvolvimento da Estatística quanto da Genética. Ele apresentou os princípios de planejamento de experimentos, introduzindo os conceitos de aleatorização e da Análise da Variância, procedimentos muito usados atualmente em Estatística.

A necessidade de expressar o grau de incerteza na ocorrência dos experimentos e de explicar o fato de duas experiências iguais poderem ter resultados diferentes conduz ao reconhecimento da racionalidade probabilística em eventos da natureza. A pesquisa em Probabilidade, no século XVIII, culmina com o notável trabalho de Pierre- Simon de Laplace (1749-1827), “Theorie Analitique de Probabilités”. À luz da concepção do cientificismo, rapidamente amplia- se o domínio de abrangência do cálculo probabilístico. Este se torna indispensável para lidar com dados relativos a temas de interesse social e econômico, tais como administração das finanças públicas, saúde coletiva, conduta de eleições e seguro de vida. Surgem as primeiras ideias do positivismo, e o Marquês de Condorcet (1743-1794) propõe uma ciência natural da sociedade, isto é, uma matemática social baseada no cálculo das probabilidades.

A inserção da Estatística no currículo educacional começou a ser discutida em 1948. Nesse debate, estimulado pela UNESCO, que incentivou as pesquisas sobre a necessidade da estatística nas suas atribuições cotidianas, foram criados comitês e associações a fim de disseminar a educação estatística.

Em meados dos anos 70, foi criado o ISI (Instituto de Educação Estatística), que tinha como objetivo incentivar e ampliar as pesquisas sobre educação estatística. A partir daí, novos meios foram criados para desenvolver o conhecimento estatístico na educação básica, dentre eles:

- produção de livros e textos nos quais a Estatística fazia parte do cotidiano do aluno;
- produção de material didático para auxiliar os professores;
- organização de encontros de interessados na área de educação estatística.

Percebendo a evolução da ideia do conhecimento estatístico, sob as perspectivas de que:

- a Probabilidade e a Estatística estavam relacionadas a quase todas as atividades da sociedade moderna;
- os estudantes usariam o conhecimento de Probabilidade e Estatística em suas profissões futuras;
- A inserção da Estatística produzirá um efeito estimulante, por ser um ramo dinâmico da Matemática, e suas aplicações seriam evidentes.

Em 1970, na 1ª conferência Comprehensive School Mathematics, a Estatística foi inserida no currículo da educação básica de alguns países.

2.1 A Estatística no Brasil

No Brasil, a Estatística tem sua história associada à história do Instituto Brasileiro de Geografia e Estatística (IBGE), cujas raízes foram implantadas ainda durante o Império. Quem primeiro coordenou e sistematizou atividades ligadas a levantamentos censitários foi a Diretoria Geral de Estatística, criada em agosto de 1872, data do "primeiro Recenseamento Geral do Império do Brasil". No período anterior a esta data (1750-1872), a Coroa Portuguesa era quem determinava levantamentos populacionais, realizados precariamente, com o objetivo maior de "conhecer a população livre e adulta apta a ser usada na defesa do território".

A partir da segunda metade do século XIX, esses levantamentos passaram a ser realizados por juízes de paz e chefes de polícia dos municípios, mais com fins eleitoreiros, constituindo-se as paróquias a base para as informações. Com o advento da República, a produção das estatísticas dispersou-se nas esferas Municipal, Estadual e Federal, quase impossibilitando a unificação dos resultados e dificultando as análises estatísticas.

Ainda de acordo com o Calendário, foi criado, em 1907, o Conselho Superior de Estatística, com vistas à padronização de conceitos e apuração de resultados em todo o território nacional.

Em 1934, foi criado o Instituto Nacional de Estatística, que só passou a existir de fato em 1936, mudando em 1938 para Instituto Brasileiro de Geografia e Estatística

(IBGE), quando os serviços geográficos foram a ele vinculados. Foi a partir de 1940 que se iniciaram os “modernos censos decenais”. Antes disso, ocorreram os de 1872, 1890, 1900 e 1920.[19]

Sob o preceito de que a Estatística estava em muitos eventos do cotidiano, ela começou a fazer parte do currículo da Educação Básica brasileira em 1997, anos após a inclusão em outros países, com a criação dos Parâmetros curriculares nacionais (PCN's), os quais estabeleciam habilidades a serem alcançadas pelos alunos da Educação Básica.

3 ALGUMAS NOÇÕES DE ESTATÍSTICA

As definições apresentadas neste capítulo são fundamentadas nas obras: "Probabilidade e Estatística: Um curso introdutório"[15] e "Estatística básica"' [12].

A Estatística é um dos ramos da Matemática que mais usamos em nosso cotidiano:

- uma empresa automobilística realiza pesquisa para saber a preferência do consumidor em assuntos tais como: cor, tipo de combustível, potência do motor, etc.
- pesquisas eleitorais são realizadas para saber a preferência por candidatos.
- o IBGE realiza pesquisas para senso demográfico, índices de preços ao consumidor (IPCA), dentre outras, que interferem diretamente na vida da população.
- emissoras de televisão utilizam pesquisas que mostram preferências dos espectadores para organizar sua programação.
- pesquisas são realizadas com atletas para mostrar e melhorar seu desempenho em partidas.

A realização de uma pesquisa envolve várias etapas, tais como: escolha da amostra, coleta e organização dos dados colhidos, resumo e apresentação dos dados, e interpretação dos resultados.

Mas, afinal, o que é Estatística? Para Marcos [12], a Estatística é um conjunto de técnicas que permite, de forma sistemática, organizar, descrever, analisar e interpretar dados oriundos de experimentos realizados em qualquer área do conhecimento, sendo dado um conjunto de valores numéricos ou não.

Podemos dividir a Estatística em três partes:

- Estatística descritiva;
- Probabilidade estatística;
- Inferência estatística.

A Estatística descritiva pode ser definida como um conjunto de técnicas destinadas a descrever e resumir os dados, afim de que se possam tirar conclusões a respeito de determinadas características de interesse.

A Probabilidade estatística é teoria da Matemática utilizada para se estudar as incertezas oriundas de fenômenos de caráter aleatório.

A Inferência estatística é o estudo de técnicas que possibilitam a extrapolação, a um grande conjunto de dados obtidos a partir de subconjuntos de valores, usualmente de dimensões muito menor.

Neste trabalho, abordaremos apenas a Estatística descritiva.

3.1 População

Denomina-se população ou universo estatístico a um conjunto constituído por elementos portadores de, pelo menos, uma característica em comum ou que têm uma mesma propriedade, sendo objetos de interesse para estudo.

A população pode ser dividida em:

- população finita: é aquela que pode ser quantificada numericamente (contada).
- população infinita: é aquela que não pode ser quantificada numericamente (contada).

Para Magno [15], o conceito de população é bem mais amplo que o uso dessa palavra. População pode ser constituída por pessoas, domicílios, peças de produção, cobaias ou qualquer outro elemento a ser investigado. A caracterização da população é realizada em função do problema a ser estudado. Para que haja uma clara definição das unidades que constituem a população, é necessária a especificação de três elementos: uma característica em comum, localização temporal e localização geográfica.

3.2 Amostra

Parte da população de nosso interesse em uma pesquisa estatística denomina-se amostra da população. Cada elemento que pertence à amostra denomina-se, por sua vez, indivíduo da amostra.

3.3 Variável

Em Estatística, uma variável é uma característica qualquer de interesse que associamos à população ou à amostra para ser estudada estatisticamente. São chamadas assim porque apresentam variação de elemento para elemento na população ou amostra de estudo. Iniciaremos nossos estudos com a seguinte situação:

Exemplo 3.1 *Um administrador da empresa J.A. representações resolveu realizar uma pesquisa socioeconômica para tentar obter propostas para realizações de atividades que melhorassem a qualidade de vida de seus funcionários. Os dados obtidos foram organizados conforme a tabela abaixo:*

Figura 2: Quadro de informações dos funcionários

Informações sobre estado civil, grau de instrução, nº de filhos, salários, idade e região de procedência dos funcionários empresa.

Nº	Estado civil	Grau de instrução	Nº de filhos	Salário (X sal. Mín)	Idade (anos)	Região de procedência
1	solteiro	Ensino fundamental	-	4,00	26	Interior
2	casado	Ensino fundamental	1	4,56	32	capital
3	casado	Ensino fundamental	2	5,25	36	capital
4	solteiro	Ensino médio	-	5,73	20	outra
5	solteiro	Ensino fundamental	-	6,26	40	outra
6	casado	Ensino fundamental	0	6,66	28	Interior
7	solteiro	Ensino fundamental	-	6,86	41	Interior
8	solteiro	Ensino fundamental	-	7,39	43	capital
9	casado	Ensino médio	1	7,59	34	capital
10	solteiro	Ensino médio	-	7,44	23	outra
11	casado	Ensino médio	2	8,12	33	Interior
12	solteiro	Ensino fundamental	-	8,46	27	capital
13	solteiro	Ensino médio	-	8,74	37	outra
14	casado	Ensino fundamental	3	8,95	44	outra
15	casado	Ensino médio	0	9,13	30	Interior
16	solteiro	Ensino médio	-	9,35	38	outra
17	casado	Ensino médio	1	9,77	31	outra
18	casado	Ensino fundamental	2	9,80	39	capital
19	solteiro	superior	-	10,53	25	Interior
20	solteiro	Ensino médio	-	10,76	37	Interior
21	casado	Ensino médio	1	11,06	30	outra
22	solteiro	Ensino médio	-	11,59	34	capital
23	solteiro	Ensino fundamental	-	12,00	41	outra
24	casado	superior	0	12,79	26	outra
25	casado	Ensino médio	2	13,23	32	Interior
26	casado	Ensino médio	2	13,60	35	outra
27	solteiro	Ensino fundamental	-	13,85	46	outra
28	casado	Ensino médio	0	14,69	29	Interior
29	casado	Ensino médio	5	14,71	40	Interior
30	casado	Ensino médio	2	15,99	35	capital
31	solteiro	superior	-	16,22	31	outra
32	casado	Ensino médio	1	16,61	36	Interior
33	casado	superior	3	17,26	43	capital
34	solteiro	superior	-	18,75	33	capital
35	casado	Ensino médio	2	19,40	48	capital
36	casado	superior	3	23,30	42	Interior

Fonte: Dados hipotéticos.

De modo geral, para cada elemento investigado em uma pesquisa, tem-se associado um (ou mais de um) resultado(s) correspondendo a realizações de uma característica. Entendem-se por realizações as respostas possíveis para determinado tipo de variável.

À variável estado civil, por exemplo, pode-se associar as realizações solteiro, casado, viúvo, etc.

Exemplo 3.2 *Na figura 2, verifique todas as variáveis pesquisadas:*

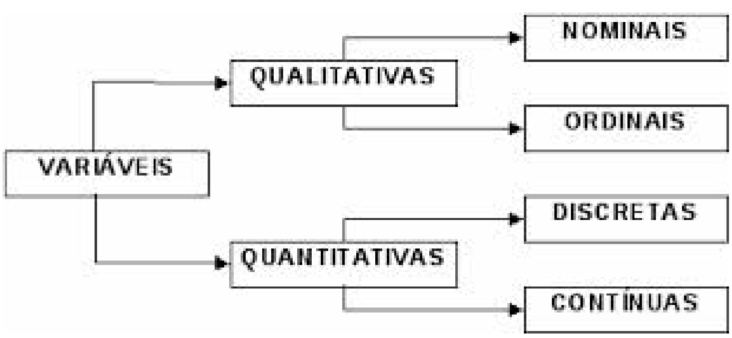
Solução:

Estado civil, grau de instrução, número de filhos, salário, idade e região de procedência.

Algumas variáveis, tais como estado civil, grau de instrução e região de procedência, apresentam como possíveis realizações (respostas) uma qualidade do indivíduo pesquisado. A esse tipo de variável denominamos variável qualitativa. Já as variáveis como números de filhos, salário e idade têm como possíveis realizações valores numéricos (contagem ou mensuração). A esse tipo de variável denominamos variável quantitativa.

As variáveis qualitativas subdividem-se em qualitativa nominal, para a qual não existe ordenação alguma nas possíveis realizações (por exemplo, estado civil), e qualitativa ordinal, para a qual existe uma ordem na obtenção dos seus resultados (por exemplo, grau de instrução). De forma análoga, as variáveis quantitativas subdividem-se em discreta, cujas possíveis realizações constituem um conjunto finito ou enumerável de números, que resultam frequentemente de uma contagem (por exemplo, o número de filhos(0,1,2,3...)), e contínua, cujas possíveis realizações pertencem a um intervalo de números reais, que resultam de uma mensuração (por exemplo, a massa de um corpo e a altura).

Figura 3: Classificação das variáveis



Exemplo 3.3 *Usando a figura 2 dada anteriormente, classifique cada uma das variáveis pesquisadas.*

- Estado civil - qualitativa nominal;
- Grau de instrução - qualitativa ordinal, pois, para obter cada grau de formação, é necessário seguir uma ordem: fundamental primeiro; em seguida, médio e, posteriormente, superior;
- Número de filhos - quantitativa discreta;
- Salário - quantitativa contínua;
- Idade - quantitativa discreta;
- Região de procedência - qualitativa nominal.

3.4 Distribuição de frequência

O principal interesse ao estudar uma variável é conhecer o seu comportamento analisando a ocorrência de suas possíveis realizações. Dessa forma, estudaremos uma maneira de dispor um conjunto de realizações para termos ideia global sobre elas, ou seja, de sua distribuição.

Existem alguns termos associados à distribuição de frequência que nos ajudam a ter uma visão sobre o comportamento da variável:

- frequência absoluta (n_i) - número de ocorrências de cada realização da variável na pesquisa.
- frequência relativa (f_i) - o valor decimal obtido para cada uma das realizações. Obtemos a frequência relativa por meio da razão $\dfrac{n_i}{\displaystyle\sum_{i=1}^{n} n_i}$.
- porcentagem ($100 \cdot f_i$) - valor percentual de cada realização da pesquisa. Obtemos a porcentagem por meio do produto $100 \cdot f_i$.

Exemplo 3.4 *Usando a figura 2, construa a tabela de frequência referente ao grau de instrução dos funcionários.*

1. Inicialmente, observamos que a tabela é organizada contendo a variável, frequência absoluta(freq. abs. (n_i)), frequência relativa (freq. rel. (f_i)) e porcentagem ($100 \cdot f_i$), da seguinte forma:

2. Calculemos os valores das frequências absoluta e relativa, e a porcentagem:

Tabela 1: Grau de instrução dos funcionários

Grau de instrução	Freq. absoluta (n_i)	Freq relativa (f_i)	Porcentagem ($100 \cdot f_i$)
Fundamental	n_1	f_1	$100.f_1$
Médio	n_2	f_2	$100 \cdot f_2$
Superior	n_3	f_3	$100 \cdot f_3$
Total	$\sum_{i=1}^{3} n_i$	$\sum_{i=1}^{3} f_i$	$\sum_{i=1}^{3} 100 \cdot f_i$

- Frequência absoluta:

$$n_1 = 12,$$
$$n_2 = 18,$$
$$n_3 = 6.$$

$$\sum_{i=1}^{3} n_i = n_1 + n_2 + n_3 = 12 + 18 + 6 = 36.$$

- Frequência relativa:

$$f_1 = \frac{n_1}{\sum_{i=1}^{3} n_i} = \frac{12}{36} = 0,334,$$

$$f_2 = \frac{n_2}{\sum_{i=1}^{3} n_i} = \frac{18}{36} = 0,500,$$

$$f_3 = \frac{n_3}{\sum_{i=1}^{3} n_i} = \frac{6}{36} = 0,166.$$

$$\sum_{i=1}^{3} f_i = f_1 + f_2 + f_3 = 0,334 + 0,500 + 0,166 = 1.$$

- Porcentagem:

$$100 \cdot f_1 = 100.0,334 = 33,4\%,$$
$$100 \cdot f_2 = 100.0,500 = 50\%,$$
$$100 \cdot f_3 = 100.0,166 = 16,4\%,$$

$$\sum_{i=1}^{3} 100 \cdot f_i = 100 \cdot f_1 + 100 \cdot f_2 + 100 \cdot f_3 = 33,4 + 50 + 16,6 = 100\%.$$

3. Por fim, preenchemos a tabela com os valores obtidos:

Tabela 2: Grau de instrução dos funcionários

Grau de instrução	Freq. absoluta (n_i)	Freq. relativa (f_i)	Porcentagem ($100 \cdot f_i$)
Fundamental	12	0,334	33,4
Médio	18	0,5	50
Superior	6	0,164	16,4
Total	36	1	100%

Exemplo 3.5 *Usando a figura 2 dada anteriormente, construa a tabela de frequências para a variável número de filhos dos funcionários casados.*

1. Frequência absoluta:

$$n_1 = 4,$$
$$n_2 = 5,$$
$$n_3 = 7,$$
$$n_4 = 3,$$
$$n_5 = 0,$$
$$n_6 = 1.$$

$$\sum_{i=1}^{6} n_i = n_1 + n_2 + n_3 + n_4 + n5 + n_6 = 4 + 5 + 7 + 3 + 0 + 1 = 20.$$

2. Frequência relativa:

$$f_1 = \frac{n_1}{\sum_{i=1}^{6} n_i} = \frac{4}{20} = 0,2,$$

$$f_2 = \frac{n_2}{\sum_{i=1}^{6} n_i} = \frac{5}{20} = 0,25,$$

$$f_3 = \frac{n_3}{\sum_{i=1}^{6} n_i} = \frac{7}{20} = 0,35,$$

$$f_4 = \frac{n_4}{\sum_{i=1}^{6} n_i} = \frac{3}{20} = 0,15,$$

$$f_5 = \frac{n_5}{\sum_{i=1}^{6} n_i} = \frac{0}{20} = 0,$$

$$f_6 = \frac{n_6}{\sum_{i=1}^{6} n_i} = \frac{1}{20} = 0,05.$$

$$\begin{aligned}\sum_{i=1}^{6} f_i &= f_1 + f_2 + f_3 + f_4 + f_5 + f_6 \\ &= 0,2 + 0,25 + 0,35 + 0,15 + 0 + 0,05 \\ &= 1.\end{aligned}$$

3. Porcentagem:

$100 \cdot f_1 = 100.0,2 = 20\%,$
$100 \cdot f_2 = 100.0,25 = 25\%,$
$100 \cdot f_3 = 100.0,35 = 35\%,$
$100 \cdot f_4 = 100.0,15 = 15\%,$
$100 \cdot f_5 = 100.0,0 = 0,$
$10 \cdot f_6 = 100.0,05 = 5\%.$

$$\begin{aligned}\sum_{i=1}^{6} 100 \cdot f_i &= 100 \cdot f_1 + 100 \cdot f_2 + 100 \cdot f_3 + 100 \cdot f_4 + 100 \cdot f_5 + 100 \cdot f_6 \\ &= 20 + 25 + 35 + 15 + 0 + 5 \\ &= 100\%.\end{aligned}$$

4. Tabela:

Tabela 3: Número de filhos dos funcionários

Número de filhos	Freq. absoluta (n_i)	Freq. relativa (f_i)	Porcentagem($100 \cdot f_i$)
0	4	0,2	20
1	5	0,25	25
2	7	0,35	35
3	3	0,15	15
4	0	0	0
5	1	0,05	5
Total	20	1	100%

Quando trabalharmos com variáveis quantitativas contínuas ou ainda com variáveis discretas que tomam muitos valores distintos, veremos que as informações serão apresentadas em uma tabela denominada tabela de classes, a qual tornará o trabalho mais organizado e informativo. Uma das noções estatísticas de que necessitamos para construção e organização da tabela de classes é o rol.

Definição 3.1 *Considere um conjunto* $X = (x_1, x_2, x_3, ..., x_n)$ *que contém as informações de uma pesquisa. Rol(R) desses dados é o conjunto* $R = (x_{(1)}, x_{(2)}, x_{(3)}, ..., x_{(n)})$ *tal que, para* $1 \leq i \leq n$*, tem-se* $x_{(i)} = x_{(j)}$*, para algum* $1 \leq j \leq n$ *e, além disso,*

$$x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$$

Ou seja, o rol é uma lista ordenada (de forma crescente ou decrescente) de todos os resultados das observações de algum fenômeno, mesmo que repetidos.

Exemplo 3.6 *A fim de melhorar a produção de leite bovino, um fazendeiro deseja suplementar a alimentação de suas matrizes leiteiras. Antes, contudo, ele observou que a quantidade produzida (em litros) por suas nove matrizes em um determinado dia foi o conjunto* $M = (m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8, m_9) = (8, 5, 7, 12, 8, 5, 5, 7, 10)$*. Obtenha o rol referente a esse conjunto.*

Solução:

De acordo com a definição de rol(R), basta ordenarmos (por exemplo, em forma crescente) os elementos do conjunto dado: $R = (m_2, m_6, m_7, m_3, m_8, m_1, m_5, m_4, m_9) = (5, 5, 5, 7, 7, 8, 8, 10, 12)$.

$$\begin{aligned} R &= (m_{(1)}, m_{(2)}, m_{(3)}, m_{(4)}, m_{(5)}, m_{(6)}, m_{(7)}, m_{(8)}, m_{(9)}) \\ R &= (5, 5, 5, 7, 7, 8, 8, 10, 12). \end{aligned}$$

3.4.1 Construção da tabela de classes

Obtido rol(R) da pesquisa, o próximo passo para a construção da tabela de classes é obter a amplitude total do conjunto.

Definição 3.2 *Amplitude de um conjunto de informações é a diferença entre a maior e a menor informação do conjunto, ou seja: dado um conjunto de informações* $X = (x_1, x_2, x_3, ..., x_n)$*, obtido o rol(R)=* $(x_{(1)}, x_{(2)}, x_{(3)}, ..., x_{(n)})$*, a amplitude total do conjunto será*

$$\Delta(M) = x_{(n)} - x_{(1)}.$$

Exemplo 3.7 *Considerando as informações do Exemplo 3.6, calcule a amplitude do conjunto mencionado.*

Solução:

Como $R = (5, 5, 5, 7, 7, 8, 8, 10, 12)$ e a amplitude é dada por $\Delta(M) = x_{(n)} - x_{(1)}$ temos que:

$$\Delta(M) = 12 - 5 \Rightarrow \Delta(M) = 7.$$

O próximo passo é a obtenção da quantidade de classes (K). É importante que a quantidade de classes da distribuição seja adequado. Se o número de classes for excessivamente pequeno, acarretará perda de detalhes, e pouca informação pode ser extraída da tabela. Por outro lado, se for utilizado um número excessivo de classes, haverá alguma classe com frequência nula ou muito pequena. Não existe uma fórmula exata para o número de classes, embora as seguintes sejam as mais usadas:

1. Regra de Sturges[4]

$$K \approx 1 + 3,3 \log n$$

2. para $n \leq 25$, consideramos $K = 5$; para $n > 25$, consideramos $K = \sqrt{n}$, em que n é o total de observações.

Por fim, calcularemos a amplitude do intervalo de classe (Δi), a qual nos informará a dimensão de cada intervalo.

$$\Delta i = \frac{\Delta(M)}{K}.$$

Organizaremos as classes da seguinte forma (o símbolo $\dashv$ representa o intervalo numérico $[a, b[$):
1ª classe : $x_1 \dashv x_1 + \Delta i$,
2ª classe : $x_1 + \Delta i \dashv x_1 + \Delta i + \Delta i = x_1 + 2 \cdot \Delta i$,
3ª classe : $x_1 + 2 \cdot \Delta i \dashv x_1 + 2 \cdot \Delta i + \Delta i = x_1 + 3 \cdot \Delta i$,
.
.
.
kª classe : $x_1 + (k-1) \cdot \Delta i \dashv x_1 + (k-1) \cdot \Delta i + \Delta i = x_1 + k \cdot \Delta i$.

Exemplo 3.8 *Usando a figura 2, construa a tabela de classes das variáveis:*

a) Salário

[4]Sturges (1882-1958): A regra proposta por Herbert Sturges em 1926 é uma regra prática sobre a quantidade de classes que deve ser considerada para ser incorporada em uma tabela de classe ou histograma.

Solução:

- Inicialmente, vamos obter o rol(R) das observações da variável:

$$R = (4; 4,56; 5,25; ...; 23,30).$$

- Amplitude do conjunto de observações (ΔM):

$\Delta M = x_{(n)} - x_{(1)} \Rightarrow \Delta(M) = 23,30 - 4 \Rightarrow \Delta M = 19,30.$

- Usando os dois métodos estudados, calculemos a quantidade de classes (K):

1. Regra de Sturges:

   $K \approx 1 + 3,3.\log n \Rightarrow K \approx 1 + 3,3.\log 36 \Rightarrow K \approx 1 + 5,13 \Rightarrow K \approx 6,13 \Rightarrow K \approx 6$ ($n = 36$: quantidade total de observações).

2. Método da raiz:

   $K = \sqrt{n} \Rightarrow K = \sqrt{36} \Rightarrow K = 6$ ($n = 36$: quantidade total de observações).

- calculemos o valor do intervalo de classe (Δi):

$$\Delta i = \frac{\Delta(M)}{K} \Rightarrow \Delta i \approx \frac{19,3}{6} \Rightarrow \Delta i \approx 3,3.$$

- Tabela de classes:

Tabela 4: Salário dos funcionários

Salário	Freq. absoluta (n_i)	Freq. relativa(f_i)	Porcentagem ($100 \cdot f_i$)
$4 \dashv 7,30$	7	0,19	20
$7,3 \dashv 10,6$	12	0,33	33
$10,6 \dashv 13,9$	8	0,22	22
$13,9 \dashv 17,2$	5	0,14	14
$17,2 \dashv 20,5$	3	0,09	9
$20,5 \dashv 23,8$	1	0,03	3
total	36	1	100%

b) Idade

Solução:

- Rol(R):

  $R = (20, 23, 25, 26, ..., 48)$.

- Amplitude do conjunto $\Delta(M)$:

  $\Delta(M) = \Delta(M) = x_{(n)} - x_{(1)} \Rightarrow \Delta(M) = 48 - 20 \Rightarrow \Delta(M) = 28.$

- Número de classes (K):

  1. Regra de Sturges:

     $K \approx 1 + 3,3 \log n \Rightarrow K \approx 1 + 3,3 \cdot log36 \Rightarrow K \approx 6.$

  2. Método da raiz:

     $K \approx \sqrt{n} \Rightarrow K \approx \sqrt{36} \Rightarrow K \approx 6.$

- Amplitude do intervalo de classe:

  $$\Delta i = \frac{\Delta(M)}{K} \Rightarrow \Delta i \approx \frac{28}{6} \Rightarrow \Delta i \approx 4,7.$$

- Tabela de Classes

Tabela 5: Idade dos Funcionários

Idade	Freq. absoluta (n_i)	Freq. relativa (f_i)	Porcentagem ($100 \cdot f_i$)
$20 \dashv 24,7$	2	0,055	5,5
$24,7 \dashv 29,4$	6	0,167	16,7
$29,4 \dashv 34,1$	10	0,277	27,7
$34,1 \dashv 38,8$	7	0,195	19,5
$38,8 \dashv 43,5$	8	0,223	22,3
$43,5 \dashv 48,2$	3	0,083	8,3
Total	36	1	100%

3.5 Medidas resumo

Quando lidamos com um conjunto finito grande, muitas vezes necessitamos de uma medida que resuma os dados nele contidos. Por exemplo, uma escola, diante de um conjunto de notas obtidas em um ano letivo por uma determinada classe, procura resumir tais notas a um único valor para que aplique o conceito aprovado ou reprovado. Uma panificadora deseja saber quantos pães deve produzir diariamente, de modo que procura resumir os valores obtidos no mês anterior.

Para Magno [15], medidas resumo são números que associamos aos conjuntos e que, de alguma maneira, descrevem o comportamento global dos seus elementos.

As medidas resumo são classificadas em dois grupos: medidas de tendência central (medidas de posição) e medidas de dispersão.

3.6 Medidas de tendência central

As medidas de tendência central são valores que, de certa forma e de maneira condensada, trazem consigo informações contidas nos dados estatísticos, sejam eles populacionais ou amostrais. Segundo Salsa [18], as medidas de tendência central funcionam como uma espécie de medidas-resumo, pois nos passam a ideia, digamos, do comportamento geral das observações estudadas. Podemos dizer ainda que elas são como valores de referência, em torno dos quais os outros se distribuem. Quando estão associadas aos dados populacionais, são denominadas parâmetros; quando são calculadas a partir de amostras, são denominadas estatísticas. Essa diferença ocorre porque os parâmetros são valores constantes (fixos), pois são calculados a partir de todos os dados de certo conjunto (a população de interesse). Porém, se trabalhamos com amostras, as medidas estatísticas obtidas variam de acordo com as observações que foram selecionadas. Por isso, elas não são valores fixos, pois dependem dos elementos da amostra particular que foi escolhida.

As medidas de posição vão nos fornecer medidas por meio das quais possamos caracterizar o comportamento dos elementos do conjunto de dados. Elas são classificadas em: média aritmética, mediana e moda.

3.6.1 Média aritmética

Definição 3.3 *Considere X um conjunto de cardinalidade finita, isto é, $X = (x_1, x_2, x_3, ..., x_n)$. Dizemos que a média aritmética dos elementos de X é o número $\overline{x}$ definido por*

$$\overline{x} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Exemplo 3.9 *Considerando o Exemplo 3.6, vamos calcular a média aritmética de M:*

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\begin{aligned}
&= \frac{\sum_{i=1}^{9} x_i}{9}\\
&= \frac{5+5+5+7+7+8+8+10+12}{9}\\
&= \frac{67}{9}\\
\overline{x} &= 7,444...
\end{aligned}$$

Portanto, a média diária de produção é, aproximadamente, $7,44$ litros.

Definição 3.4 *Considere um conjunto de cardinalidade finita*[5] *que representa a variedade de informações em um processo de coleta de dados, isto é,* $X = (x_1, x_2, x_3, ..., x_n)$*, em que* x_i *tem uma frequência* n_i *no rol que contém as informações do processo (*n_i *representa o peso de cada realização). Denominamos média aritmética ponderada:*

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i}.$$

Exemplo 3.10 *Um professor de Matemática resolveu dividir sua nota bimestral em três avaliações:*

1. *prova - peso 5;*
2. *trabalho - peso 3;*
3. *participação nas atividades - peso 2.*

Sabendo que todas as avaliações eram pontuadas de zero (0) a dez (10), calcule a nota final de cada aluno abaixo:

a) João: 4 na prova, 2 no trabalho e 2 nas participações.

Solução:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{3} x_i n_i}{\sum_{i=1}^{3} n_i} = \frac{4 \cdot 5 + 2 \cdot 3 + 2 \cdot 2}{5+3+2} = \frac{30}{10} = 3.$$

[5]Conjunto com quantidade finita de elementos.

Portanto, a média final de João é igual a 3 pontos.

b) Maria: 8 na prova, 9 no trabalho e 7 nas participações.

Solução:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{3} x_i n_i}{\sum_{i=1}^{3} n_i} = \frac{8 \cdot 5 + 9 \cdot 3 + 7 \cdot 2}{5 + 3 + 2} = \frac{81}{10} = 8,1.$$

Portanto, a média final de Maria é igual a 8,1 pontos.

3.6.2 Mediana

Definição 3.5 *Dizemos que a mediana de um rol de dados* $X = (x_{(1)}, x_{(2)}, x_{(3)}, ..., x_{(n)})$ *é a medida que centraliza o rol, isto é,* x_{med} *tal que:*

$$x_{med} \leq 50\% \quad dos \quad dados \quad ; \quad x_{med} \geq 50\% \quad dos \quad dados.$$

Como consequência temos,

$$x_{med} = \begin{cases} x_{\frac{n+1}{2}}, & se \quad n \quad ímpar \\ \dfrac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2}, & se \quad n \quad par \end{cases}$$

como uma mediana para o rol.

A mediana é o valor em que x_{med} representa sua posição no rol. Logo, o valor da mediana sé maior que ou igual a 50% dos dados e menor que ou igual aos outros 50% dos dados.

Exemplo 3.11 *Calcule a mediana do rol(R)* $= (5, 5, 5, 7, 7, 8, 8, 10, 12)$ *referente ao Exemplo 3.6.*

Como $n = 9$ é ímpar, temos que,

$$x_{med} = x_{(\frac{n+1}{2})} = x_{(\frac{9+1}{2})} = x_{(\frac{10}{2})} = x_5.$$

Assim, a mediana é o termo que ocupa a quinta(5ª) posição no rol. Portanto, $x_5 = 7$. Mediana igual a 7 litros.

3.6.3 Moda

Definição 3.6 *Considere* $X = (x_{(1)}, x_{(2)}, x_{(3)}, ..., x_{(n)})$ *um conjunto de informações. Denominamos moda*(x_{mod}) *o elemento do conjunto* X *que apresenta a maior frequência absoluta.*

Exemplo 3.12 *Dado o rol(R)* $= (5, 5, 5, 7, 7, 8, 8, 10, 12)$, *calcule a moda.*

Solução:

Como a moda é a observação de maior frequência absoluta, temos que $x_{mod} = 5$.

3.6.4 Média, moda e mediana em uma tabela de frequência

Dada uma tabela de frequência de variável quantitativa, podemos identificar a média, a moda e a mediana dos dados à qual ela se refere.

Exemplo 3.13 *Calcule a média, a moda e a mediana na tabela de frequência dada a seguir:*

Número de filhos	Freq. absoluta (n_i)	Freq. relativa (f_i)	Porcentagem ($100 \cdot f_i$)
0	4	0,2	20
1	5	0,25	25
2	7	0,35	35
3	3	0,15	15
4	0	0	0
5	1	0,05	5
Total	20	1	100%

- Média aritmética

$$\begin{aligned}
\overline{x} &= \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i} \\
&= \frac{4 \cdot 0 + 5 \cdot 1 + 7 \cdot 2 + 3 \cdot 3 + 0 \cdot 4 + 1 \cdot 5}{4 + 5 + 7 + 3 + 0 + 1} \\
&= \frac{33}{20} \\
\overline{x} &= 1,65.
\end{aligned}$$

Portanto, a média aritmética é 1,65 filho.

- Moda
  Como a moda trata da realização de maior frequência absoluta, temos que $x_{mod} = 2$.

  Portanto, a moda é 2 filhos.

- Mediana
  Como $n = 20$ é par, temos que x_{med} será:

$$
\begin{aligned}
x_{med} &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} \\
&= \frac{x_{(\frac{20}{2})} + x_{(\frac{20}{2}+1)}}{2} \\
&= \frac{x_{(10)} + x_{(11)}}{2} \\
&= \frac{2+2}{2} \\
x_{med} &= 2.
\end{aligned}
$$

Portanto, a mediana é igual à 2 filhos.

3.6.5 Média, moda e mediana na tabela de classes

Definição 3.7 *Dado o intervalo [a,b[, o valor médio desse intervalo é dado por* $m_i = \frac{a+b}{2}$.

Exemplo 3.14 *Dada a tabela abaixo, calcule a média, a moda e a mediana.*

Salário	Freq. absoluta (n_i)	Freq. relativa (f_i)	Porcentagem($100 \cdot f_i$)
$4 \dashv 7,3$	7	0,19	20
$7,3 \dashv 10,6$	12	0,33	33
$10,6 \dashv 13,9$	8	0,22	22
$13,9 \dashv 17,2$	5	0,14	14
$17,2 \dashv 20,5$	3	0,09	9
$20,5 \dashv 23,8$	1	0,03	3
Total	36	1	100%

Inicialmente, calculemos os valores médios de cada intervalo de classes da tabela (m_i) como segue

$$
\begin{aligned}
m_1 &= \frac{4+7,3}{2} \Rightarrow m_1 = 5,65, \\
m_2 &= \frac{7,3+10,6}{2} \Rightarrow m_2 = 8,95, \\
m_3 &= \frac{10,6+13,9}{2} \Rightarrow m_3 = 12,25, \\
m_4 &= \frac{13,9+17,2}{2} \Rightarrow m_4 = 15,55,
\end{aligned}
$$

$$\begin{aligned} m_5 &= \frac{17,2+20,5}{2} \Rightarrow m_5 = 18,85, \\ m_6 &= \frac{20,5+23,8}{2} \Rightarrow m_6 = 22,15. \end{aligned}$$

- Média aritmética

$$\begin{aligned} \overline{x} &= \frac{\sum_{i=1}^{n} m_i n_i}{\sum_{i=1}^{n} n_i} \\ &= \frac{5,65 \cdot 7 + 8,95 \cdot 12 + 12,25 \cdot 8 + 15,55 \cdot 5 + 18,85 \cdot 3 + 22,15 \cdot 1}{7+12+8+5+3+1} \\ &= \frac{401,4}{36} \\ \overline{x} &= 11,15. \end{aligned}$$

Portanto, a média de salários dos funcionários é 11,15 salários mínimos.

- Moda
A moda será o valor de m_i de maior frequência absoluta. Logo, $x_{mod} = 8,95$.

Logo a moda é igual a 8,95 salários mínimos.

- Mediana
Como $n = 36$, temos que,

$$\begin{aligned} x_{med} &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} \\ &= \frac{x_{(\frac{36}{2})} + x_{(\frac{36}{2}+1)}}{2} \\ &= \frac{x_{(18)} + x_{(19)}}{2} \\ &= \frac{8,95+8,95}{2} \\ x_{med} &= 8,95. \end{aligned}$$

Portanto, a mediana dos salários é 8,95 salários mínimos.

Obs: x_{18} e x_{19} são as informações que ocupam, respectivamente, a 18ª e a 19ª posições na tabela de classes. Por isso, $x_{18} = 8,95$ e $x_{19} = 8,95$.

3.7 Medidas de dispersão

Ao obter um conjunto de informações, os primeiros descritores que aplicamos sobre esse conjunto são as medidas de posição. Mas será que essas medidas sempre garantem a boa representatividade do conjunto? Segundo Magno [15], não, pois a boa representatividade de um conjunto está diretamente ligada à variabilidade da informação que ele traz consigo. Para termos mais clareza sobre esse assunto, analisemos o exemplo a seguir:

Exemplo 3.15 *Uma empresa de entretenimento recebe listas de dois grupos de indivíduos (dados a seguir), as quais trazem informações sobre a idade de cada um deles. Tal empresa deve realizar atividades específicas voltadas a cada grupo.*

- *Grupo 1: A = (48, 52, 54, 56, 60).*
- *Grupo 2: B = (12, 38, 54, 66, 100).*

Vamos calcular a média aritmética e a mediana de cada grupo para tentar descrever o comportamento de cada um deles.

1. Grupo 1: (48, 52, 54, 56, 60)

   - Média aritmética

$$\begin{aligned}
\overline{x} &= \frac{\displaystyle\sum_{i=1}^{n} x_i}{n} \\
&= \frac{\displaystyle\sum_{i=1}^{5} x_i}{5} \\
&= \frac{48+52+54+56+60}{5} \\
&= \frac{270}{5} \\
\overline{x} &= 54.
\end{aligned}$$

   - Mediana

   Como $n = 5$ é ímpar, temos que

$$x_{med} = x_{(\frac{n+1}{2})} \Rightarrow x_{med} = x_{(\frac{5+1}{2})} \Rightarrow x_{med} = x_3.$$

   A mediana será o valor que ocupa a 3ª posição no rol: $x_{med} = 54$.

2. Grupo 2: (12, 38, 54, 66, 100)

- Média aritmética

$$\begin{aligned}
\overline{x} &= \frac{\sum_{i=1}^{n} x_i}{n} \\
&= \frac{\sum_{i=1}^{5} x_i}{5} \\
&= \frac{12+38+54+66+100}{5} \\
&= \frac{270}{5} \\
\overline{x} &= 54.
\end{aligned}$$

- Mediana

  Como $n = 5$ é ímpar, temos que

$$x_{med} = x_{\left(\frac{n+1}{2}\right)} \Rightarrow x_{med} = x_{\left(\frac{5+1}{2}\right)} \Rightarrow x_{med} = x_3.$$

  A mediana será o valor que ocupa a 3ª posição no rol: $x_{med} = 54$.

Percebemos que a média aritmética e a mediana dos grupos dados coincidem. Então, a partir desses dados (medidas de posição), a empresa poderá organizar atividades iguais para os dois grupos? Observamos que os dados do Grupo 2 são mais dispersos (mais heterogêneos) que os do Grupo 1, de modo que essa verificação seria inviável caso existissem apenas a média e mediana dos grupos. A partir daí surge a necessidade de calcularmos novas medidas, as quais sejam úteis para indicar tais dispersões.

Para Magno [15], apenas o conhecimento das medidas de posição não nos oferece subsídio para realizarmos uma avaliação séria sobre os efeitos das metodologias apresentadas. É nesse sentido que necessitamos definir novas medidas capazes de melhorar nossa compreensão de situações como as do exemplo acima. Assim, estudaremos um conjunto de novas medidas denominadas medidas de dispersão: desvio médio absoluto, variância e desvio padrão.

Definição 3.8 *Cosideremos um conjunto de informações* $X = (x_1, x_2, x_3, ..., x_n)$ *qualquer, cuja média aritmética é* $\overline{x}$. *O desvio* d_i *que a informação* x_i *apresenta em relação à média* $\overline{x}$ *é*

$$d_i = x_i - \overline{x}.$$

Observamos que o valor d_i será negativo para qualquer $x_i \leq \overline{x}$, e positivo para qualquer $x_i \geq \overline{x}$.

Propriedade: a soma de todos os desvios em relação à média de um conjunto de informações $X = (x_1, x_2, x_3, ..., x_n)$ é zero. Isto é,

$$\sum_{i=1}^{n} d_i = 0.$$

Demonstração:

$$\begin{aligned}
\sum_{i=1}^{n} d_i &= \sum_{i=1}^{n} (x_i - \overline{x}) \\
&= [(x_1 - \overline{x}) + (x_2 - \overline{x}) + (x_3 - \overline{x}) + ... + (x_n - \overline{x})] \\
&= (x_1 + x_2 + x_3 + ... + x_n) - (\underbrace{\overline{x} + \overline{x} + \overline{x} + ... + \overline{x}}_{n \quad vezes}) \\
&= \sum_{i=1}^{n} x_i - n\overline{x} \\
&= \sum_{i=1}^{n} x_i - n\frac{\sum_{i=1}^{n} x_i}{n} \\
&= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i \\
\sum_{i=1}^{n} d_i &= 0.
\end{aligned}$$

Ao analisarmos a propriedade que acabamos de demonstrar, verificamos que o desvio médio apresenta importância quando tratamos cada informação, em particular em relação à média, mas não quanto ao conjunto total de informações. Para tanto, se considerarmos, em vez do desvio em relação à média, a distância da informação à média, isto é, o desvio absoluto $|d_i| = |x_i - \overline{x}|$, poderemos definir o desvio médio absoluto como:

3.7.1 Desvio médio absoluto

Definição 3.9 *Dado um conjunto de informações $X = (x_1, x_2, x_3, ..., x_n)$, o desvio médio absoluto é definido como a média das distâncias que essas informações apresentam em relação à média, isto é,*

$$\delta(X) = \frac{\sum_{i=1}^{n} |x_i - \overline{x}|}{n}.$$

Exemplo 3.16 *Retornando ao Exemplo 3.15 que traz as informações das idades dos grupos, sabendo que $\overline{x}(A) = \overline{x}(B) = 54$, temos*

$$\delta(A) = \frac{|48-54|+|52-54|+|54-54|+|56-54|+|60-54|}{5} = \frac{16}{5} = 3,2$$

e

$$\delta(B) = \frac{|12-54|+|38-54|+|54-54|+|66-54|+|100-54|}{5} = \frac{116}{5} = 23,2.$$

Observemos que as idades do Grupo 2 (B) são mais dispersas que as do Grupo 1 (A). O desvio médio absoluto capta justamente essa distorção entre as informações dos grupos, a cada vez que o desvio se aproxima ou se distancia do zero: quanto mais próximo de zero, menor a dispersão. O desvio absoluto nos oferece clareza, porém, em questões mais bem elaboradas estatisticamente, o uso dos módulos pode dificultar a manipulação. Assim, necessitamos definir a medida de dispersão mais conhecida: a variância.

3.7.2 Variância

Definição 3.10 *Dado um conjunto de informações* $X = (x_1, x_2, x_3, ..., x_n)$*, a variância é definida como a média dos quadrados das distâncias que cada uma dessas informações apresenta em relação à média, isto é,*

$$VAR(X) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}.$$

Exemplo 3.17 *Retornando novamente ao Exemplo 3.15, em que* $A = (48, 52, 54, 56, 60)$ *e* $B = (12, 38, 54, 66, 100)$*, calculemos as variâncias desses dois grupos.*

$$\begin{aligned} VAR(A) &= \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n} \\ &= \frac{(48-54)^2+(52-54)^2+(54-54)^2+(56-54)^2+(60-54)^2}{5} \\ &= \frac{80}{5} \\ VAR(A) &= 16 \quad anos^2 \end{aligned}$$

e

$$\begin{aligned} VAR(B) &= \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n} \\ &= \frac{(12-54)^2+(38-54)^2+(54-54)^2+(66-54)^2+(100-54)^2}{5} \end{aligned}$$

$$\begin{aligned} &= \frac{4280}{5} \\ VAR(B) &= 856 \quad anos^2. \end{aligned}$$

De modo análogo ao desvio médio absoluto, a variância pode representar a dispersão do grupo usando os dados obtidos. Para Magno [15], o uso da variância gera certo desconforto por distorcer as escalas. Isso porque a unidade de medida da variância é o quadrado da unidade da variável. Nesse caso, a unidade ficaria anos ao quadrado, nada conveniente.

Por essa razão, definimos o desvio padrão como a raiz quadrada da variância. Dessa forma, permanece o tratamento matemático dos dados e desaparecem as distorções das escalas.

3.7.3 Desvio padrão

Definição 3.11 *Dado um conjunto de informações* $X = (x_1, x_2, x_3, ..., x_n)$*, o desvio padrão DP é definido como sendo a raiz quadrada da variância, isto é,*

$$DP(X) = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}} = \sqrt{VAR}$$

Exemplo 3.18 *Retornando, mais uma vez, ao Exemplo 3.15, temos*

$$DP(A) = \sqrt{16} = 4 \quad \text{anos}$$

e

$$DP(B) = \sqrt{856} \approx 29,26 \quad \text{anos.}$$

De modo geral, ao calcularmos as medidas de dispersão, quanto mais o valor que obtivermos se aproxima de zero, mais homogêneo é o conjunto de informações. E, reciprocamente, quanto mais se distancia de zero, mais heterogêneo (disperso) é o conjunto de informações.

3.7.4 Desvio padrão nas tabelas de frequência e de classes

A variância na tabela de frequência é dada por

$$VAR(X) = \frac{\sum_{i=1}^{n} n_i(x_i - \overline{x})^2}{n}.$$

Como o desvio padrão é definido como sendo a raiz quadrada da variância, temos que

$$DP(X) = \sqrt{\frac{\sum_{i=1}^{n} n_i(x_i - \overline{x})^2}{n}}$$

Exemplo 3.19 *Calcule o desvio padrão das informações constantes da tabela a seguir.*

Número de filhos	Freq. absoluta (n_i)	Freq. relativa (f_i)	Porcentagem ($100 \cdot f_i$)
0	4	0,2	20
1	5	0,25	25
2	7	0,35	35
3	3	0,15	15
4	0	0	0
5	1	0,05	5
Total	20	1	100%

- Variância

  Como $\overline{x} = 1,65$ (calculada no Exemplo 3.13), temos que

  $$VAR(X) = \frac{\sum_{i=1}^{n} n_i(x_i - \overline{x})^2}{n}$$

  $$= \frac{4 \cdot (0 - 1,65)^2 + 5 \cdot (1 - 1,65)^2 + 7 \cdot (2 - 1,65)^2 + 3 \cdot (3 - 1,65)^2 + 1 \cdot (5 - 1,65)^2}{20}$$

  $$= \frac{30,55}{20}$$

  $$VAR(X) = 1,5275 \quad filho^2$$

- Desvio padrão

  Como, $DP(X) = \sqrt{VAR(X)}$, temos que

  $$DP(X) = \sqrt{1,5275} \Rightarrow DP(X) \approx 1,24 \quad \text{filho}.$$

Em se tratando de tabela de classes, assim como nas medidas de tendência central, consideraremos o valor médio de cada intervalo m_i. Logo, a variância em uma tabela de classe será definida por

$$VAR(X) = \frac{\sum_{i=1}^{n} n_i(m_i - \overline{x})^2}{n}.$$

Por consequência, o desvio padrão será definido por

$$DP = \sqrt{\frac{\sum_{i=1}^{n} n_i(m_i - \overline{x})^2}{n}}.$$

Exemplo 3.20 *Calcule o desvio padrão das informações constantes da tabela a seguir.*

Salário	Freq. absoluta (n_i)	Freq. relativa (f_i)	Porcentagem($100 \cdot f_i$)
$4 \dashv 7,3$	7	0,19	20
$7,3 \dashv 10,6$	12	0,33	33
$10,6 \dashv 13,9$	8	0,22	22
$13,9 \dashv 17,2$	5	0,14	14
$17,2 \dashv 20,5$	3	0,09	9
$20,5 \dashv 23,8$	1	0,03	3
Total	36	1	100%

- Variância

Pelo exemplo 3.14 temos que:
$\overline{x} = 11,15$ e $m_1 = 5,65$, $m_2 = 12,25$, $m_3 = 15,55$, $m_4 = 18,85$, $m_5 = 22,15$ e $m_6 = 22,15$, então:

$$VAR(X) = \frac{\sum_{i=1}^{n} n_i.(m_i - \overline{x})^2}{n}$$

$$= \frac{(5,65-11,15)^2.7+(8,95-11,15)^2.12+(12,25-11,15)^2.8+(15,55-11,15)^2.5+(18,85-11,15)^2.3+(22,15-11,15)^2.1}{7+12+8+5+3+1}$$

$$= \frac{675,18}{36}$$

$VAR(X) \approx 18,75 \quad salarios^2$

- Desvio padrão

Como $DP(X) = \sqrt{VAR(X)}$, temos que

$$DP(X) = \sqrt{18,755} \Rightarrow DP \approx 4,33 \quad \text{salários.}$$

3.8 Representações gráficas

As definições encontradas neste subcapítulo são fundamentadas na obra "Conexões com a matemática"[8].

A representação gráfica compreende o conjunto de ferramentas gráficas utilizadas em Estatística para condensar dados utilizando figuras geométricas relacionadas. Seu objetivo principal é proporcionar uma visualização objetiva dos dados coletados em relação à variável estudada. Para que a representação gráfica seja o mais eficaz possível, é necessário que atenda a alguns critérios estabelecidos:

- Simplicidade - O gráfico deve ser destituído de detalhes de importância secundária ou traços desnecessários que induzam o observador ao erro.

- Clareza - O gráfico deve apresentar uma correta interpretação dos valores estudados.

- Veracidade - O gráfico deve apresentar a verdade sobre o(s) fenômeno(s) estudado(s).

3.8.1 Gráficos de barras (horizontal ou vertical)

Os gráficos de barras verticais apresentam os dados por meio de colunas dispostas em posição vertical. A altura de cada coluna corresponde à frequência (absoluta ou relativa) dos valores observados. O exemplo abaixo nos fornece a ideia desse tipo de gráfico.

Exemplo 3.21 *O Instituto Nacional de Pesquisas Espaciais (INPE) informou que o desmatamento na região da Amazônia, medido entre agosto de 2007 e julho de 2008, foi de 11.968* km^2*, de acordo com o resultado do Projeto de monitoramento na Amazônia Legal (Prodes).*

Tabela 6: Área desmatada na Amazônia

Ano	Área desmatada(km^2)
2003	25282
2004	27423
2005	18729
2006	14039
2007	11224
2008	11968

O gráfico de barras verticais que representa a tabela citada é o seguinte.

Figura 4: Desmatamento na Amazônia

O gráfico apresentado representa o desmatamento da Amazônia a cada ano (de 2003 a 2008), durante os meses de agosto a julho, e a frequência absoluta é dada em km^2. Podemos observar que o gráfico de barras verticais é simples e de fácil interpretação, como neste exemplo.

Exemplo 3.22 *Com base no gráfico "Desmatamento na Amazônia", responda às seguintes perguntas:*
a) Em que ano tivemos a maior e a menor área devastada?
b) Sabendo que as dimensões de um campo de futebol são de $52m \times 110m$*, a área devastada em 2005 equivale a quantos campos de futebol?*

Solução: a) Observando as relações entre cada ano e o eixo vertical, percebemos que o ano de maior frequência absoluta foi 2004, com cerca de 27500 km^2 devastados, e o de menor frequência foi o de 2007, com cerca de 12000 km^2.

Solução: b) A área de um campo de futebol (A_c) com as dimensões apresentadas é

$$A_c = (52m) \cdot (110m) \Rightarrow A_c = 5720m^2$$

Agora, a área devastada (A_d) em 2003 em m^2 é

$$A_d = (25000m) \cdot (1000000m) \Rightarrow A_d = 25000000000m^2$$

Por fim, para obter a quantidade de campos de futebol (Q_c) devastados, basta dividir (A_d) por (A_c):

$$Q_c = \frac{A_d}{A_c} = \frac{25000000000}{5720} \approx 4370629$$

.

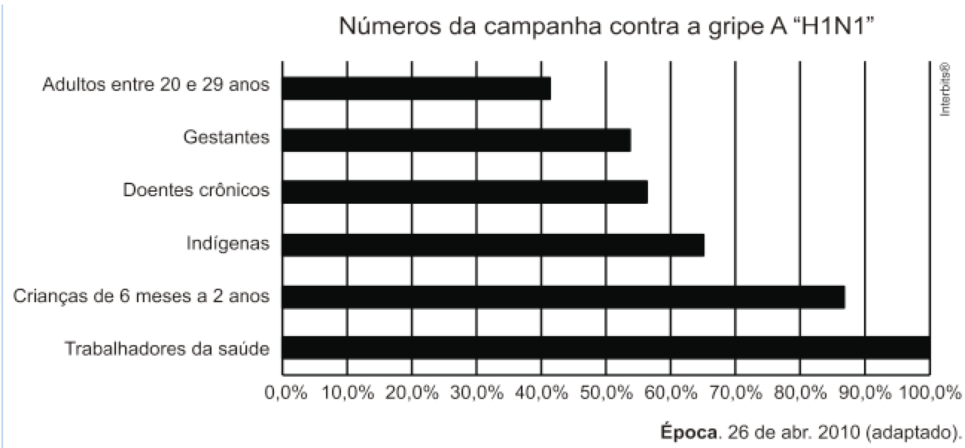
Portanto, a área devastada de floresta no ano de 2003 foi de, aproximadamente, 4370639 campos de futebol.

Outra forma de apresentar as informações coletadas é por meio do gráfico de barras horizontais. Esse tipo de gráfico apresenta os dados por meio de barras na posição horizontal. Os comprimentos das barras correspondem à frequência (absoluta ou relativa) dos valores observados.

Podemos observar que tal como o gráfico de barras verticais, o gráfico de barras horizontais é simples e de fácil interpretação, como no exemplo a seguir.

Exemplo 3.23 *A gripe H1N1 (ou influenza A) é provocada pelo vírus H1N1 da influenza do tipo A. Ele é resultado da combinação de segmentos genéticos do vírus humano da gripe, do vírus da gripe aviária e do vírus da gripe suína, que infectaram porcos simultaneamente. O período de incubação desse vírus varia de 3 a 5 dias. A transmissão desse tipo de gripe pode ocorrer antes de aparecerem os sintomas. Ela se dá pelo contato direto com os animais ou com objetos contaminados e de pessoa para pessoa, por via aérea ou por meio de partículas de saliva e de secreções das vias respiratórias. Experiências recentes indicam que esse vírus não é tão agressivo quanto se imaginava.*

Figura 5: Grupo de pessoas imunizadas a gripe H1N1



O gráfico de barras horizontais acima apresenta os percentuais dos grupos de pessoas imunizadas em uma campanha de vacinação no ano de 2010.

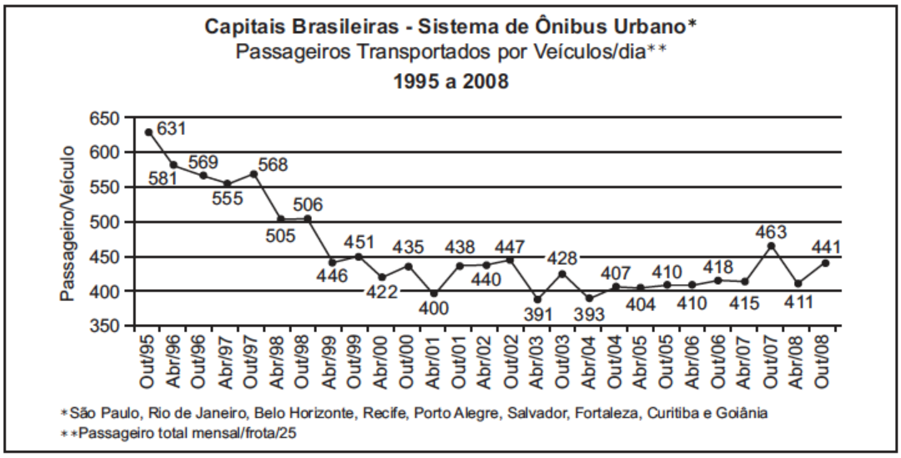
3.8.2 Gráfico de segmentos

Os gráficos de segmentos (ou gráficos de linha) são muito empregados para representar o comportamento de um conjunto de dados ao longo de um período. Para construir um gráfico de segmentos, adotamos um referencial parecido com o plano cartesiano, no qual os pontos correspondentes aos dados são marcados e, em seguida, unidos por segmentos de reta.

Observemos um exemplo de gráfico de segmentos.

Exemplo 3.24 *Dados da Associação Nacional de Empresas de Transportes Urbanos (ANTU) mostram que o número de passageiros transportados mensalmente nas principais regiões metropolitanas do país vem caindo sistematicamente. Eram 476,7 milhões de passageiros em 1995, e esse número caiu para 321,9milhões em abril de 2001. Nesse período, o tamanho da frota de veículos mudou pouco, tendo no final de 2008 praticamente o mesmo tamanho que tinha em 2001.*

Figura 6: Sistema de ônibus urbano

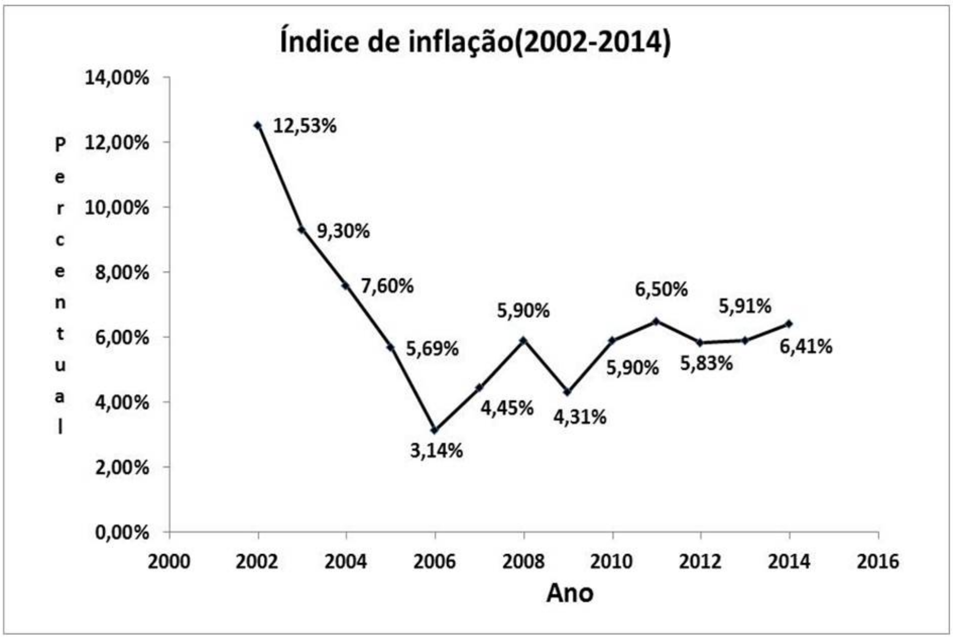


Disponível em: http://www.ntu.org.br. Acesso em 16 jul. 2009 (adaptado).

Este gráfico apresenta o quantitativo de passageiros de ônibus que utilizaram este tipo de transporte nos meses de outubro de 1995, e abril e outubro de 1996 a 2008, nas principais capitais brasileiras.

Exemplo 3.25 *Inflação é um conceito econômico que representa o aumento de preços dos produtos num determinado país ou região, durante certo período. Em um processo inflacionário, o poder de compra da moeda diminui. O gráfico de linhas a seguir mostra a oscilação da inflação entre 2002 e 2014.*

Figura 7: Índice de inflação (2002-2014)



Podemos extrair, de forma bastante clara, muitas informações desse gráfico. Questionamentos como:

- Em qual ano tivemos o maior índice de inflação? 2002, com índice igual a 12,53%

- Em qual ano tivemos o menor índice de inflação? 2006, com 3,14%

- Qual a amplitude do conjunto de dados? $\Delta(M) = 12,53 - 3,14 = 9,39\%$

3.8.3 Gráfico de setores

O Gráfico de setores (Gráfico de pizza) é usado para mostrar as diferenças entre as proporções de uma totalidade. Nele os valores são normalmente expressos em porcentagem; tem o formato de uma circunferência e é dividida em setores, com ângulos centrais proporcionais às frequências (ocorrências) das classes (100% correspondem a 360°). Consideremos o seguinte exemplo.

Exemplo 3.26 *Os produtos de alta tecnologia, como os aviões, não têm grande participação nas exportações brasileiras: somente 3,73% das vendas nacionais são de mercadorias de alta tecnologia.*

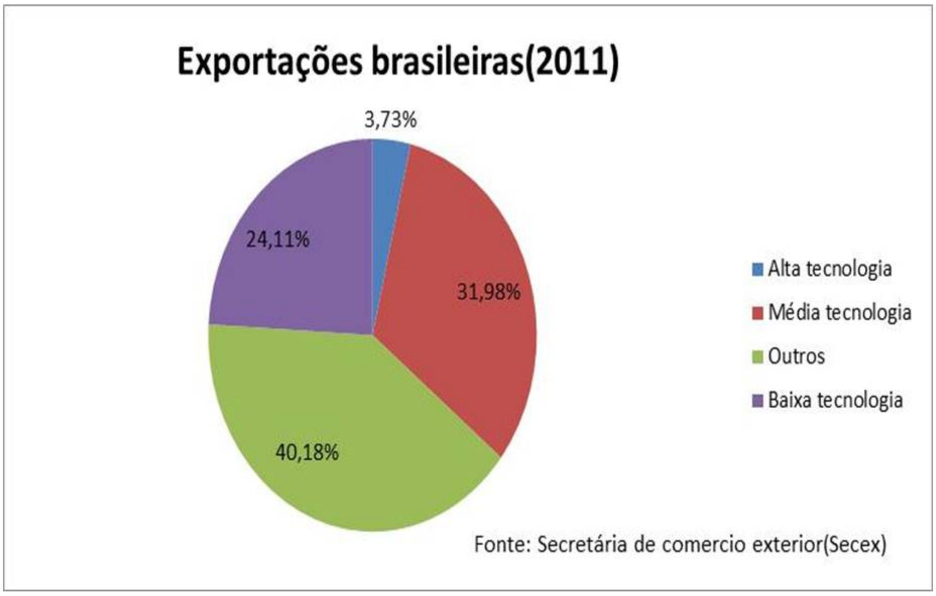
A tabela abaixo apresenta a participação de cada tipo de produto (em porcentagem) nas exportações brasileiras.

Tabela 7: Exportações brasileiras (2011)

Tipo de produto	Porcentagem
alta tecnologia	3,73%
média tecnologia	31,98%
baixa tecnologia	24,11%
outros[6]	40,18%
Total	100%

O gráfico de setores que representará o exemplo será:

Figura 8: Exportações brasileiras (2011)



Calculamos o ângulo x de um setor que representa determinada realização por meio de proporcionalidade: $360°$ está para 100%, assim como x está para a realização (em porcentagem) de interesse. Por exemplo, $360°$ está para 100%, assim como x está para $3,73\%$, ou seja, $100 \cdot x = 360 \cdot 3,73 \Rightarrow x \approx 13°$.

Analogamente, podemos calcular todos os demais ângulos dos setores que compõem o gráfico dado.

3.8.4 Histograma

Quando queremos representar uma distribuição de frequências cuja variável tem seus valores agrupados em intervalos (tabelas de classes), costumamos usar o histograma.

Esse tipo de gráfico é formado por retângulos justapostos cujas bases são construídas sobre o eixo das abcissas, cada retângulo representando um intervalo ($[a, b[$) da tabela.

Exemplo 3.27 *A poluição sonora ocorre quando num determinado ambiente o som altera a condição normal de audição. Embora ela não se acumule no meio ambiente, como outros tipos de poluição, causa vários danos ao corpo e à qualidade de vida das pessoas.*

O ruído é o que mais colabora para a existência da poluição sonora. Ele é provocado pelo som excessivo das indústrias, canteiros de obras, meios de transporte, áreas de recreação, etc. Estes ruídos provocam efeitos negativos para o sistema auditivo das pessoas, além de provocar alterações comportamentais e orgânicas.

A OMS (Organização Mundial de Saúde) considera que um som deve ficar em até 50 db (decibéis, unidade de medida do som) para não causar prejuízos ao ser humano. A partir de 50 *db, os efeitos negativos começam. Alguns problemas podem ocorrer a curto prazo, outros levam anos para serem notados.*

A tabela de classes a seguir, mostram os resultados de uma aferição dos ruídos sonoros(em decibéis) durante 50 horários diferentes em uma movimentada avenida de São Paulo.

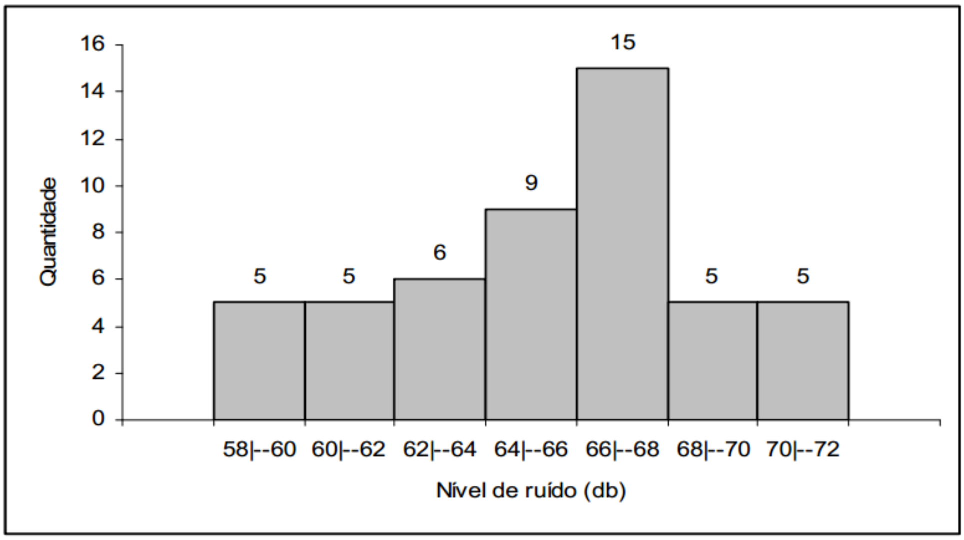
Tabela 8: Ruídos sonoros(em decibéis)

Decibéis(db)	Freq. absoluta (n_i)	Freq. relativa (f_i)	Porcentagem($100 \cdot f_i$)
$58 \dashv 60$	5	0,1	10
$60 \dashv 62$	5	0,1	10
$62 \dashv 64$	6	0,12	12
$64 \dashv 66$	9	0,18	18
$66 \dashv 68$	15	0,3	30
$68 \dashv 70$	5	0,1	10
$70 \dashv 72$	5	0,1	10
Total	50	1	100%

Cada retângulo justaposto representra um intervalo de classe diferente. Por exemplo, o primeiro retângulo refere-se ao intervalo $[58, 60[$, que representa os ruídos maiores que ou iguais a 58 *db* e menores que 60 *db*. Para todos os outros retângulos componentes do histograma, a análise é análoga à do primeiro.

Portanto o histograma que representa a tabela de classes é dado por:

Figura 9: Ruídos sonoros(em decibéis)



4 ANÁLISE DE LIVROS DIDÁTICOS

Neste capítulo, trataremos da análise dos livros didáticos adotados atualmente no Ensino Médio, no que diz respeito à Estatística. Comentaremos os pontos positivos, bem como os negativos, de tais livros. Para tal, baseamo-nos no livro "Exame de textos: Análise de livros de Matemática para o Ensino Médio", de Elon Lages Lima *et al*[9].

4.1 Fundamentos para análise dos livros-texto de Matemática para o Ensino Médio

Segundo Elon [9], a análise dos livros-texto para o Ensino Médio deve considerar sua adequação às três componentes básicas seguintes:

- Conceituação
  Compreende a formulação de definições, enunciados de proposições, estabelecimento de conexões entre os diversos conceitos, bem como a interpretação e a reformulação desses conceitos sob diferentes aspectos. No que diz respeito à conceituação, o analisador do livro deve se atentar para:

  1. Erros: dentre os erros, os mais visados, e que devem ter maior atenção ao analisar o material didático, são:

     - erros provenientes de desatenção;

     - erros de raciocínio;

     - erros de definição;

     - erros resultantes de conceitos mal formulados.

  2. Excesso de formalismo: isso se deve a, muitas vezes, utilizar, por exemplo, definições desnecessárias, as quais não acrescentaram nada em relação ao conhecimento matemático.

  3. Linguagem inadequada: está relacionada tanto à linguagem gramatical expressa no material, quanto na representação das sentenças matemáticas.

  4. Imprecisão: principalmente nas definições. Muitas dessas deixam de ser definidas corretamente ou até mesmo deixam de ser apresentadas.

  5. Obscuridade: muitas vezes, os textos matemáticos apresentam trechos ambíguos, contraditórios. A didática deve aliar-se à conceituação para esclarecer tais problemas.

6. Objetividade: muitas vezes, os textos trazem fórmulas, definições, dentre outros, cuja aplicação no cotidiano ou até mesmo em estudos posteriores os estudantes não conseguem perceber. Sugere-se que o material dê relevância a termos mais objetivos, que apresentem real necessidade serem estudados.

7. Conexões: os conteúdos estudados devem ter conexões entre eles, ou até mesmo com outras áreas de conhecimento.

- Manipulação

  Está diretamente relacionada à habilidade no manuseio de equações, fórmulas, operações e construções geométricas elementares. Essas manipulações, realizadas de forma adequada, farão com que o estudante crie atividades mentais automáticas, facilitando, assim, o desenvolvimento do raciocínio matemático.

- Aplicação
  O material deve nos oferecer meios para que possamos aplicar os conhecimentos adquiridos, desde problemas do cotidiano até questões relacionadas a outras áreas do conhecimento.

4.2 Matemática: Contexto e Aplicações (Luiz Roberto Dante)

Ao analisar o livro "Matemática: Contexto e Aplicações", de autoria de Luiz Roberto Dante, no que tange as três componentes básicas, podemos ressaltar:

4.2.1 Conceituação

O material traz um breve histórico do conteúdo associado a um exemplo, mostrando assim, a objetividade do estudo da estatística. As definições são bem claras em relação aos termos estatísticos: população, amostra e indivíduos de uma pesquisa Os mesmos são apresentados sem excesso de formalismo. Ao tratar as variáveis, percebemos a conexão e a dinamicidade de sua apresentação, exemplos envolvendo várias áreas do conhecimento, justificam tais qualidades, facilitando assim, o entendimento do leitor.

Percebemos algumas imprecisões quando o material trata das tabelas de classes, o autor não justifica o rol da pesquisa, outro fator é a quantidade de classes da tabela, simplesmente discorre "Utilize um valor maior que quatro". O ideal seria a utilização de algum método para esse cálculo (Sturges ou método da raiz).

Três fatores são bastante visíveis quando o autor retrata as representações gráficas: a linguagem adequada, a objetividade ao mostrar através de exemplos como usaremos as definições até então apresentadas e a conexão da estatística com as mais diversas áreas de conhecimento através dos mesmos exemplos.

As medidas resumo, em específico as de tendência central, o livro define apenas a média aritmética, deixando à mediana e a moda sem definições, gerando assim certa imprecisão, referência-la apenas como “termo central (caso número de observações for ímpar) ou média dos termos centrais (caso número de observações for par)”pode causar dúvidas ao leitor, o ideal seria encontrar o rol do conjunto de informações e achar a posição do(s) termo(s) que compõe(m) a mediana através de sua definição. Quanto às medidas de dispersão (variância e desvio padrão), há a preocupação em mostrar o que significa e para que sirva cada uma delas, para isso utiliza-se um exemplo que o auxilia a fazer a verificação.

4.2.2 Manipulação

O material é bem claro ao manipular as definições nele encontradas, porém por deixar de apresentar outras de importância relevante, pode causar dificuldades na construção do raciocínio para resolução de algumas atividades propostas pelo autor.

4.2.3 Aplicação

Apesar de ser uma das principais deficiências dos livros didáticos brasileiro segundo Lima [9], a aplicação do assunto abordado é vista em todo desenvolvimento do capítulo, através de seus exemplos e atividades, e em algumas atividades específicas como:

1. “Elaborando uma pesquisa escolar”' onde buscam-se temas relacionados ao dia dia e organiza-se os dados estatisticamente.
2. “Estatística no computador”a aplicação ds estatística nas ciências da computação.
3. “Pensando no ENEM”em que o autor simula uma situação problema com características voltadas ao exame.

As quais darão subsídios necessários pra que o leitor perceba a relação da Estatística e seu cotidiano.

4.3 Matemática Paiva (Manoel Paiva)

Ao analisar o livro “Matemática Paiva”de Manoel Paiva, com relação as três componentes estabelecidas, temos:

4.3.1 Conceituação

Ao definir a estatística fazendo uma alusão a uma situação problema, o autor mostra a objetividade no estudo da mesma. Porém o mesmo peca em não fazer nenhuma menção à história da estatística deixando termos como: Como surgiu? Pra que surgiu? Sem respostas para o leitor, ou seja, não dá significado ao estudo do conteúdo. O material discorre de forma objetiva aos termos da pesquisa estatística (população, amostra, variáveis e rol) trazendo exemplos para mostrar cada um deles.

Outro ponto importante dentro da conceituação é a conexão entre as tabelas de frequência e de classes com os gráficos estatísticos, ambos são trabalhados paralelamente, mostrando assim que eles são formas de organização e apresentação de dados de uma pesquisa. Percebem-se algumas imprecisões quando o material aborda a construção das tabelas de classes.

1. Não apresenta nenhuma forma para determinar a quantidade de classes. O ideal seria a apresentação do método mais utilizado (Sturges).

2. A forma que o material apresenta a amplitude dos intervalos de classes é bastante confusa, atribuem-se valores aleatórios sem dar fundamentos ao mesmo.

3. O material apresenta cada classe utilizando a forma $[a, b[$, e deixa de utilizar e mencionar o símbolo $\dashv$ que possui mesmo significado e é o mais usado principalmente em exames futuros.

Ao tratar das medidas de posição, o livro traz situações cotidianas antes de citar qualquer uma delas, posteriormente cada uma das medidas são definidas, e assim o autor relaciona as mesmas as situações citadas. Isso proporciona que leitor perceba a real utilidade dos conceitos aprendidos. Os exemplos e os exercícios apresentados são atuais e tratam de temas relevantes, tais como economia, esportes, dados geográficos, etc. E mostra assim a conexão da estatística com outras áreas de conhecimento. Por fim o autor trata das medidas de dispersão, e analogamente as medidas de posição o livro define tais medidas voltadas a situações problemas cotidianas mostrando as conexões entre as áreas do conhecimento.

4.3.2 Manipulação

O autor é muito preciso na manipulação das definições apresentadas, principalmente no que trata das medidas resumo. Todas são definidas claramente, sem excessos de formalismo, dando assim mecanismos para que o leitor desenvolva raciocínios para desenvolvimento das atividades propostas.

4.3.3 Aplicação

O material preocupa-se com a aplicação do conteúdo desde o inicio do capítulo, ao trazer situações problemas relacionadas às mais diversas áreas do conhecimento, como: economia, esportes e ciências humanas. Em uma atividade específica bastante interessante“matemática sem fronteiras”o autor apresenta um texto no qual é perceptível a aplicação da estatística na política, analisando minuciosamente uma pesquisa eleitoral.

4.4 Matemática: ensino médio (Kátia Stocco Smoles e Maria Ignez Diniz)

Analisando o livro “Matemática: ensino médio”de Kátia Stocco e Maria Ignez Diniz em relação às três componentes básicas estabelecidas temos:

4.4.1 Conceituação

As autoras iniciam o capítulo com uma introdução trazendo um exemplo muito interessante, "A urbanização nos continentes", no qual já é feita conexões do conteúdo a ser estudado a várias áreas das ciências humanas. Ao definir os termos estatísticos, percebemos a imprecisão do material ao deixar de citar alguns. Por exemplo, define-se população mas não relata a amostra, o que seria importante para entender como de fato é feita uma pesquisa estatística. As variáveis são definidas apenas como quantitativas e qualitativas não há a subdivisão das mesmas. Existem vários pontos que apresentam mais imprecisões, quando o material trata das tabelas de classes.

1. O livro não trata da amplitude do conjunto de informações, logo não menciona o rol dos daos da pesquisa.

2. Não é apresentado nenhum método para o cáculo da quantidade de classes da tabela.

3. A amplitude dos intervalos de classe também são obtidos através de valores aleatórios, não é apresentado nenhuma justificativa para tais valores.

Outros problemas são encontrados ao analisar as medidas de posição. São mostrados através de exemplos a importância, e a para que sirvam os valores encontrados para essas medidas, o problema encontra-se justamente na forma em que as autoras encontram esses valores. O material não define nenhuma delas, simplesmente as calculam em cima dos exemplos citados, causando dúvidas em relação à generalização.

Em relação as medidas de dispersão, as mesmas tem seus objetivos mostrados com clareza. Perguntas como: O que são as medidas de dispersão? Para que servem as medidas? São compreendidas ao lê-se o texto introdutório. Porém analogamente as medidas de posição, não são definidas.

4.4.2 Manipulação

Por apresentar vários problemas na conceituação, a manipulação das atividades também fica comprometida, principalmente no que nos remete a generalização do conhecimento adquirido. Por exemplo, por não apresentar o rol, como o leitor manipularia uma sequência em que não estivesse em ordem crescente? Ou seja, o material não dá mecanismos para o usuário desenvolver corretamente o raciocínio. Ao não apresentar as definições das medidas resumo, em questões um pouco mais complexas certamente geraria muitas dificuldades de resolução.

O ideal seria que as autoras definissem todos os termos estatísticos básicos, em especial o rol e as medidas resumo, dando assim melhores condições para o desenvolvimento do raciocínio.

4.4.3 Aplicação

Apesar dos problemas detectados na conceituação e posteriormente na manipulação, o que percebemos e que as autoras preocupam-se muito com relação à aplicação da estatística em assuntos cotidianos, nas mais diversas áreas do conhecimento. Desde as poucas definições apresentadas, mostram-se exemplos em áreas como; economia, saúde e ciências humanas. Outro fator é a preocupação com exames futuros, trazendo várias atividades adaptadas de vestibulares e ENEM. Algumas atividades específicas mostram a aplicação do conteúdo:"ler para resolver"onde é mostrada a estatística na produção de alimentos e "conexão: matemática, meteorologia e climatologia", onde a estatística é usada para verificar a confiabilidade das previsões do tempo.

4.5 Matemática: ciência e aplicações (Gelson Iezzi)

Ao analisar o livro "Matemática: ciência e aplicações"de Gelson Iezzi, no que tange as três componentes básicas temos:

4.5.1 Conceituação

O material introduz o conteúdo a ser trabalhado de forma breve e objetiva, os termos estatísticos (população, amostra e variáveis), e as tabelas de frequência e de classes, são apresentados, sendo que essas últimas sem nenhum exemplo mostrado em relação a sua construção(posteriormente nos exercícios tratados é pedido para construir tabelas).

Percebem-se algumas imprecisões, em particular, a abordagem feita pelos autores sobre as tabelas de classes, simplesmente é apresentado o significado do símbolo $\dashv$, utilizados nos intervalos de classes. Questões como: quantidades de intervalos, amplitude dos intervalos, e até mesmo a organização da tabela ficam de fora do conteúdo trabalhado.

As representações gráficas também são apresentadas resumidamente, os gráficos (barras, linha, setores, pictogramas e histogramas) são expressos através de exemplos, nos quais os autores fazem referências aos mesmos. Posteriormente as atividades apresentadas são bem interessantes e mostram a conexão do conteúdo a outras áreas de conhecimento.

O material mostra-se completo ao tratar das medidas resumo (posição e dispersão), todas são bem definidas e os exemplos apresentados são bem explorados deixando assim o conteúdo muito bem explicado. As medidas de posição (média, moda e mediana) são definidas, com isso facilita a generalização do conhecimento em questões mais complexas (apresentadas nas atividades propostas). Analogamente, as medidas de dispersão têm suas definições e principalmente seus objetivos em uma pesquisa muito bem expostos pelos autores, facilitando assim a resolução de atividades propostas.

4.5.2 Manipulação

Ao ser muito breve ao apresentar algumas definições, ou até mesmo, não apresentá-las, o material não dá condições de resolvermos algumas atividades propostas posteriormente. Porém no que remete as medidas resumo, por serem muito bem conceituadas o leitor não se limita a certo pontos, utilizando os conhecimentos adquiridos conseguem desenvolver as atividades propostas no final do capítulo.

4.5.3 Aplicação

Apesar de a conceituação ser breve em alguns pontos, o que podemos perceber é que o material analisado preocupa-se em aplicar o conhecimento transmitido. Todos os exemplos são bastante atuais e com temas relevantes ao cotidiano. Mas duas atividades em específico dão noção da aplicabilidade da estatística.

1. As pesquisas eleitorais, um texto explicando termos bastante divulgados pela mídia e pouco esclarecidos em materiais didáticos, as margens de erros da pesquisa e o nível de confiança da mesma.

2. Como estimar públicos em grandes eventos, no qual o autor mostra como a estatística é aplicada neste cálculo.

4.6 Novo olhar matemática (Joamir Roberto de Souza)

Ao analisar o material em relação as três componentes básica vemos:

4.6.1 Conceituação

O autor inicia o capítulo sobre Estatística de maneira bastante interessante, por meio do texto "a internet", no qual é apresentada a história e a evolução da internet, e, a partir disso, ele explora dois gráficos mostrando, antes mesmo de definir os termos estatísticos, e a conexão do conteúdo ao tema abordado no texto.

O autor apresenta os termos estatísticos fazendo alusão a assuntos atuais e muito discutidos pela sociedade, mostrando, assim, a conexão da Estatística com as mais diversas áreas de conhecimento. O consumo de água e as fontes energéticas são exemplos dos temas tratados. Os exercícios explorados não apresentam erros e seguem a mesma linha de raciocínio da parte explicativa, dando relevância a temas atuais.

Ao apresentar um texto sobre a importância dos impostos no Brasil, no qual o autor discorre sobre arrecadação, distribuição e a importância da nota fiscal, são organizadas tabelas de frequências para expor os dados extraídos desse texto. A área da saúde também é explorada ao ser apresentada a tabela de classes. Um texto sobre o triglicerídeo é apresentado, e, a partir dele, o autor fornece a construção da tabela de classes. Porém, alguns pontos nesse processo de construção podem ser melhorados para fornecer mais clareza ao leitor. São eles:

1. os dados da pesquisa já são apresentados ordenados (em ordem crescente). O fato é que, se esse dados são fornecidos de forma desordenada (como geralmente ocorre), podem causar dúvidas ao leitor. O ideal é definir o rol de uma pesquisa estatística.

2. a quantidade de classes da tabela não está bem definida, causando dúvidas ao leitor. A apresentação aleatória de um valor não fornece condições para calcular a quantidade de classes caso o número de observações seja muito grande. O ideal é mostrar o método mais usado para esse cálculo (Regra de Sturges).

Paralelamente ao estudo das tabelas de frequência e de classes, o autor faz análise de vários tipos de gráficos (setores, barras, linhas e histogramas), utilizando-se dos mesmos exemplos, o que facilita o entendimento, pois as tabelas e os gráfico têm a função de expressar e apresentar, de forma clara, os resultados de uma pesquisa estatística.

As medidas de tendência central são bem definidas, apresentando exemplos em que são utilizados os conhecimentos adquiridos. Esses exemplos abordam temas que fazem parte do cotidiano do aluno. Ao abordar, em específico, a mediana, as observações já são apresentadas ordenadamente, e, com isso, o autor deixa de apresentar o rol, que é bastante importante caso essas observações sejam fornecidas desordenadamente. As medidas de dispersão também são apresentadas formalmente, precedidas de exemplos nos quais o autor mostra os objetivos dos cálculos de tais medidas.

4.6.2 Manipulação

Por deixar de definir alguns termos estatísticos, em particular o rol, a manipulação de algumas atividades propostas cobradas nos exercícios, acabam sendo prejudicadas. Porém ao conceituar as medidas de resumo de forma clara, fornece mecanismos necessários para resolução dos exrcícios relacionados a estes termos.

4.6.3 Aplicação

O que percebemos no desenvolver do capítulo voltado a Estatística, é a preocupação do autor em mostrar as aplicações do conteúdo nas mais diversas áreas de conhecimento. Na saúde com o texto sobre triglicerídeos, relacinando com tabelas de frequências. O desperdício de alimentos, as fontes enegéticas, a arrrecadação de tributos e a incidência de raios no território brasileiro, são temas bastante atuais, no qual a aplicação da Estatística é fundamental na análise dos mesmos.

5 A ESTATÍSTICA NO ENEM

O conteúdo contido neste capítulo foi fundamentado nas seguintes referências: [3], [7] e [10].

5.1 Breve histórico e objetivos

O Exame Nacional do Ensino Médio (ENEM) é um exame individual, de caráter voluntário, oferecido anualmente aos estudantes que estão concluindo ou que já concluíram o Ensino Médio. Seu objetivo principal é possibilitar uma referência para autoavaliação, a partir das competências e habilidades que estruturam o exame. Tais competências e habilidades são bem claras, em particular, no que diz respeito ao tema Estatística, na matriz de referência do ENEM [10]:

Competência de área 6 - Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extrapolação, interpolação e interpretação.

- H24 - Utilizar informações expressas em gráficos ou tabelas para realizar inferências.

- H25 - Resolver problemas com dados apresentados em tabelas ou gráficos.

- H26 - Analisar informações expressas em gráficos ou tabelas como recurso para a construção de argumentos.

Competência de área 7 - Compreender o caráter aleatório e não determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas, determinação de amostras e cálculos de probabilidade para interpretar informações de variáveis apresentadas em uma distribuição estatística.

- H27 - Calcular medidas de tendência central ou de dispersão de um conjunto de dados expressos em uma tabela de frequências de dados agrupados (não em classes) ou em gráficos.

- H28 - Resolver situações-problema que envolvam conhecimentos de Estatística e Probabilidade.

- H29 - Utilizar conhecimentos de Estatística e Probabilidade como recurso para a construção de argumentação.

- H30 - Avaliar propostas de intervenção na realidade utilizando conhecimentos de Estatística e Probabilidade.

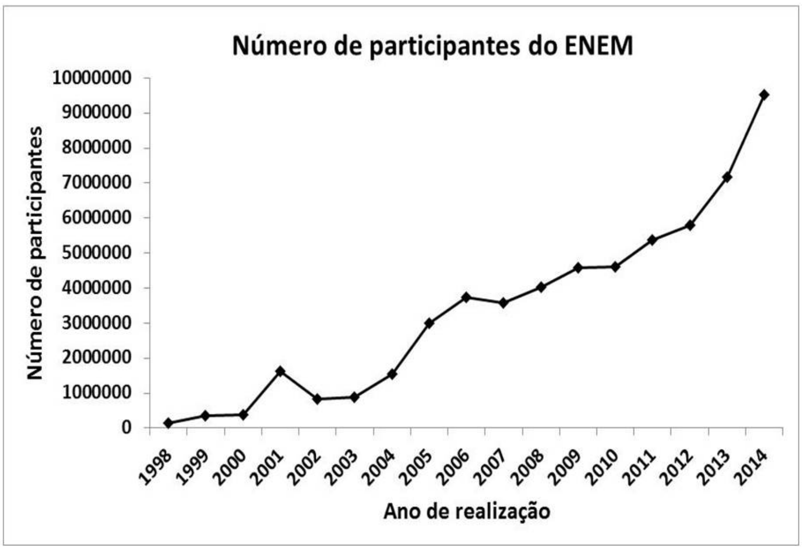
Criado em 1998, o ENEM teve como princípio avaliar o aluno ao concluir o Ensino Médio e, assim, auxiliar o Ministério da Educação e Cultura (MEC) a elaborar novas políticas educacionais que interferissem positivamente no desenvolvimento da educação. O ENEM foi dividido em duas fases:

1. Velho ENEM, aplicado entre 1998 e 2008, cujas provas eram compostas de 63 questões, aplicadas em um dia de prova, e os resultados não serviam para ingressar no Ensino Superior.

2. Novo ENEM, com provas aplicadas a partir de 2009, compostas por 180 questões, divididas em eixos de conhecimentos[7] e aplicadas em dois dias. Tem o intuito de unificar os vestibulares das universidades públicas e, posteriormente, distribuir bolsas parciais ou integrais em instituições privadas de Ensino Superior por meio do Programa Universidade para Todos (Prouni).

O gráfico abaixo mostra a evolução do Enem, em relação ao número de participantes, de acordo com o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP):

Figura 10: Participantes do ENEM (1998-2014)



O Novo ENEM cria, por certo, demandas que não existiam, exige dos professores uma série de posturas que antes não lhes eram comuns, peculiares. Estipula a necessidade

[7]Os eixos são: Linguagens, Códigos e suas Tecnologias, Ciências da Natureza e suas Tecnologias, Ciências Humanas e suas Tecnologias, e Matemática e suas Tecnologias.

de leitura e atualização constante por parte dos estudantes (e, em contrapartida, pelos educadores com os quais estarão trabalhando). Propõe, através de suas questões, o desenvolvimento do raciocínio, da capacidade de se relacionar, da possibilidade de ir além da mera memorização de fórmulas e dados[5]. Grandes dificuldades são encontradas pelos alunos com relação a resolução das questões propostas pelo exame, e isso é comprovado através das análises das médias obtidas pelos estudantes que fizeram a avaliação, segundo o INEP[7].

Tabela 9: Média em Matemática no ENEM

Ano de realização	Média em matemática
2012	569,55
2013	508,06
2014	473,5

Segundo Freitas[5] as dificuldades na resolução das questões, estão diretamente ligadas à: extensão das questões, interpretração de gráficos e tabelas e interdisciplinaridade envolvidas nas mesmas.

Percebendo a evolução do Exame, e as dificuldades apresentadas pelos discentes, resolvemos relacionar o conteúdo abordado neste trabalho e a avaliação do eixo Matemática e suas Tecnologias do novo ENEM. Inicialmente, observamos que esse eixo de conhecimento é composto por 45 (quarenta e cinco) questões de múltipla escolha, e que cerca de 20% dessas estão, direta ou indiretamente, relacionadas à Estatística.

A partir dessas informações, selecionamos e resolvemos em detalhes algumas dessas questões do novo ENEM.

5.2 Questões do ENEM sobre Estatística

As questões resolvidas e analisadas neste subcapítulo foram extraídas das avalições anuias do novo ENEM (2009-2014) [7].

Exemplo 5.1 *Questão 161 (ENEM 2009, caderno azul, página 25) Suponha que a etapa final de uma gincana escolar consista em um desafio de conhecimentos. Cada equipe escolheria 10 alunos para realizar uma prova objetiva, e a pontuação da equipe seria dada pela mediana das notas obtidas pelos alunos. As provas valiam, no máximo, 10 pontos cada. Ao final, a vencedora foi a equipe Ômega, com 7,8 pontos, seguida pela equipe Delta, com 7,6 pontos. Um dos alunos da equipe Gama, a qual ficou na terceira e última colocação, não pôde comparecer, tendo recebido nota zero na prova. As notas obtidas pelos 10 alunos da equipe Gama foram 10; 6,5; 8; 10; 7; 6,5; 7; 8; 6; 0. Se o aluno da equipe Gama que faltou tivesse comparecido, essa equipe:*
A) teria a pontuação igual a 6,5 se ele obtivesse nota 0.
B) seria a vencedora se ele obtivesse nota 10.
C) seria a segunda colocada se ele obtivesse nota 8.

D) permaneceria na terceira posição, independentemente da nota obtida pelo aluno.
E) empataria com a equipe Ômega na primeira colocação se o aluno obtivesse nota 9.

Solução:

O primeiro passo para calcular a mediana de um conjunto de dados é obter o rol(R):

$$R = (0; 6; 6,5; 6,5; 7; 7; 8; 8; 10; 10).$$

Como o número de observações do conjunto é par ($n = 10$), a mediana do grupo caso o aluno tire zero é

$$\begin{aligned}
x_{med} &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} \\
&= \frac{x_{(\frac{10}{2})} + x_{(\frac{10}{2}+1)}}{2} \\
&= \frac{x_5 + x_6}{2} \\
&= \frac{7+7}{2} \\
x_{med} &= 7.
\end{aligned}$$

Percebemos que os termos centrais não se alteram, caso a nota seja maior que ou igual a 8. Logo:

$$\begin{aligned}
x_{med} &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} \\
&= \frac{x_{(\frac{10}{2})} + x_{(\frac{10}{2}+1)}}{2} \\
&= \frac{x_5 + x_6}{2} \\
&= \frac{7+8}{2} \\
x_{med} &= 7,5.
\end{aligned}$$

Logo, o valor máximo da mediana é 7,5. Portanto, o resultado não se altera, qualquer que seja o valor da nota do aluno. Resposta: alternativa D.

Exemplo 5.2 *Questão 148 (ENEM 2009, caderno azul, página 22): A tabela a seguir mostra alguns dados da emissão de dióxido de carbono de uma fábrica, em função do número de toneladas produzidas.*

Produção (em ton)	Emissão de dióxido de carbono (em ppm)
1,1	2,14
1,2	2,30
1,3	2,46
1,4	2,64
1,5	2,83
1,6	3,03
1,7	3,25
1,8	3,48
1,9	3,73
2,0	4,00

Os dados na tabela indicam que a taxa média de variação entre a emissão de dióxido de carbono (em ppm) e a produção (em toneladas) é
A) inferior a 0,18.
B) superior a 0,18 e inferior a 0,50.
C) superior a 0,50 e inferior a 1,50.
D) superior a 1,50 e inferior a 2,80.
E) superior a 2,80.

Solução:

Inicialmente, calculemos a média aritmética da produção($\overline{xp}$) e, em seguida, a média aritmética da emissão de dióxido de carbono ($\overline{xe}$).
Produção:

$$\begin{aligned}\overline{x_p} &= \frac{\sum_{i=1}^{n} x_i}{n}\\ &= \frac{1,1+1,2+1,3+1,4+1,5+1,6+1,7+1,8+1,9+2}{10}\\ &= \frac{15,5}{10}\\ \overline{x_p} &= 1,55.\end{aligned}$$

Emissão:

$$\begin{aligned}\overline{x_e} &= \frac{\sum_{i=1}^{n} x_i}{n}\\ &= \frac{2,14+2,30+2,46+2,64+2,83+3,03+3,25+3,48+3,73+4}{10}\\ &= \frac{29,86}{10}\\ \overline{x_e} &= 2,986.\end{aligned}$$

Por fim, calculemos a taxa média (T), que é a razão entre a média da emissão e a média produção:

$$T = \frac{\overline{x_e}}{\overline{x_p}} \Rightarrow T = \frac{2,986}{1,55} \Rightarrow T \approx 1,93.$$

Portanto, a alternativa correta é D.

Exemplo 5.3 *Questão 168 (ENEM 2009, caderno azul, página 27) Na tabela a seguir, são apresentados dados da cotação mensal do ovo extra branco vendido no atacado, em Brasília, em reais, por caixa de 30 dúzias de ovos, em alguns meses dos anos 2007 e 2008.*

Mês	Cotação (em reais)	Ano
Outubro	83,00	2007
Novembro	73,10	2007
Dezembro	81,60	2007
Janeiro	82,00	2008
Fevereiro	85,30	2008
Março	84,00	2008
Abril	84,60	2008

De acordo com esses dados, o valor da mediana das cotações mensais do ovo extra branco nesse período era igual a:
A) 73,10.
B) 81,50.
C) 82,00.
D) 83,00.
E) 85,30.

Solução:

Para calcularmos a mediana de um conjunto de dados, obteremos inicialmente o seu rol(R).

$$R = (73,10; 81,60; 82; 83; 84; 84,60; 85,30).$$

Como o número de observações é ímpar ($n = 7$), temos que

$$\begin{aligned} x_{med} &= x_{\frac{n+1}{2}} \\ &= x_{\frac{7+1}{2}} \\ &= x_{\frac{8}{2}} \\ x_{med} &= x_4. \end{aligned}$$

Portanto, a mediana será o termo que ocupa a 4ª posição no rol: $x_{med} = 83$. Resposta: alternativa D.

Exemplo 5.4 *Questão 155 (ENEM 2010, caderno azul, página 24) O IGP-M é um índice da Fundação Getúlio Vargas, obtido por meio da variação dos preços de alguns setores da economia, do dia vinte e um do mês anterior ao dia vinte do mês de referência. Ele é calculado a partir do Índice de Preços por Atacado (IPA-M), que tem peso de 60% do índice, do Índice de Preços ao Consumidor (IPC-M), que tem peso de 30%, e do Índice Nacional de Custo de Construção (INCC), representando 10%. Atualmente, o IGP-M é o índice para a correção de contratos de aluguel e o indexador de algumas tarifas, como energia elétrica.*

INCC

Mês/ano	Índice do mês (em %)
Mar/2010	0,45
Fev/2010	0,35
Jan/2010	0,52

IPC-M

Mês/Ano	Índice do mês (em %)
Mar/2010	0,83
Fev/2010	0,88
Jan/2010	1,00

IPA-M

Mês/Ano	Índice do mês (em %)
Mar/2010	1,07
Fev/2010	1,42
Jan/2010	0,51

A partir das informações acima, é possível determinar o maior IGP-M mensal desse primeiro trimestre, cujo valor é igual a:
A) 7,03%.
B) 3,00%.
C) 2,65%.
D) 1,15%.
E) 0,66%.

Solução:

Calcularemos a média aritmética ponderada, pois, para cada observação x_i, temos o peso ni ($n_1 = 0,6;\ n_2 = 0,3;\ n_3 = 0,1$), e assim, determinarmos o IGP-M de cada mês.

- Março 2010
  $x_1 = 1,07;\ x_2 = 0,83;\ x_3 = 0,45.$

$$\begin{aligned}\overline{x} &= \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i}\\ &= \frac{0,6 \cdot 1,07 + 0,83 \cdot 0,3 + 0,1 \cdot 0,45}{0,6 + 0,3 + 0,1}\\ \overline{x} &= 0,936.\end{aligned}$$

- abril 2010

$x_1 = 1,42;\ x_2 = 0,88;\ x_3 = 0,35.$

$$\begin{aligned}
\overline{x} &= \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i} \\
&= \frac{0,6 \cdot 1,42 + 0,88 \cdot 0,3 + 0,1 \cdot 0,35}{0,6 + 0,3 + 0,1} \\
\overline{x} &= 1,151.
\end{aligned}$$

- Maio 2010
  $x_1 = 0,51;\ x_2 = 1;\ x_3 = 0,52.$

$$\begin{aligned}
\overline{x} &= \frac{\sum_{i=1}^{n} x_i . n_i}{\sum_{i=1}^{n} n_i} \\
&= \frac{0,6.0,51 + 1.0,3 + 0,1.0,52}{0,6 + 0,3 + 0,1} \\
\overline{x} &= 0,658.
\end{aligned}$$

Portanto, o maior valor para o IGP-M no primeiro trimestre de 2010 ocorreu em abril, com valor igual a 1,151%. Resposta: alternativa D.

Exemplo 5.5 *Questão 158 (ENEM 2010, caderno azul, página 24) Em uma corrida de regularidade, a equipe campeã é aquela em que o tempo dos participantes mais se aproxima do tempo fornecido pelos organizadores em cada etapa. Um campeonato foi organizado em 5 etapas, e o tempo médio de prova indicado pelos organizadores foi de 45 minutos por prova. No quadro, estão representados os dados estatísticos das cinco equipes mais bem classificadas.* ***Dados estatísticos das equipes mais bem classificadas (em minutos)***

Equipes	Média	Moda	Desvio padrão
Equipe I	45	40	5
Equipe II	45	41	4
Equipe III	45	44	1
Equipe IV	45	44	3
Equipe V	45	47	2

Utilizando os dados estatísticos do quadro, a campeã foi a equipe:
A) I.
B) II.

C) III.
D) IV.
E) V.

Solução:

A regularidade de um conjunto de dados está diretamente relacionada às medidas de dispersão (desvio médio, variância e desvio padrão): quanto mais se aproxima de zero, menor será a variabilidade do conjunto de dados, ou seja, mais regular será o conjunto. A questão apresenta o desvio padrão como a única medida de dispersão, e como as médias de todas as equipes são iguais Logo, o desvio padrão em relação a média representará a variabilidade do conjunto. Como a equipe III apresenta o menor desvio padrão (1), a equipe mais regular será a mesma. Resposta: Alternativa C.

Exemplo 5.6 *Questão 154 (ENEM 2011, caderno azul, página 24) A participação dos estudantes na Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP) aumenta a cada ano. O quadro indica o percentual de medalhistas de ouro, por região, nas edições da OBMEP de 2005 a 2009.*

Região	2005	2006	2007	2008	2009
Norte	2%	2%	1%	2%	1%
Nordeste	18%	19%	21%	15%	19%
Centro-Oeste	5%	6%	7%	8%	9%
Sudeste	55%	61%	58%	66%	60%
Sul	21%	12%	13%	9%	11%

Em relação às edições de 2005 a 2009 da OBMEP, qual o percentual médio de medalhistas de ouro da região Nordeste?
A) 14,6%
B) 18,2%
C) 18,4%
D) 19%
E) 21,0%

Solução:

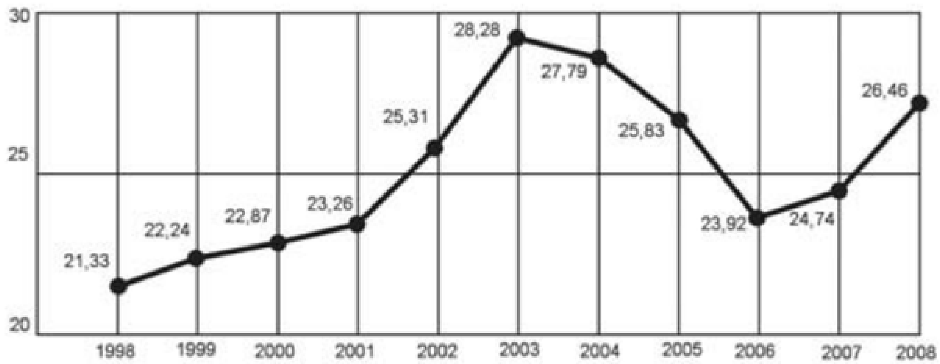
Para obtermos o percentual médio de medalhistas de ouro na região Nordeste, calcularemos a média aritmética das observações dadas na tabela:

$$\begin{aligned} \overline{x} &= \frac{\sum\limits_{i=1}^{n} x_i}{n} \\ &= \frac{18+19+21+15+19}{5} \end{aligned}$$

$$= \frac{92}{5}$$
$$\overline{x} = 18,4.$$

Portanto, o percentual médio de medalhistas de ouro na região Nordeste é 18,4%. Resposta: alternativa C.

Exemplo 5.7 *Questão 173 (ENEM 2011, caderno azul, página 29) O termo agronegócio não se refere apenas à agricultura e à pecuária, pois as atividades relacionadas a essa produção incluem fornecedores de equipamentos, serviços para a zona rural, industrialização e comercialização dos produtos.*
O gráfico seguinte mostra a participação percentual do agronegócio no PIB brasileiro:



Esse gráfico foi usado em uma palestra em que o orador ressaltou uma queda da participação do agronegócio no PIB brasileiro e a posterior recuperação dessa participação, em termos percentuais.
Segundo esse gráfico, o período de queda ocorreu entre os anos de:
A)1998 e 2001
B)2001 e 2003
C)2003 e 2006
D)2003 e 2007
E)2003 e 2008

Solução:

Basta analisarmos o gráfico dado, o qual apresenta no seu eixo horizontal os períodos das observações, e o no seu eixo vertical, os valores percentuais das observações. Assim, é fácil percebermos em quais períodos há um decrescimento nos valores percentuais. Em 2003, foi de 28,28%; em 2004, de 27,79%; em 2005, de 25,83%, e em 2006, de 23,92%. A partir de 2006, esses valores passam a crescer novamente. Resposta: alternativa C.

Exemplo 5.8 *Questão 176 (ENEM 2012, caderno azul, página 30) Um produtor de café irrigado em Minas Gerais recebeu um relatório de consultoria estatística, constando, dentre outras informações, o desvio padrão das produções de uma safra dos talhões de sua*

propriedade. Os talhões têm a mesma área de $30000m^2$*, e o valor obtido para o desvio padrão foi de 90 kg/talhão. O produtor deve apresentar as informações sobre a produção e a variância dessas produções em sacas de 60 kg por hectare* ($10000\ m^2$)*.*
A variância das produções dos talhões expressa em $(sacas/hectare)^2$ *é*
A) 20,25.
B) 4,50.
C) 0,71.
D) 0,50.
E) 0,25.

Solução:
Como 1 talhão(t) = 30000 m^2 e 1 hectare(hec) = 10000 m^2, temos que:

$$\begin{aligned} t &= 30000\ m^2 \\ &= 3.10000\ m^2 \\ t &= 3\ hec. \end{aligned}$$

De acordo com o enunciado do problema, o desvio padrão (DP) é igual 90kg/talhão(t). Assim:

$$\begin{aligned} DP &= 90kg/t \\ &= 90kg/3hec \\ &= 30kg/hec. \end{aligned}$$

Como o relatório final será em sacas de $60kg$, fazendo a conversão para sacas, obtemos

$$DP = 0,5\ sacas/hec.$$

Agora, podemos calcular a variância (VAR):

$$\begin{aligned} DP &= \sqrt{VAR} \\ 0,5\ sacas/hec &= \sqrt{VAR} \\ (0,5\ sacas/hec)^2 &= (\sqrt{VAR})^2 \\ VAR &= (0,5)^2(sacas/hec)^2 \\ VAR &= 0,25(sacas/hec)^2. \end{aligned}$$

Portanto, a alternativa correta é E.

Exemplo 5.9 *Questão 142 (ENEM 2011, caderno azul, página 21) A equipe de especialistas do centro meteorológico de uma cidade mediu a temperatura do ambiente, sempre no mesmo horário, durante 15 dias intercalados, a partir do primeiro dia de um mês. Esse tipo de procedimento é frequente, uma vez que os dados coletados servem de referência para estudos e verificações de tendências climáticas ao longo dos meses e anos.*
As medições ocorridas nesse período estão indicadas na tabela a seguir:

Dia do mês	Temperatura (em °C)
1	15,5
3	14
5	13,5
7	18
9	19,5
11	20
13	13,5
15	13,5
17	18
19	20
21	18,5
23	13,5
25	21,5
27	20
29	16

Em relação à temperatura, os valores da média, mediana e moda são, respectivamente, iguais a:
A) 17 °C, 17 °C e 13,5 °C.
B) 17 °C, 18 °C e 13,5 °C.
C) 17 °C, 13,5 °C e 18 °C.
D) 17 °C, 18 °C e 21,5 °C.
E) 17 °C, 13,5 °C e 21,5° C.

Solução:

Inicialmente, obtemos o rol(R) do conjunto de observações:

$R = (13,5; 13,5; 13,5; 13,5; 14; 15,5; 16; 18; 18; 18,5; 19,5; 20; 20; 20; 21,5)$.

Agora, calcularemos as medidas de tendência central.

- Média aritmética

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i}$$

$\overline{x} = \frac{4 \cdot 13{,}5 + 1 \cdot 14 + 1 \cdot 15{,}5 + 1 \cdot 16 + 2 \cdot 18 + 1 \cdot 18{,}5 + 1 \cdot 19{,}5 + 3 \cdot 20 + 1 \cdot 21{,}5}{4+1+1+1+2+1+1+3+1}$

$\overline{x} = \dfrac{255}{15}$

$\overline{x} = 17.$

Portanto, a média será 17°C.

- Moda

A moda é o valor x_i (Realização) de maior frequência n_i. Portanto, a observação de maior frequência é igual a 13,5°C.

- Mediana

Como o número total de observações é ímpar ($n = 15$), a mediana é:

$$\begin{aligned} x_{med} &= x_{\frac{n+1}{2}} \\ &= x_{\frac{15+1}{2}} \\ &= x_{\frac{16}{2}} \\ x_{med} &= x_8, \end{aligned}$$

o 8º termo do rol(R). Mediana igual a 18°C. Resposta: alternativa C.

Exemplo 5.10 *Questão 174 (ENEM 2012, caderno azul, página 29) A tabela a seguir mostra a evolução da receita bruta anual nos três últimos anos de cinco microempresas (ME) que estão à venda.*

ME	2009 (em milhares)	2010 (em milhares)	2011 (em milhares)
Alfinetes V	200	220	240
Balas W	200	230	200
Chocolates X	250	210	215
Pizzaria Y	230	230	230
Tercelagem Z	160	210	245

Um investidor deseja comprar duas das empresas listadas na tabela. Para tal, ele calcula a média da receita bruta anual dos últimos três anos (de 2009 até 2011) e escolhe as duas empresas de maior média anual.
As empresas que esse investidor escolhe comprar são:
A) Balas W e Pizzaria Y.
B) Chocolates X e Tecelagem Z.

C) Pizzaria Y e Alfinetes V.
D) Pizzaria Y e Chocolates X.
E) Tecelagem Z e Alfinetes V.

Solução:

Calcularemos a média aritmética anual das empresas listadas na tabela dada:

$$\begin{aligned}
\overline{x}_V &= \frac{\sum_{i=1}^{n} x_i}{n} = \frac{200+220+240}{3} = \frac{660}{3} = 220,\\
\overline{x}_W &= \frac{\sum_{i=1}^{n} x_i}{n} = \frac{200+230+200}{3} = \frac{630}{3} = 210,\\
\overline{x}_X &= \frac{\sum_{i=1}^{n} x_i}{n} = \frac{250+210+215}{3} = \frac{675}{3} = 225,\\
\overline{x}_Y &= \frac{\sum_{i=1}^{n} x_i}{n} = \frac{230+230+230}{3} = \frac{690}{3} = 230,\\
\overline{x}_Z &= \frac{\sum_{i=1}^{n} x_i}{n} = \frac{160+210+245}{3} = \frac{615}{3} = 205.
\end{aligned}$$

Logo, as empresas de maiores médias são: Y e X (nessa ordem). Portanto, a alternativa correta D.

Exemplo 5.11 *Questão 141 (ENEM 2014, caderno azul, página 20) Os candidatos K, L, M, N e P estão disputando uma única vaga de emprego de uma empresa e fizeram as provas de Português, Matemática, Direito e Informática. A tabela a seguir apresenta as notas obtidas por esses cinco candidatos.*

Candidato	Português	Matemática	Direito	Informática
K	33	33	33	34
L	32	39	33	34
M	35	35	36	34
N	24	37	40	35
P	36	16	26	41

Segundo o edital de seleçãoo, o candidato aprovado será aquele para o qual a mediana das notas obtidas por ele nas quatro disciplinas for a maior. O candidato aprovado será:
A) K.
B) L.

C) M.
D) N.
E) P.

Solução:

Inicialmente, obteremos o rol(R) de cada conjunto de informações dadas:

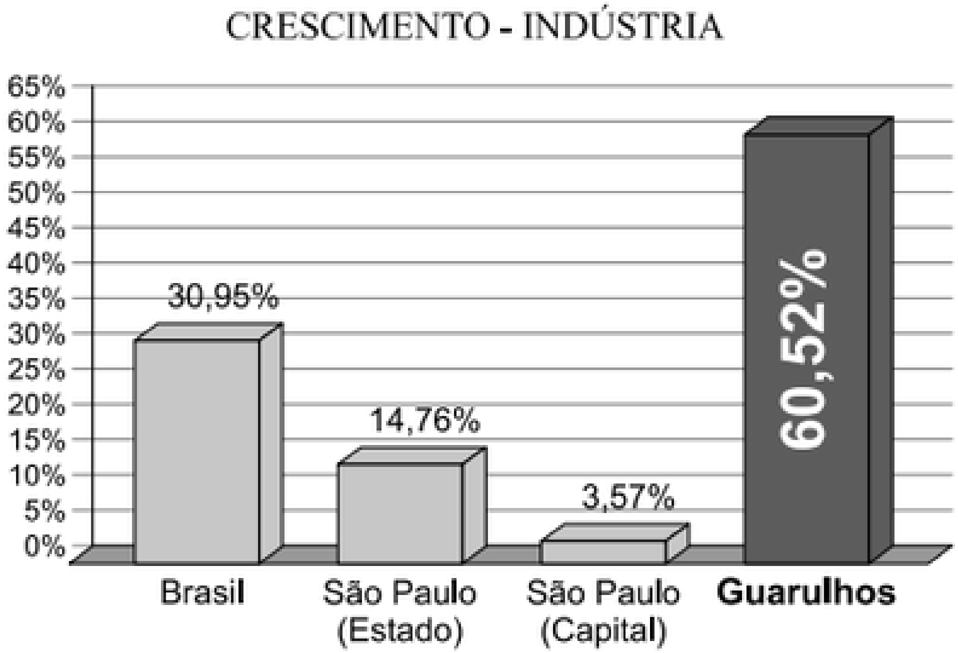
$$\begin{aligned} R_K &= (33, 33, 33, 34), \\ R_L &= (32, 33, 34, 39), \\ R_M &= (34, 35, 35, 36), \\ R_N &= (24, 35, 37, 40), \\ R_P &= (16, 26, 36, 41). \end{aligned}$$

Como o número de informações é par $n = 4$, igual para todos os conjuntos, temos que

$$\begin{aligned} (x_{med})_K &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} = \frac{x_2 + x_3}{2} = \frac{33 + 33}{2} = 33, \\ (x_{med})_L &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} = \frac{x_2 + x_3}{2} = \frac{33 + 34}{2} = 33,5, \\ (x_{med})_M &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} = \frac{x_2 + x_3}{2} = \frac{35 + 35}{2} = 35, \\ (x_{med})_N &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} = \frac{x_2 + x_3}{2} = \frac{35 + 37}{2} = 36, \\ (x_{med})_P &= \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} = \frac{x_2 + x_3}{2} = \frac{26 + 36}{2} = 31. \end{aligned}$$

Logo, o candidato com maior mediana é o candidato N. Portanto, a alternativa correta é D.

Exemplo 5.12 *Questão 144 (ENEM 2013, caderno azul, página 21) A cidade de Guarulhos (SP) tem o 8º PIB municipal do Brasil, além do maior aeroporto da América do Sul. Em proporção, tem a economia que mais cresce em indústrias, conforme mostra o gráfico a seguir.*



Analisando os dados percentuais do gráfico acima, qual a diferença entre o maior e o menor centro em crescimento no polo das indústrias?
A) 75,28
B) 64,09
C) 56,95
D) 45,76
E) 30,07

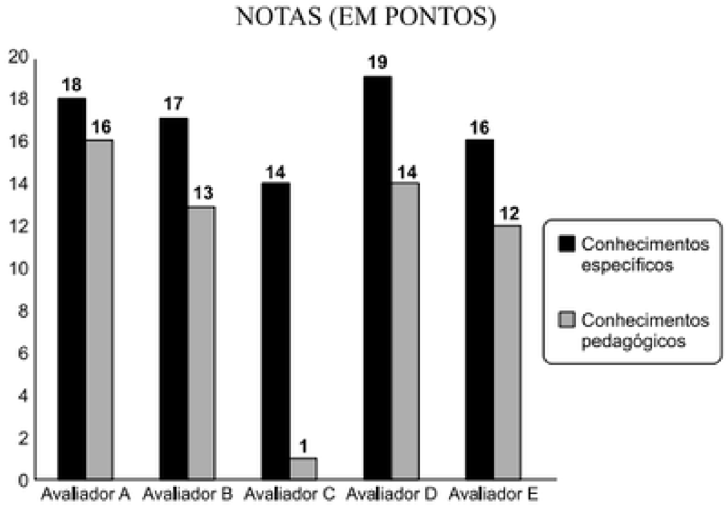
Solução:

Ao analisar o gráfico de barras dado, é fácil perceber que o maior polo de crescimento é Guarulhos, com 60,52%, e o menor é São Paulo (capital), com 3,57%. Como a questão pede a diferença (d) entre a maior e a menor taxa, temos que:

$$d = 60,52 - 3,57 \Rightarrow d = 56,95\%$$

Portanto, a alternativa correta é C.

Exemplo 5.13 *Questão 137 (ENEM 2013, caderno azul, página 19) As notas de um professor que participou de um processo seletivo, em que a banca avaliadora era composta por cinco membros, são apresentadas no gráfico a seguir. Sabe-se que cada membro da banca atribuiu duas notas ao professor, uma relativa aos conhecimentos específicos da área de atuação e outra, aos conhecimentos pedagógicos, e que a média final do professor foi dada pela média aritmética de todas as notas atribuídas pela banca avaliadora.*



Utilizando um novo critério, essa banca avaliadora resolveu descartar a maior e a menor notas atribuídas ao professor.
A nova média, em relação à média anterior, é
A) 0,25 ponto maior.
B) 1,00 ponto maior.
C) 1,00 ponto menor.
D) 1,25 ponto maior.
E) 2,00 pontos menor.

Solução:

Com as observações fornecidas pelo gráfico de barras dado, calcularemos a média aritmética de todas as notas obtidas (em um total de $n = 10$ notas).

$$\begin{aligned}\overline{x} &= \frac{\sum_{i=1}^{n} x_i}{n}\\ &= \frac{18+16+17+13+14+1+19+14+16+12}{10}\\ &= \frac{140}{10}\\ \overline{x} &= 14.\end{aligned}$$

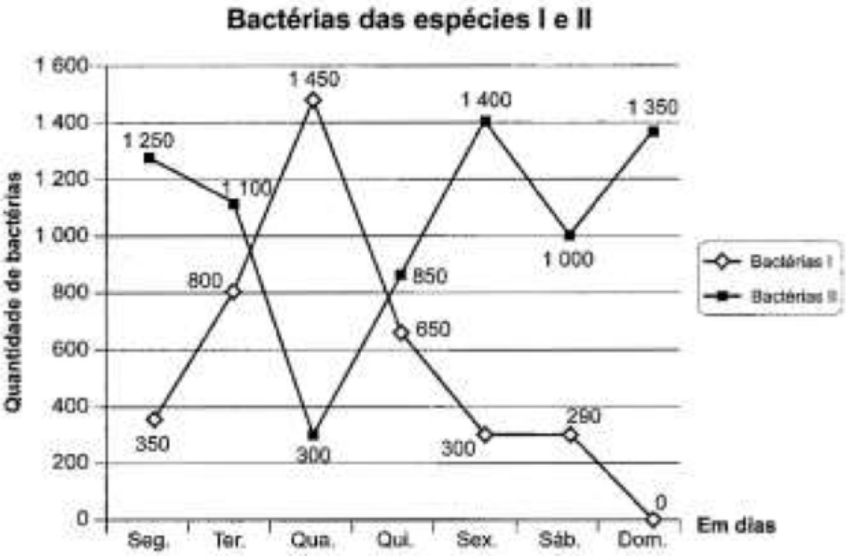
Agora, adotando o novo critério, descartaremos a nota 1 (menor nota) e a nota 19 (maior nota) (totalizando, agora, $n = 8$ notas).

$$\begin{aligned}\overline{x} &= \frac{\sum_{i=1}^{n} x_i}{n}\\ &= \frac{18+16+17+13+14+14+16+12}{8}\\ &= \frac{120}{8}\end{aligned}$$

$$\overline{x} = 15.$$

Logo, a nova média é 1 ponto maior que a antiga. Portanto, a alternativa correta é B.

Exemplo 5.14 *Questão 176 (ENEM 2014, caderno azul, página 30) Um cientista trabalha com as espécies I e II de bactérias em um ambiente de cultura. Inicialmente, existem 350 bactérias da espécie I e 1250 bactérias da espécie II. O gráfico a seguir representa as quantidades de bactérias de cada uma dessas espécies, em função do dia, durante uma semana.*



Em que dia dessa semana a quantidade total de bactérias nesse ambiente de cultura foi máxima?
A) Terça-feira. B) Quarta-feira. C) Quinta-feira. D) Sexta-feira. E) Domingo.

Solução:

Analisando o gráfico de segmentos dado, calcularemos o valor total de bactérias (Q_b) a cada dia da semana. Para isso, adicionamos a quantidade de bactérias da espécie I àquela da espécie II:

$$\begin{aligned}
(Q_b)_{Segunda} &= 1250 + 350 \Rightarrow Q_b = 1600, \\
(Q_b)_{Terca} &= 1100 + 800 \Rightarrow Q_b = 1900, \\
(Q_b)_{Quarta} &= 1450 + 300 \Rightarrow Q_b = 1750, \\
(Q_b)_{Quinta} &= 850 + 650 \Rightarrow Q_b = 1500, \\
(Q_b)_{Sexta} &= 1400 + 300 \Rightarrow Q_b = 1700, \\
(Q_b)_{Sabado} &= 1000 + 200 \Rightarrow Q_b = 1200, \\
(Q_b)_{Domingo} &= 1350 + 0 \Rightarrow Q_b = 1350.
\end{aligned}$$

Logo, o valor máximo será na terça, igual a 1900 bactérias. Portanto, a alternativa correta é A.

6 CONSIDERAÇÕES FINAIS

Ao realizar um levantamento das provas do ENEM, podemos observar que cerca de 20% das questões de Matemática (eixo Matemática e suas Tecnologias)estavam relacionadas, direta ou indiretamente, à Estatística. Daí, a escolha do tema abordado neste trabalho.

A busca pela melhor compreensão dos conteúdos matemáticos inseridos em nosso cotidiano de modo geral mostra que a Estatística, em particular, está diretamente relacionada ao nosso dia a dia. Exploração de tabelas e gráficos, estimativas futuras, dentre outros, fazem com que este trabalho seja de real utilidade, já que mostra toda a relação da Estatística com o cotidiano.

Além disso, por meio das análises realizadas nos livros didáticos recomendados pelo PNLD atual, tendo como referência o livro *Exames de textos: análise dos livros de Matemática para o Ensino Médio*[9], podemos nortear, favorecendo assim, o professor de Matemática na escolha do material a ser adotado em suas aulas, atribuindo mais objetividade e clareza a elas. Considerando questões selecionadas do ENEM sobre Estatística, mostramos, por meio de suas soluções em detalhes, que elas, em sua maioria, são de fácil interpretação e, por consequência, também de fácil solução.

Deste estudo apresentado, podemos perceber que a Estatística é de fundamental importância no cotidiano. Devido a tal importância, é imprescindível que ela seja abordada considerando as três componentes básica proposta por Lima[9]: conceituação, manipulação e aplicação, para que facilite a compreensão do conteúdo transmitido. No que diz respeito aos livros didáticos, percebemos muitos problemas relacionados, em especial, na conceituação de alguns termos estatísticos, e que de certa forma interferiram negativamente na manipulação de algumas atividades propostas. Diferentemente do que afirma Lima[9], no que diz respeito a aplicação, o que percebemos é que esta componente está bastante presente nos materiais analisados, justificando assim a Estatística como um dos ramos da matemática mais aplicáveis no contexto social.

Esperamos que com este trabalho o professor de Matemática tenha uma ferramenta que lhe sirva de norte ao trabalhar a Estatística, e, assim, desenvolva suas aulas de forma mais dinâmica e objetiva, mostrando o quanto a Estatística, que muitas vezes não é nem abordada, trata de um dos conteúdos mais aplicáveis e explorados em nosso cotidiano.

Referências

[1] BAYER, A. *Estatística na escola: importância dos conteúdos de Estatística no Ensino Fundamental e Médio.* Artigo científico, v.5, n.1, jan./jun. Natal: EDUFRN (Editora da UFRN), 2005.

[2] DANTE, Luiz Roberto. *Matemática: contexto e aplicações.* 2. ed. São Paulo: Ática, 2013.

[3] *Exame Nacional do Ensino Médio.* Disponível em http://pt.wikipedia.org/wiki/Exame-Nacional-do-Ensino-Médio (Acesso em 02/03/2015).

[4] *Exame Nacional do Ensino Médio.* Disponível em http://portal.inep.gov.br/web/enem/conteudo-das-provas (Acesso em 02/03/2015)

[5] FREITAS, Sunny Karelly S. de.*Dificuldade dos alunos de Ensino Médio na resolução das questões do Enem.* Pesquisa de campo.

[6] IEZZI, Gelson*et al. Matemática: ciência e aplicações.* v. 3. 7ª ed. São Paulo: Saraiva 2013.

[7] INEP. *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.* http://portal.inep.gov.br/, acesso em 02/04/2015.

[8] LEONARDO, Fábio Martins de (editor responsável). *Conexões com a matemática / organizadora Editora Moderna; obra coletiva concebida, desenvolvida e produzida por Editora Moderna.* 2ª ed. São Paulo: Moderna, 2013.

[9] LIMA, Elon Lages. *Exame de Textos: Análise de livros de Matemática para o Ensino Médio.* Rio de Janeiro: SBM, 2001.

[10] *Matriz de referência ENEM* Ministério da Educação e Cultura. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Páginas 6 e 7.

[11] MLODINOW, Leonardo.*O andar do bêbado.*

[12] MORETTIN, Pedro Alberto, BUSSAB, Wilton O. *Estatística básica*, 6. ed. São Paulo: Saraiva, 2010.

[13] NOGUEIRA, Paulo Apolinário, VICTER, Eline das Flores, NOVIKOFF, Cristina. *Roteiro didático para o ensino de Estatística: A cidadania na/pela Matemática.* Disponível em $www.pucrs.br/famat/viali/tic_literatura/.../produto - paulo - apolinario.pdf$ (Acesso em 03/03/15).

[14] NUNES César A. A. *Criação, produção e uso de Objetos de Aprendizagem.* . Disponível em objectosaprendizagem.no.sapo.pt/ppw/ppcn.ppt (Acesso em 03/03/15).

[15] OLIVEIRA, Magno Alves de. *Probabilidade e Estatística: Um curso introdutório.* Brasília: Editora IFB, 2011.

[16] PAIVA, Manoel. *Matemática: Paiva/Manoel Paiva*, 2. ed. São Paulo: Moderna, 2013.

[17] PCNEM *Parâmetros Curriculares Nacional do Ensino Médio* Disponível em: http://portal.mec.gov.br/seb/arquivos/pdf/ciencian.pdf. Consulta em 03/03/2015.

[18] SALSA, Ivone da silva, MOREIRA, Janete Alves, PEREIRA, Marcelo Gomes. *Matemática e realidade: interdisciplinar* Artigo científico. Natal: EDUFRN (Editora da UFRN), 2005.

[19] *Semana da Estatística 2013: A história da Estatística no Brasil*. Disponível em http://www.estatistica.ccet.ufrn.br/historia.php (Acesso em 03/03/15).

[20] SOUZA, Joamir Roberto de. *Novo olhar matemática*.v.3. 2. ed. São Paulo: FTD, 2013.

[21] STOCCO, Kátia Smole e DINIZ, Maria Ignez. *Matemática: ensino médio*. v.3. 8ª ed. São Paulo: Saraiva, 2013.