**CLÁUDIA BORIM DA SILVA**

# Pensamento Estatístico e Raciocínio sobre variação: um estudo com professores de Matemática

**DOUTORADO EM EDUCAÇÃO MATEMÁTICA**

**Pontifícia Universidade Católica**

**São Paulo**

**2007**

**CLÁUDIA BORIM DA SILVA**

# Pensamento Estatístico e Raciocínio sobre variação: um estudo com professores de Matemática

**DOUTORADO EM EDUCAÇÃO MATEMÁTICA**

Tese apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de Doutor em Educação Matemática, sob a orientação da Professora Doutora Cileda de Queiroz e Silva Coutinho

**Pontifícia Universidade Católica**

**São Paulo**

**2007**

BANCA EXAMINADORA

Á minha mãe, Elisabeth Antão Borim Lopes (*in memorian*), que me apresentou a vida de maneira brilhante, acreditando no amor e na educação.

Á minha mãe científica, Professora Doutora Anna Regina Lanner de Moura, que me apresentou a área de Educação Matemática de maneira entusiasmada, acreditando na ciência e seus benefícios.

## AGRADECIMENTOS


Agradeço ao meu marido e minhas filhas, pela compreensão, apoio, ajuda e pelo aprendizado de viver sem mim em muitos momentos.

Agradeço ao meu pai, que sempre me incentivou e rezou por mim.

Agradeço à Dona Marilene, que soube me substituir tão bem.

Agradeço à minha orientadora, que aprendeu a ser terapeuta, me consolando nos momentos de desespero e desânimo.

Agradeço ao Professor Saddo, a quem admiro e com quem eu aprendi muito.

Agradeço ao Professor Michel Henry e à professora Maria Inês pelas infindáveis contribuições à parte teórica deste trabalho.

Agradeço à Universidade São Judas Tadeu, pelas horas de pesquisa concedidas para a realização deste trabalho, sem as quais seria impossível sua conclusão.

Agradeço ao Professor Silva, que acreditou em mim e me apoiou sempre.

Agradeço às minhas amigas: Ana Lucia, Zezé, Vera, Rosaninha, Michele, Lu, Irene, sem as quais eu não teria conseguido elaborar este trabalho.

Agradeço à Professora Márcia Brito, que além de minha eterna orientadora, também colaborou muito para que este trabalho findasse.

Agradeço a professora Lisbeth, que me incentivou e contribuiu para este trabalho.

Agradeço à professora Célia, pelas contribuições na banca.

Agradeço aos meus amigos da São Judas: Adriano, Fátima, Vivian, Simone, Cláudio, Vilma, Rodrigo, Arlete, Petri, que alguma maneira colaboraram para que este trabalho fosse realizado.

Agradeço às editoras FTD, Saraiva e Moderna pela concessão dos livros de Matemática – exemplar do professor, para que fosse possível elaborar a análise praxeológica.

Agradeço especialmente a todos os professores que participaram desta pesquisa, pois sem eles não existiria trabalho.

SILVA, Cláudia Borim da. *Pensamento estatístico e raciocínio sobre variação*: um estudo com professores de Matemática. 2007. 354 f. Tese (doutorado em Educação Matemática) Pontifícia Universidade de São Paulo, São Paulo.

## RESUMO

Devido à dificuldade encontrada por alunos de graduação para a compreensão do desvio padrão, este trabalho teve como objetivo verificar o raciocínio sobre variação e variabilidade nas etapas do ciclo investigativo do pensamento estatístico. Foram participantes da pesquisa nove professores de Matemática da escola básica e dois alunos de Matemática da Universidade de São Paulo. O trabalho seguiu os pressupostos de uma pesquisa-ação e a fase de implementação teve duração de quarenta e oito horas, divididas em dezesseis encontros de três horas cada. Foram discutidos os conteúdos estatísticos: distribuição de freqüência simples e com dados agrupados, representações gráficas, medidas de tendência central e dispersão. Os níveis de raciocínio sobre variação foram classificados de acordo com o modelo proposto por Garfield (2002). O diagnóstico identificou a ausência de raciocínio sobre variação, exceção feita a um professor que apresentava raciocínio idiossincrático. Durante a fase de sensibilização da pesquisa-ação e planejamento do ciclo investigativo, os professores apresentaram naturalmente o raciocínio sobre variabilidade, mas não sobre variação. Entretanto, a experiência com a elaboração de uma pesquisa, desde a definição dos objetivos até a coleta e montagem do banco de dados permitiu um avanço no desenvolvimento do pensamento estatístico dos professores, que já transitavam em três das quatro dimensões de sua estrutura elaborada por Wild e Pfannkuch (1999). Não obstante, o desenvolvimento do pensamento estatístico não implicou diretamente em um nível mais avançado do raciocínio de variação, observado durante a fase de análise dos resultados da pesquisa. Para a comparação de três distribuições de freqüências simples de variável discreta foram utilizadas a percepção da moda, a observação dos valores máximo e mínimo e da menor freqüência e a elaboração de um intervalo de variação composto pelos valores da variável que tinham freqüência nas três distribuições, conjuntamente, que foram categorizados como raciocínio verbal de variação até raciocínio de procedimento, respectivamente. A discussão sobre as medidas de tendência central permitiu observar a interpretação equivocada de média como maioria, que se refere à moda, que foi um fator impeditivo para a percepção da necessidade de uma medida de variação. A utilização do correto significado de média motivou os professores a utilizarem medidas complementares como a moda e os valores máximo e mínimo, mas não o desvio padrão. O significado atribuído ao desvio padrão foi, predominantemente, uma medida da variação entre as observações indicando homogeneidade da amostra, aspecto reforçado pelos livros didáticos de Matemática do ensino médio e categorizado como raciocínio verbal de variação. A composição do intervalo de um desvio padrão da média não surgiu naturalmente e mesmo os participantes que compreenderam esta interpretação do desvio padrão, apresentaram dificuldade para identificar o que tinha no intervalo. Acredita-se que o desenvolvimento de aplicativos computacionais para trabalhar o conceito de intervalo em torno da média possa auxiliar na aquisição deste raciocínio, considerado um raciocínio completo de variação. Conclui-se que a linguagem “maior variação” pode induzir dois diferentes raciocínios idiossincráticos: a maior variação das freqüências em alguma categoria ou valor da variável de uma distribuição de freqüências e a maior variação de observações diferentes na amostra, ambas não relacionadas com a medida de tendência central.

**Palavras-chave**: Pensamento Estatístico; Nível de Raciocínio sobre variação; desvio padrão; professores de Matemática; pesquisa-ação.

SILVA, Cláudia Borim da. *Statistical Thinking and variation reasoning: a study with Mathematics Teachers*. 2007. 354 f. Thesis (Doctoral in Mathematics Education) Pontifícia Universidade de São Paulo, São Paulo.

## ABSTRACT

Due to student difficulty with understanding standard deviation, this work aimed to identify the reasoning about variation and variability in all parts of the investigation cycle of the statistical thinking. Nine middle and high school Mathematics teachers and two mathematics students of University of São Paulo participated in an action research, 3 hr meetings, lasting in total for 48 hrs. The contents were simple and grouping data frequency distribution, graphics, center and spread measures. The reasoning levels were classified using the general model developed by Garfield (2002). The teachers showed no variaton reasoning during the first week, except for a teacher with idiosyncratic reasoning. During the action research sensibility phase and planning of investigative cycle phase, the teachers developed variability reasoning naturally, but not about variation. However, this experience promoted an upgrade of teachers statistical thinking, that used three (between four) dimensions created by Wild e Pfannkuch (1999). Nevertheless, the statistical thinking upgrade did not implicate a gain in variation reasoning level, observed during the data analysis phase. To compare three discret variable frequency distribution were done using the perception of mode, minimum and maximum values and minimum frequency and use of the distribution chunk with range was organized with existence of the frequency in all groups, understood like verbal until procedural reasoning, respectively. The center measures discussion showed the misconception of mean, which was understood as the mode, and this inhibited necessity perception of a spread measure. The use of correct mean of arithmetic mean induced the teachers use complement measures as the mode and minimum and maximum values, but not the standard deviation. The mean fo standard deviation was predominantely a measure of number of differents observations, signal of homogeneous sample, as many Mathematics textbooks introduced the concept of variation. The comprehension of one standard deviation interval towards mean didn´t develop naturally and the teachers who understood this mean of standard deviation had difficulty to understand what was in the interval, which suposed to develop this integrated reasoning process with the educational softwares created for this intention. In conclusion, the term “more variation” can cause wastly differing results due to personal interpretation of the phrase ´more variation´ and idiosyncratic reasoning process involved in analysing complex mathematical data: more variation between frequency in only the variable category or variable value in comparing frequency distributions and more variation between sample different observations, both without use of variation from mean.

**Key words**: Statistical Thinking; Variation Reasoning level; standard deviation; Mathematics teacher; action research.

## SUMÁRIO

## LISTA DE FIGURAS

# LISTA DE TABELAS

## LISTA DE QUADROS

“*Incerteza e variabilidade estão diretamente relacionadas: devido a existência da variabilidade, nós vivemos em incerteza e como nem tudo é determinado ou certo, há variabilidade.*” (BAKKER, 2004, p. 14)

## Introdução

Este trabalho está inserido no projeto *O Pensamento Matemático – Formação de um Núcleo de Ensino, Aprendizagem e Pesquisa,* coordenado pelo Professor Doutor Saddo Ag Almouloud e que conta com quatro linhas de pesquisa: Geometria, Pensamento Algébrico, Pensamento Numérico e Tratamento da Informação.

Desde o início dos trabalhos deste grupo de pesquisa (2000), dois objetivos foram claramente apresentados: proporcionar uma formação continuada gratuita a professores de Matemática da rede pública do Estado de São Paulo e desenvolver pesquisas na área de Educação Matemática.

Motivada pelos resultados encontrados em seu trabalho de mestrado (SILVA, 2000) e pelas inquietações como professora de Estatística em cursos de graduação, esta autora viu a oportunidade de desenvolver um trabalho na área de didática da Estatística que, segundo Batanero (2000), ainda é uma área em desenvolvimento.

Devido à grande dificuldade que os alunos apresentam com o cálculo e, principalmente, com a interpretação do desvio padrão, escolheu-se pesquisar sobre o tema variação[1] (dispersão), embora a formação continuada oferecida aos professores de Matemática tenha englobado outros conteúdos da Análise Exploratória de Dados.

Como este trabalho refere-se à variação, é importante a definição do termo. Segundo Reading e Shaughnessy (2004, p. 201) “variação é um substantivo usado para descrever o ato de variar ou mudar de condição”. Suponha que uma empresa de recrutamento de trabalhadores tenha publicado o salário médio para

[1] Neste trabalho, o termo variação é sinônimo de dispersão. Segundo Reading e Shaughnessy (2004, p. 203) “as medidas de dispersão são freqüentemente referenciadas como *variabilility* ou *spread*”.

uma determinada vaga. Se a medida de variação também for publicada, o candidato terá condições de estimar seu salário futuro, caso seja selecionado.

Conforme salientado por Ben-Zvi e Garfield (2004) e Reading e Shaughnessy (2004), muitos pesquisadores têm usado os termos variação (*variation*) e variabilidade (*variability*) como sinônimos. Neste trabalho pretende-se fazer a distinção entre eles.

Segundo Reading e Shaughnessy (2004, p. 201), "variabilidade é uma forma substantiva do adjetivo variável, significando que alguma coisa está pronta ou tem propensão para variar ou mudar". Usando o mesmo exemplo do recrutamento de candidatos, para um mesmo cargo, o salário é uma entidade propensa a variar, o que permite dizer que há variabilidade no salário para determinada vaga.

Neste trabalho, esses termos serão usados no mesmo sentido que Reading e Shaughnessy (2004, p. 202) usaram. "O termo variabilidade será usado referindo-se à característica da entidade que é observável e o termo variação será usado como descrição ou medida desta característica".

O estudo das medidas de variação neste trabalho está limitado aos conteúdos sugeridos pelos Parâmetros Curriculares Nacionais de Matemática para o ensino médio - PCN+ (Brasil, 2002), já que eram os conteúdos de interesse dos professores de Matemática participantes da pesquisa.

Os estudos sobre o conceito de variação (e variabilidade) são recentes. Segundo Reading (2004), o estudo de Shaughnessy et al. [2] (1999) foi uma das primeiras experiências para descobrir, de maneira sistemática, o que estava acontecendo no entendimento de variação apresentado pelos estudantes.

Anos depois, o raciocínio sobre variabilidade foi o tema central do terceiro *Fórum de Pesquisa em Alfabetização, Pensamento e Raciocínio Estatístico* (SRTL-3) que aconteceu na Universidade de Nebraska, Estados Unidos, em Julho de 2003 e também foi o tema central das edições de Novembro de 2004 e de Maio de 2005 do SERJ (*Statistics Education Research Journal*). Segundo Pfannkuch (2005) a edição de Novembro de 2004 do SERJ é um marco no

[2] Não foi possível consultar este trabalho, que foi publicado num evento do *National Council of Teachers of Mathematics* (NCTM). Trabalhos do mesmo autor, sobre a mesma temática, posteriores a este, foram consultados e estão apresentados no Capítulo 4 desta pesquisa.

estabelecimento das fundações iniciais do raciocínio sobre variação e variabilidade.

Não só os estudos sobre o raciocínio de variação/variabilidade na área de educação são recentes como também a inserção de conteúdo estatístico na disciplina Matemática. As medidas de variação passaram a fazer parte do conteúdo a ser trabalhado na disciplina Matemática do Ensino Médio a partir de 2002, quando da publicação dos Parâmetros Curriculares Nacionais do Ensino Médio - PCN+ (BRASIL, 2002).

Esta inserção de conteúdos estatísticos na disciplina Matemática tem exigido que os professores se adaptem a essa realidade, aprimorando seu conhecimento sobre esse conteúdo de maneira que lhe permitam desenvolver atividades didáticas motivadoras e que tenham significado para o aluno.

Segundo Garfield e Ben-Zvi (2005, p. 92), é necessário preparar os professores de Matemática para lidar com o conteúdo de Estatística, pois variabilidade deveria ser enfatizada de maneira central desde as séries iniciais (com atividades e discussões formais e informais) até o ensino médio e o início da graduação, tarefa que pode ser difícil para o professor de Matemática que, possivelmente, teve apenas uma disciplina de Estatística em sua formação.

Bakker (2004) conta sobre sua experiência numa sala de 5ª série do ensino fundamental, na qual o professor fez uma pergunta que se assemelhava com Estatística e um aluno respondeu *médiamodamediana*, como se fosse uma única palavra.

> Este incidente exemplifica o que uma grande quantidade de pesquisas em educação estatística tem relatado: muito freqüentemente os alunos aprendem estatística como um conjunto de técnicas e eles não aplicam-nas sensivelmente. Mesmo que eles tenham aprendido a calcular média, mediana, moda e a elaborar histogramas e *boxplots*, a maioria não entende que eles podem usar a média como uma representação do grupo quando comparando dois conjuntos de dados. (BAKKER, 2004, p. 64)

Segundo Bakker (2004), esse triste cenário deve-se à falta de entendimento conceitual para analisar dados com as técnicas aprendidas. Ou seja, é preciso ensinar mais do que simplesmente o cálculo das medidas e elaboração de gráficos. É preciso discutir sobre o significado e a aplicabilidade das medidas e representações e, principalmente, relacionar esses conceitos.

Portanto, o objetivo deste trabalho foi identificar a percepção de variação (variabilidade) e o nível de raciocínio sobre sua medida (variação) apresentados pelos professores de Matemática durante uma formação continuada.

O Capítulo 1 apresenta uma distinção entre os termos pensamento, raciocínio e letramento estatísticos, pois a variação é componente de todos. O Capítulo 2 faz uma revisão teórica sobre diferentes medidas de variação, que permite ao leitor compreender as discussões sobre estudos já realizados em ensino-aprendizagem de variação e variabilidade, que estão apresentados no Capítulo 4 e comparar com as tarefas, técnicas e aporte teórico atualmente apresentados nos livros didáticos de Matemática do Ensino Médio, que estão no Capítulo 3.

O quinto capítulo apresenta a problemática da pesquisa e o sexto capítulo apresenta a metodologia empregada para alcançar tais problemas.

Os resultados obtidos em cada etapa do ciclo investigativo (idealizado e implementado pelos próprios participantes da pesquisa) estão apresentados do sétimo ao décimo quarto capítulos e o Capítulo 15 apresenta as considerações finais desta autora, levantando resultados importantes, limitações deste estudo e possíveis pesquisas futuras.

## 1 O Papel da Variação no Pensamento, Raciocínio e Letramento estatísticos

Variação (e variabilidade) é um conceito que está diretamente envolvido no pensamento estatístico (*statistical thinking*), no raciocínio estatístico (*statistical reasoning*) e no letramento estatístico (*statistical literacy*) e por esse motivo, este capítulo teve o objetivo de apresentar a definição desses três temas e a relação da variação (e da variabilidade) com cada um deles.

### 1.1 Letramento Estatístico

Dar-se-á início a esta discussão com o conceito de letramento estatístico e a justificativa por esta maneira de tradução.

Soares (2004) diferencia os termos alfabetização e letramento. Segundo a autora, enquanto em outros países tais termos são independentes, no Brasil estão inter-relacionados sob o conceito de alfabetização funcional na qual "após alguns anos de aprendizagem escolar, o indivíduo terá não só aprendido a ler e escrever [alfabetização], mas também a fazer uso da leitura e da escrita [letramento]" (SOARES, 2004, p.7).

Carvalho (2006), autora portuguesa, utiliza o termo literacia e diz que se refere não à própria aquisição de conhecimentos, mas sim à mobilização de competências, o saber em ação. Ou seja, mesmo um indivíduo escolarizado (teoricamente alfabetizado) pode não ser capaz de mobilizar os conhecimentos adquiridos para interpretar situações cotidianas.

Soares (2004) explica que letramento e literacia são sinônimos. Pelo fato de não existir o termo literacia nos dicionários da língua portuguesa no Brasil, este trabalho adota o termo letramento.

Gal (2002) define letramento estatístico pensando numa pessoa adulta que vive numa sociedade industrializada. Segundo o autor, letramento estatístico tem dois componentes inter-relacionados:

> a) competência das pessoas para interpretar e avaliar criticamente a informação estatística, os argumentos relacionados aos dados ou a fenômenos estocásticos, que podem se apresentar em qualquer contexto e, quando relevante, b) competência das pessoas para discutir ou comunicar suas reações para tais informações estatísticas, tais como seus entendimentos do significado da informação, suas opiniões sobre as implicações desta informação ou suas considerações acerca da aceitação das conclusões fornecidas. (GAL, 2002, p. 2-3).

Ser letrado estatisticamente auxilia o indivíduo a entender fenômenos e tendências de relevância social e pessoal tais como as taxas de criminalidade, o crescimento populacional, a produção industrial, o aproveitamento educacional, etc. (GAL, 2002). Os muitos contextos em que o letramento estatístico pode ser ativado indica que a maioria dos adultos são consumidores (ao invés de produtores) da informação estatística.

Gal (2002) propõe um modelo de letramento estatístico composto por cinco elementos cognitivos responsáveis pela competência das pessoas para compreender, interpretar e avaliar criticamente informações estatísticas, e por dois elementos de disposição, responsáveis pela postura ativa diante da informação estatística.

Os elementos cognitivos do modelo de letramento estatístico proposto por Gal (2002) são: o próprio letramento, que é a capacidade de ler informações textuais, em gráficos e/ou tabelas; os conhecimentos estatístico, matemático e do contexto; e a competência de elaborar questões críticas.

Um pré-requisito óbvio para compreender e interpretar informações estatísticas é o conhecimento de conceitos e procedimentos básicos de estatística e probabilidade que, segundo Gal (2002), não pode ser discutido em termos absolutos, mas dependente do nível de letramento estatístico esperado pelos cidadãos. Esse autor propõe cinco tópicos do conhecimento básico de Estatística: a) conhecimento dos motivos e das maneiras pelas quais a coleta de dados aconteceu; b) familiaridade com termos e idéias básicas relacionadas à Estatística Descritiva; c) familiaridade com termos e idéias básicas relacionadas às apresentações gráficas e tabulares; d) compreensão das noções básicas de probabilidade e e) conhecimento sobre como as conclusões e inferências estatísticas são obtidas.

Sobre tais noções básicas de Estatística, Gal (2002) salienta alguns pontos importantes: conhecimento, pelo menos informalmente, das idéias-chave da investigação estatística, em que a primeira na lista dos estatísticos é a existência da variação; entendimento que a média aritmética e a mediana são, simplesmente, meios para resumir um conjunto de dados a partir de sua medida de tendência central e que a média aritmética é mais afetada que a mediana; é também importante saber que a mesma informação pode ser apresentada em uma tabela ou em um gráfico e saber o que significa eventos aleatórios ou ao acaso.

No que diz respeito ao conhecimento matemático necessário para ser letrado estatisticamente, Gal (2002) lembra o artigo de Moore (1997) discutindo que é mais importante desenvolver o pensamento estatístico numa primeira disciplina Estatística do que demonstrar matematicamente um conceito. Mas Gal (2002) lembra a importância de se conhecer como se calcula uma média aritmética, por exemplo, pois torna possível compreender que um valor extremo a influencia.

Essa colocação de Gal (2002) é ratificada por esta autora, considerando a importância da articulação entre o conceito e o algoritmo no processo de construção do campo conceitual de variação, no sentido de Vergnaud (apud MOREIRA, 2002)[3].

Além de conhecer tópicos matemáticos e estatísticos, é necessário conhecer basicamente o contexto. Gal (2002, p. 17) salienta que

> o conhecimento do contexto é o principal determinante da familiaridade do consumidor com as fontes de variação e erro, pois ele pode imaginar porque uma diferença entre grupos pode ocorrer ou imaginar a razão de um estudo estar errado.

Desta forma, a relação entre a “leitura” de informações estatísticas e a compreensão do contexto pode permitir a construção do conceito de variabilidade.

O último elemento cognitivo do modelo de letramento estatístico proposto por Gal (2002) é a habilidade de criticar, ou seja, receber uma informação estatística e analisá-la criticamente. O autor salienta a importância desse elemento cognitivo na leitura de informações publicadas, por exemplo, em revistas semanais.

Além dos elementos cognitivos, há os elementos de disposição: postura crítica, atitudes e crenças. A postura crítica é a propensão de um adulto ter um comportamento questionador diante de informações quantitativas que podem ser unilaterais, viesadas ou incompletas, seja de maneira intencional ou não.

---

[3] Para Vergnaud (apud MOREIRA, 2002, p. 127) um campo conceitual é um conjunto de problemas e situações cujo tratamento requer conceitos, procedimentos e representações de tipos diferentes, mas intimamente ligados.

Quanto às crenças e às atitudes, se um indivíduo acredita ser capaz de interpretar informações estatísticas (crença) e tem uma atitude positiva[4] em relação às investigações estatísticas, ele tende a apresentar uma postura crítica em relação às informações estatísticas. O modelo de letramento estatístico proposto por Gal (2002) está interpretado na Figura 1.

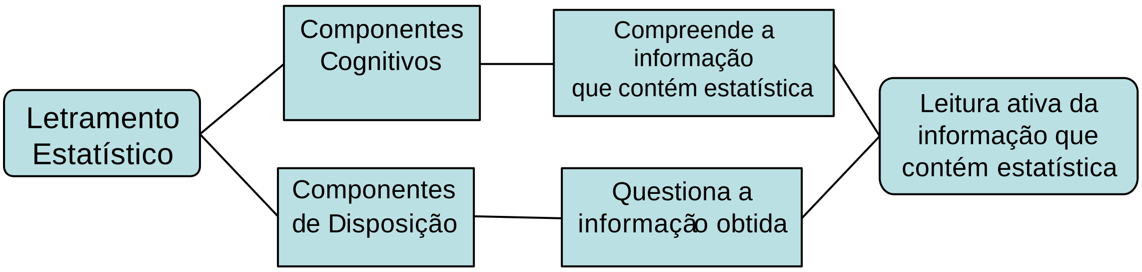



Figura 1 - Estrutura do Letramento Estatístico, segundo Gal (2002)

Gal (2002) alerta que os componentes cognitivos apresentados por ele podem sofrer alteração conforme o contexto cultural de interesse ou a sofisticação do letramento estatístico esperada do cidadão ou do trabalhador. Ou seja, os elementos constituintes do letramento estatístico podem ser diferentes quando observado um contexto de trabalho, um contexto pessoal, um contexto público e um contexto de aprendizagem formal.

Com relação à sofisticação do letramento estatístico, Watson e Callingham (2003) sugerem uma seqüência hierárquica do letramento estatístico, com seis níveis de tarefas (idiossincrático, informal, inconsistente, consistente e não-crítico, crítico e matematicamente crítico).

Watson (2003) alerta sobre a importância do contexto no letramento estatístico em que, nos dois primeiros níveis, os alunos se esforçam para interpretar a situação, mas ficam no nível do contexto. No terceiro e quarto nível, os alunos lidam com os conceitos estatísticos em diferentes contextos, mas não criticam, por exemplo, reportagens questionáveis. Só nos dois últimos níveis de

[4] Atitude é a prontidão de uma pessoa para responder a determinado objeto de maneira favorável ou desfavorável. A dissertação de Mestrado desenvolvida por esta autora, sob a orientação da Professora Dr. Márcia Regina Ferreira de Brito, foi sobre as atitudes em relação à Estatística apresentadas por alunos de graduação.

letramento estatístico os alunos compreendem os conceitos estatísticos envolvidos no relatório e apresentam postura crítica.

Para se ter adultos estatisticamente letrados, Watson (2003) diz que é necessário começar no nível escolar. Neste sentido, Rumsey (2002) apresenta uma proposta de disciplina Estatística com dois objetivos: letrar estatisticamente os alunos e desenvolver habilidades científicas de pesquisa.

Rumsey (2002) troca o termo letramento estatístico por outros dois: competência estatística e cidadania estatística e diz que esta última requer alto nível de raciocínio e pensamento estatístico. Para esta autora, competência estatística envolve cinco componentes, sintetizados no Quadro 1.

Quadro 1 - Síntese da competência estatística apresentada por Rumsey (2002)

| | |
|---|---|
| Atenção aos dados | Promove motivação aos alunos, pois os dados estão presentes na vida diária, são freqüentemente subutilizados e as decisões baseadas em dados podem ter um impacto forte em nossas vidas. |
| Entendimento básico de estatística | É a capacidade de relacionar o conceito dentro de um tema não estatístico; explicar o que o conceito significa, usá-lo em uma sentença ou dentro de um problema maior e responder questões sobre ele. Não significa saber calcular, por exemplo, o desvio padrão, mas sim compreendê-lo. |
| Coleta de dados e resultados | "Dar a oportunidade ao estudante para coletar seus próprios dados e achar os resultados estatísticos básicos pode ajudar os alunos a se apropriar de sua própria aprendizagem" (RUMSEY, 2002, p. 8). |
| Interpretação num nível básico | Saber interpretar resultados estatísticos (gráficos, tabelas, etc) com suas próprias palavras. |
| Habilidades básicas de comunicação | Leitura, escrita, demonstração da informação estatística. Transmitir para uma outra pessoa a informação estatística. |

Rumsey (2002) explica que essa competência estatística é a base para o raciocínio e pensamento estatísticos, que serão necessários para atingir sua segunda meta na disciplina, o desenvolvimento de habilidades científicas de pesquisa, que é a capacidade de explicar, julgar, avaliar e tomar decisões sobre a informação. Ou seja, essas são as habilidades que devem ser inicialmente desenvolvidas em um nível de letramento estatístico.

Watson (2003) explica que foi somente após a introdução de conteúdos estatísticos na disciplina Matemática na Austrália, no ano de 1991, que foi possível dar início a esse processo de letramento estatístico.

No Brasil, a inserção de conteúdos estatísticos na disciplina Matemática no ensino fundamental foi sugerida em 1998 (BRASIL, 1998) e em 2002 para o ensino médio (BRASIL, 2002).

Logo, devido à recente implantação dos conteúdos estatísticos na disciplina Matemática, é necessário letrar estatisticamente os professores de Matemática, atualmente responsáveis por letrar estatisticamente seus alunos. Como lembra Gal (2002), não é possível garantir que aprender fatos, regras e procedimentos estatísticos ou obter experiência estatística a partir de um projeto de análise de dados num contexto formal de sala de aula pode gerar um nível adequado de letramento estatístico.

Como salienta Rumsey (2002), para se ter cidadãos letrados estatisticamente, é necessário investir no ensino para o desenvolvimento do raciocínio e pensamento estatísticos.

## 1.2 Pensamento Estatístico

Snee (1990) definiu pensamento estatístico como o processo de pensamento que reconhece a presença da variação em torno de tudo o que se faz. Segundo ele, a identificação, a caracterização, a quantificação, o controle e a redução da variação criam a oportunidade de melhorar o processo[5].

Segundo Snee (1990), os elementos do pensamento estatístico são: reconhecimento da variação presente em todo o processo, a necessidade de

---

[5] Snee desenvolveu seus trabalhos na área de Controle de Qualidade e, portanto, refere-se a processo de produção.

dados para medir a variação e o uso de métodos e ferramentas estatísticas para quantificar e entender a variação, permitindo uma tomada de decisão.

Num artigo de 1997, Moore apresentou as três recomendações do comitê de currículo da *American Statistical Association* (ASA) e *Mathematical Association of America* (MAA): 1) enfatizar os elementos do pensamento estatístico, a saber: a necessidade de dados; a importância da produção de dados; a onipresença da variabilidade; a medida e modelagem da variabilidade ; 2) Incorporar mais dados e conceitos, menos receitas e deduções e 3) Adotar aprendizagem ativa. É interessante observar que os elementos do pensamento estatístico são os mesmos definidos por Snee (1990).

A definição de pensamento estatístico apresentada por Snee (1990) é ampliada por Wild e Pfannkuch (1999) que apresentam uma estrutura baseada em quatro dimensões: o ciclo investigativo, os tipos de pensamento, o ciclo interrogativo e as disposições.

O ciclo investigativo foi adaptado do modelo do PPDAC (*Problem, plan, data, analysis, conclusions*), que segundo Wild e Pfannkuch (1999), tem o objetivo de resolver um problema real, geralmente com a intenção de mudar um *sistema* para melhorar alguma coisa. Essa definição vai ao encontro do que Snee (1990) chamou de melhorar um processo.

A segunda dimensão de Wild e Pfannkuch (1999) é denominada de tipos de pensamento: pensamento geral e pensamento fundamental.

O pensamento geral refere-se ao planejamento do ciclo investigativo: o que vai ser feito? Como? O que já se conhece do assunto? Quanto custará? O que será necessário (material)? Os conceitos estatísticos do problema, pois isto influencia na maneira como se coleta e analisa os dados e a aplicação prática de uma técnica ou conceito, que terá sua interpretação do resultado.

O pensamento fundamental é o reconhecimento da necessidade de dados, a transnumeração, que se refere à possibilidade de mudar a representação para melhorar a compreensão do problema, a consideração da variação a partir da tomada de decisão em situações de incerteza, o uso de modelos estatísticos e a integração da estatística com o contexto.

É possível notar que o pensamento fundamental relaciona-se diretamente com os elementos do pensamento estatístico definidos por Snee (1990). É por este motivo que Wild e Pfannkuch (1999) chamaram-nos de subconjunto destes elementos.

A terceira dimensão, denominada de ciclo interrogativo, diz respeito aos questionamentos macro e micro que são delineados pelo *pensador* enquanto resolve o problema. Nesse ciclo, o *pensador* produz possibilidades, que podem ser de cunho contextual, dos dados ou estatístico, busca informação e idéias, para posteriormente interpretar o resultado estatístico, checa a informação obtida com uma referência interna (o que conhecia) e externa (literatura, outras pessoas, etc), para tomar a decisão sobre o que deve ser mantido, continuado a pesquisar, etc.

A quarta dimensão, denominada de *disposições,* pode ser entendida como o compromisso do *pensador* com o problema. Ele pode ser curioso e querer investigar mais; pode ser imaginativo e procurar enxergar o problema sob diferentes pontos de vista; pode ser céptico e questionar se as conclusões alcançadas são justas; pode tentar entender se existe uma fonte para aquela idéia observada; pode permitir que novas idéias sejam confrontadas com suas próprias e pode ser perseverante.

De acordo com estas definições, pode-se entender o pensamento estatístico como as estratégias mentais utilizadas pelo indivíduo para tomar decisão em toda a etapa de um ciclo investigativo. Uma interpretação da definição de pensamento estatístico apresentada por Wild e Pfannkuch (1999) está apresentada na Figura 2.

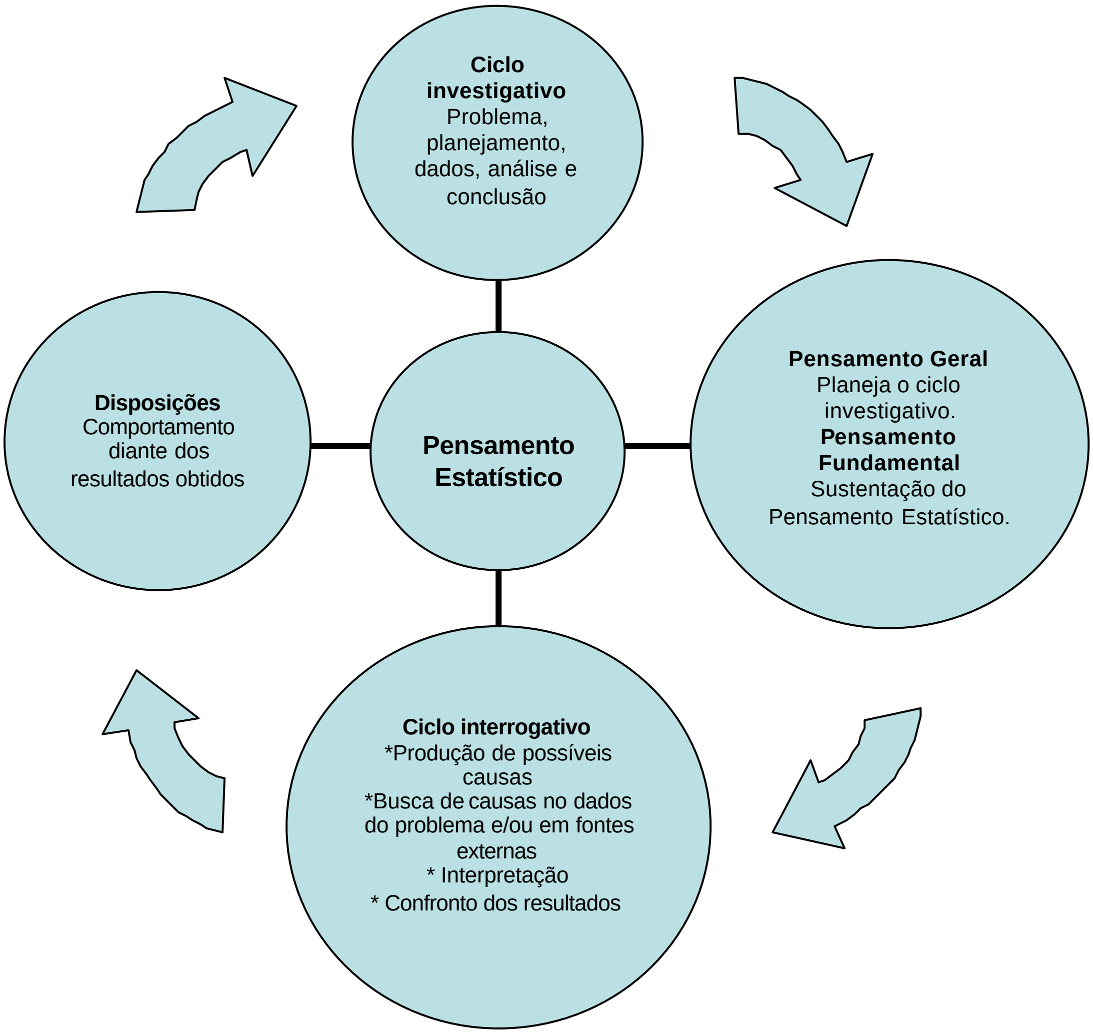



Figura 2 - Estrutura do pensamento estatístico conforme Wild e Pfannkuch (1999)

Como apresentado na Figura 2, sempre que está se fazendo uma pesquisa, está sendo usado o pensamento estatístico, mesmo de forma inconsciente. Como lembram Wild e Pfannkuch (1999, p.246),

> em muitos ambientes de pesquisa, o pensamento estatístico é como a respiração – todo mundo faz o tempo todo, raramente se lembra que está fazendo. Estatística, a disciplina, deveria ser ensinada para que as pessoas *respirem* mais efetivamente.

Esses autores consideram o ensino de Estatística como o momento propício para proporcionar condições de desenvolvimento do pensamento estatístico, fundamental para que o cidadão seja letrado estatisticamente. Como lembra Wells (apud SNEE,1990, p.117), “um dia o pensamento estatístico será necessário para o cidadão da mesma maneira que ler e escrever o são”.

Até este momento, pode-se entender que existe uma relação direta entre pensamento estatístico e letramento estatístico. Quanto mais uma disciplina

Estatística estimular o desenvolvimento do pensamento estatístico, há maior probabilidade de que os futuros cidadãos apresentem níveis de letramento estatístico mais avançados. E a variação e a variabilidade são o elemento central do pensamento estatístico e conteúdo essencial para que um indivíduo seja letrado estatisticamente.

Agora, resta introduzir o terceiro termo deste capítulo: raciocínio estatístico.

### 1.3 Raciocínio estatístico

Antes de discutir o raciocínio estatístico, é interessante apresentar a definição de raciocínio.

Segundo Costa e Capovilla (1997, p. 120), “raciocínio refere-se aos processos pelos quais as pessoas avaliam e geram argumentos lógicos, aplicando o conhecimento na consecução de metas”.

Para alguns autores, raciocínio e argumento são sinônimos, mas para Walton (1990) o raciocínio ocorre dentro de um discurso ou um argumento, ou seja, o raciocínio é usado no argumento. Para o autor “raciocínio é a elaboração de suposições denominadas premissas (ponto de partida) e o processo de mover estas premissas para a conclusão (ponto de chegada) por meio de regras” (WALTON, 1990, p. 403).

Vale ressaltar que nem todo raciocínio se manifesta na forma de argumento. Para Walton (1990, p. 411), “um indivíduo pode raciocinar em um jogo de xadrez (...) ou o raciocínio pode ocorrer para oferecer ou entender uma explicação”.

Das definições apresentadas por esses dois autores é possível entender que o raciocínio é um processo interno, mental, cujo argumento (ou o entendimento de uma explicação, ou uma ação numa situação) permite inferi-lo.

Costa e Capovilla (1997, p. 120) explicam que o estudo sobre raciocínio está intimamente ligado ao estudo de resolução de problemas e citam Evans (1993) que propõem três categorias amplas de raciocínio:

> a) os estudos sobre o <u>raciocínio dedutivo</u>, que procuram compreender como as pessoas inferem as conseqüências das informações que são dadas; ou seja, como as pessoas avaliam a validade de argumentos lógicos; b) os estudos sobre o <u>raciocínio indutivo</u>, que procuram compreender como as pessoas formulam e testam hipóteses de maneira a descobrir regras gerais e c) os estudos sobre <u>raciocínio estatístico</u>, que procuram compreender como as pessoas fazem inferências de natureza probabilística.

Especificamente sobre o raciocínio estatístico, Garfield e Gal apud Garfield (2002, p.1) explicam que é “a maneira com que as pessoas raciocinam com idéias estatísticas e como percebem a informação estatística”.

Para que uma pessoa desenvolva um raciocínio estatístico mais avançado, o ensino deve proporcionar condições para que o aluno compare conceitos, avalie a maneira mais adequada de analisar uma variável ou um conjunto de variáveis (um banco de dados), mude de representação, entenda os contra-exemplos, etc. Garfield (2002) afirma que os professores de Estatística tendem a ensinar os conceitos e os procedimentos, inclusive com a utilização de dados e aplicativos, mas esperam que o raciocínio se desenvolva como um conseqüência imediata.

Para se compreender melhor o que é raciocínio estatístico, Garfield (2002) cita algumas pesquisas desenvolvidas em diferentes temáticas: a) equívocos conceituais sobre as medidas de tendência central (confusão sobre a média como o ponto central); b) boas amostras são aquelas que apresentam uma alta porcentagem da população, sem se preocupar como foram selecionadas; c) equívoco conceitual da representatividade, em que em n jogadas de uma moeda honesta, é mais provável uma seqüência com número de caras próximo do número de coroas do que uma seqüência com todos os resultados iguais a cara, por exemplo. Sobre o raciocínio com medidas estatísticas, Garfield (2002) apresenta uma série de pontos a serem pesquisados: compreender o motivo pelo qual as medidas de tendência central, dispersão e posição apresentam diferentes informações sobre uma mesma variável; compreender quando uma medida representa (adequadamente ou não) uma variável; compreender que o uso de medidas-resumo para fazer predições será mais preciso para amostras grandes do que para amostras pequenas; saber porque uma boa medida-resumo de uma variável inclui uma medida de centro e uma medida de variação e porque as medidas de centro e dispersão podem ser usadas para comparar duas distribuições.

É possível observar que apenas com as medidas de tendência central e dispersão há uma série de possibilidades de pesquisa sobre o raciocínio estatístico e que quanto mais o aluno tiver oportunidade de vivenciar tais situações, mais refinado será seu raciocínio estatístico.

Baseada em sua experiência com o raciocínio sobre distribuição amostral, Garfield (2002) propõe um modelo geral do raciocínio estatístico, composto por cinco níveis. O Quadro 2 sintetiza esse modelo.

Quadro 2 - Síntese do modelo de raciocínio estatístico desenvolvido por Garfield (2002)

| Nível | Título | Descrição | Exemplo |
|---|---|---|---|
| 1 | Raciocínio idiossincrático | Conhecimento de algumas palavras e símbolos estatísticos, utilizados sem um entendimento completo e, freqüentemente, de maneira incorreta. | Comparar o valor da média com o valor do desvio padrão ou fazer julgamento sobre uma boa média e um bom desvio padrão. |
| 2 | Raciocínio verbal | Entendimento verbal de alguns conceitos, sem conseguir aplicá-lo a um procedimento real.<br>O indivíduo escolhe ou comunica uma definição correta, mas sem apreender seu significado. | Por quê a média é maior que a mediana em distribuições assimétricas positivas? |
| 3 | Raciocínio transitório | Capacidade de identificar corretamente uma ou duas dimensões de um conceito estatístico, sem integrá-los completamente. | Uma amostra grande apresenta um intervalo de confiança mais estreito. Um erro padrão[6] menor lida com um intervalo de confiança mais estreito.<br>Não relaciona estas duas dimensões. |
| 4 | Raciocínio de procedimento | Capacidade de identificar corretamente as dimensões de um conceito ou processo estatístico, sem integrá-los completamente ou sem entender o processo. | Um aluno sabe que a correlação não implica em causa, mas não consegue explicar a razão. |
| 5 | Raciocínio completo do processo ou conceito | Entendimento completo do processo ou conceito estatístico, coordenando as regras e os procedimentos, usando suas próprias palavras para explicar um conceito. | Explicar o que um intervalo de confiança de 95% significa em termos do processo de amostras repetidas da população. |

Usando sua experiência com pesquisa sobre o raciocínio de distribuição amostral, Garfield (2002) afirma que é necessário que um aluno tenha experiência

---

[6] Segundo Bussab e Morettin (2003, p. 309), “Se T for um estimador do parâmetro θ, chamaremos de erro padrão de T a quantidade $EP(T)=\sqrt{Var(T)}$ ”, em que Var(T) é a variância da distribuição amostral do estimador T. Por exemplo, considerando o parâmetro μ, a raiz quadrada da variância da distribuição amostral de seu estimador ($\overline{X}$) é $\sqrt{Var(\overline{X})}=\sqrt{\frac{\sigma^2}{n}}=\frac{\sigma}{\sqrt{n}}$, como, pode ser observado na página 49 deste trabalho.

Assim, $EP(\overline{X})=\sqrt{Var(\overline{X})}=\frac{\sigma}{\sqrt{n}}$.

com uma variedade de atividades: textos e explicações verbais, atividades concretas envolvendo amostragem de populações finitas, e interações com populações simuladas e distribuições amostrais quando os parâmetros são variados para que ele apresente o nível 5 do Modelo de Raciocínio.

Essa explicação de Garfield nos remete à teoria dos Campos Conceituais desenvolvida por Vergnaud (apud MOREIRA, 2002) em que, para construir um campo conceitual, é necessário trabalhar em mais de uma representação do conceito. Ou seja, para construir um conceito (ou raciocinar completamente sobre um conteúdo estatístico) é necessário vivenciar diferentes atividades, em diferentes representações.

A Figura 3 apresenta uma interpretação para a relação das três temáticas apresentadas neste capítulo.

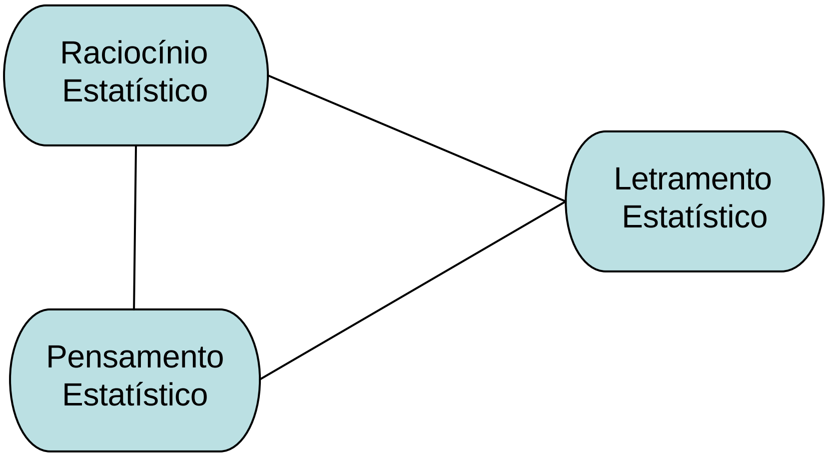



Figura 3 - Relação entre Raciocínio, Pensamento e Letramento estatísticos

Como apresentado na Figura 3, à medida que um indivíduo apresenta um nível de raciocínio mais avançado (segundo o modelo de Garfield, 2002) e pensa estatisticamente (conforme modelo de Wild e Pfannkuch, 1999), seu nível de letramento estatístico será maior. Ou seja, o nível de letramento estatístico é dependente do raciocínio e pensamento estatísticos. Por outro lado, à medida que o nível de letramento estatístico aumenta, o raciocínio e o pensamento estatístico tornam-se mais apurados.

Outra interpretação da autora é que o raciocínio e o pensamento estatísticos são mutuamente relacionados. À medida que um indivíduo apresenta um raciocínio estatístico mais avançado, pode desenvolver também o pensamento estatístico. Do mesmo modo, desenvolvendo o pensamento estatístico pode elevar seu raciocínio estatístico a um nível mais avançado.

Nessa relação entre as temáticas, deve-se destacar que a variação é o elemento central do pensamento estatístico e é um dos conteúdos necessários para que um indivíduo seja estatisticamente letrado. As pesquisas em relação ao raciocínio sobre variação podem colaborar para o desenvolvimento do pensamento e letramento estatísticos.

Para se dar início ao trabalho sobre variação, o Capítulo 2 apresenta uma revisão teórica sobre a temática, o Capítulo 3 apresenta uma análise das tarefas, das técnicas e do suporte teórico do objeto variação nos livros de Matemática do ensino médio e o Capítulo 4 apresenta uma revisão dos estudos publicados acerca da variação e da variabilidade.

## 2 As Medidas de Variação

A medida de variação mais conhecida e utilizada nas análises estatísticas é o desvio padrão. Porém, existem outras medidas tais como a variância, a amplitude total, o desvio médio, o coeficiente de variação e o intervalo interquartílico que compõem o que se chamam de medidas de variação ou medidas de dispersão. Todas essas medidas são importantes em uma análise de dados, pois permitem uma caracterização geral da distribuição da variável observada e, por esse motivo, são apresentadas neste trabalho.

O objetivo deste capítulo é apresentar, teoricamente, cada uma das medidas de variação citadas e discutir a importância de conhecer todas elas e sua representação gráfica de maneira a possibilitar a construção do campo conceitual de variação.

A exceção se faz ao intervalo interquartílico, que devido à falta de tempo, não fez parte do estudo com os professores e, portanto, não está apresentado neste trabalho.

Para dar início a esta apresentação, é necessário fazer uma primeira distinção: entre população e amostra. Segundo Bussab e Morettin (2003, p. 256), “população é o conjunto de todos os elementos ou resultados sob investigação” e amostra é qualquer subconjunto da população.

Na maioria das situações reais não é possível coletar dados de toda a população, ou seja, não é possível encontrar os parâmetros da população, o que justifica a necessidade de selecionar uma amostra.

Segundo Mood, Graybill e Boes (1974), uma amostra aleatória de uma variável aleatória X é uma n-upla ordenada ($X_1$, $X_2$, ..., $X_n$) em que cada $X_i$, i = 1, 2, ...., n é uma observação de X e, portanto, uma variável aleatória com a mesma distribuição de X e duas a duas estatisticamente independentes[7].

Os conceitos de população, amostra e variável aleatória estão tratados a partir de um exemplo de Mood, Graybill e Boes (1974). Considere que dez milhões de sementes (**população**) estão estocadas em um armazém e que cada semente pode gerar uma flor branca ou vermelha. Associando-se à semente que gera flor branca o valor um e à semente que gera flor vermelha o valor zero,

---

[7] Utiliza-se n (minúsculo) para representar o tamanho de uma amostra e N (maiúsculo) no caso de uma população.

pode-se considerar que a variável X (número associado à cor da flor gerada pela semente) é uma **variável aleatória** com distribuição de Bernoulli, com parâmetro p (proporção de flores brancas no armazém) que é desconhecido.

Uma **amostra aleatória** de X, de tamanho n, é composta de n sementes selecionadas do armazém, que pode ser indicada por $(X_1, X_2, ..., X_n)$ em que cada $X_i$ (i = 1, 2, ..., n) pode assumir o valor 0 ou 1, conforme a cor da flor que a semente gera. Assim, cada $X_i$ (i = 1, 2, ..., n) é uma variável aleatória com distribuição de Bernoulli e parâmetro p (o mesmo da população de sementes no armazém) e é independente das demais. Segundo Mood, Graybill e Boes (1974 p. 223),

> depois que a amostra é observada, os valores atuais de $X_1, X_2, \ldots, X_n$ são conhecidos e são denotados por $x_1, x_2, \ldots, x_n$. Algumas vezes as observações $x_1, x_2, ..., x_n$ são chamadas de amostra aleatória se $x_1, x_2, \ldots, x_n$ são os valores de $X_1, X_2, \ldots, X_n$ ,onde $X_1, X_2, \ldots, X_n$ é uma amostra aleatória.

Para a amostra aleatória $(X_1, X_2, \ldots, X_n)$ do exemplo citado, note que é possível obter os valores $(x_1, x_2, ..., x_n)$ tais como (0, 0, ...,0) ou (1, 0, ....,0) ou ainda (0, 0,....,1), etc., de acordo com as cores das flores que as sementes podem gerar.

Nessa explicação de amostra aleatória foi utilizado o termo **parâmetro** que, segundo Bussab e Morettin (2003 p. 265), é "uma medida usada para descrever uma característica da população". No exemplo acima, p é parâmetro e refere-se à *proporção de sementes que produzirão flores brancas* dentre as 10 milhões de sementes.

Estatística é uma variável aleatória, função das variáveis da amostra aleatória, por exemplo, $\frac{X_1+X_n}{2}$, $\frac{X_1+X_2+...+X_n}{n}$, etc. Por ser uma variável aleatória, pode-se falar em sua distribuição de probabilidades.

**Estimador** é uma estatística associada a um parâmetro populacional desconhecido. Assim, o estimador de um parâmetro θ (indicado por $\hat{q}$), "é qualquer função das observações da amostra" e "**estimativa** é o valor numérico do estimador para uma dada amostra aleatória" (Bussab e Morettin, 2003 p. 291).

Vários estimadores podem ser construídos para um mesmo parâmetro; por exemplo, $\frac{X_1+X_n}{2}$ e $\frac{X_1+X_2+...+X_n}{n}$ podem ser estimadores da média populacional. A fim de selecionar um, dentre vários, há a necessidade de introduzir qualidades que um estimador deve satisfazer, a fim de diferenciá-lo entre outros possíveis[8]. Dentre elas, uma refere-se ao fato de ser não viesado (não viciado), isto é, a média do estimador deve coincidir com o parâmetro que ele pretende estimar. Como todo estimador é uma estatística e, como tal, uma variável aleatória, pode-se falar em valor esperado de um estimador, que se denota por $E(\hat{q})$. Assim, um estimador é não viesado quando $E(\hat{q}) = q$ .

Para se tratar de qualidade de um estimador, faz-se necessário o conceito de média ou valor esperado. Para tanto, antes de dar início à apresentação do estimador da variância, é feita uma breve apresentação da média (ou esperança), suas propriedades e seu estimador.

## 2.1 Média

A média de uma variável aleatória X é uma medida de tendência central, também denominada valor esperado de X, e representada por E(X). É definida como a média aritmética dos valores de X ($X_1, X_2, ..., X_N$), isto é, $E(X) = \frac{1}{N}\sum_{i=1}^{N} X_i$ .

Algumas propriedades da média serão utilizadas na discussão sobre o estimador da variância e, por esse motivo, estão apresentadas aqui.

Seja k uma constante real,

**P1-** $E(k) = k$

Demonstração:

$$E(k) = \frac{1}{N}\sum_{i=1}^{N} k = \frac{k}{N}\sum_{i=1}^{N} 1 = \frac{k \cdot N}{N} = k$$

Para exemplificar essa primeira propriedade, suponha o conjunto numérico (7, 7, 7). A média é $E(X) = \frac{1}{3}\sum_{i=1}^{3} X_i = \frac{7+7+7}{3} = 7$.

**P2 -** $E(kX) = k \cdot E(X)$

Demonstração:

$$E(kX) = \frac{1}{N}\sum_{i=1}^{N} (kX_i) = k \cdot \left( \frac{1}{N}\sum_{i=1}^{N} X_i \right) = k \cdot E(X)$$

---

[8] Outras qualidades dos estimadores não são tratadas neste trabalho, por fugir aos propósitos deste estudo.

Por exemplo, para os valores ( 0, 1, 2, 5 e 7), a média é $E(X)=\frac{0+1+2+5+7}{5}=\frac{15}{5}=3$.

Considerando k = 7, obtém-se o novo conjunto numérico (0, 7, 14, 35, 49), em que a média é $E(X)=\frac{0+7+14+35+49}{5}=\frac{105}{5}=21$. Observando-se que 21 = (7).3, nota-se que $E(7X)=7\cdot E(X)$.

**P3** - $E(X+k)=E(X)+k$

Demonstração:

$$E(X+k)=\frac{1}{N}\sum_{i=1}^{N}(X_i+k)=\frac{1}{N}\left(\sum_{i=1}^{N}X_i+\sum_{i=1}^{N}k\right)=\frac{1}{N}\sum_{i=1}^{N}X_i+\frac{1}{N}k\sum_{i=1}^{N}1=E(X)+\frac{k\cdot N}{N}=E(X)+k$$

Por exemplo, para os valores ( 0, 1, 2, 5 e 7), a média é $E(X)=\frac{0+1+2+5+7}{5}=\frac{15}{5}=3$.

Considerando k = -7, obtém-se o novo conjunto numérico (-7, -6, -5, -2, 0), em que a média é $E(X)=\frac{-7+(-6)+(-5)+(-2)+0}{5}=\frac{-20}{5}=-4$. Observando-se que -4 = 3 + (-7) , tem-se $E(X-7)=E(X)-7$.

**P4** – No caso em que, para os elementos pesquisados, há interesse em estudar o comportamento conjunto de duas variáveis aleatórias X e Y, tem-se:

Tabela 1 – Tabela de dados de X e Y

| Variável | Elemento | | | | |
|---|---|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ | ... | $E_N$ |
| X | $X_1$ | $X_2$ | $X_3$ | ... | $X_N$ |
| Y | $Y_1$ | $Y_2$ | $Y_3$ | ... | $Y_N$ |
| $X\pm Y$ | $X_1\pm Y_1$ | $X_2\pm Y_2$ | $X_3\pm Y_3$ | ... | $X_N\pm Y_N$ |

Nesses casos, $E(X\pm Y)=E(X)\pm E(Y)$ (Por indução, essa propriedade pode ser estendida para um número finito de variáveis aleatórias).

Demonstração:

$$E(X\pm Y)=\frac{1}{N}\cdot\sum_{i=1}^{N}(X_i\pm Y_i)=\frac{1}{N}\cdot\left(\sum_{i=1}^{N}(X_i)\pm\sum_{i=1}^{N}(y_i)\right)=\frac{1}{N}\sum_{i=1}^{N}(X_i)\pm\frac{1}{N}\sum_{i=1}^{N}(y_i)=E(X)\pm E(Y)$$

Para exemplificar, considere os valores apresentados na Tabela 2.

Tabela 2 – Tabela de dados da variável X, Y e (X + Y).

| Variável | Elemento | | | | |
|---|---|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ |
| X | 0 | 1 | 2 | 5 | 7 |
| Y | 3 | 5 | 7 | 9 | 4 |
| X + Y | 3 | 6 | 9 | 14 | 11 |

A média da variável X é $E(X) = \frac{0+1+2+5+7}{5} = \frac{15}{5} = 3$ e a média da variável Y é $E(Y) = \frac{3+5+7+9+4}{5} = \frac{28}{5} = 5{,}6$. Utilizando o somatório das esperanças, tem-se $E(X)+E(Y) = 3+5{,}6 = 8{,}6$.

A nova variável (X + Y) assume os valores (3, 6, 9, 14, 11), cuja média é $E(X+Y) = \frac{3+6+9+14+11}{5} = \frac{43}{5} = 8{,}6$.

Assim sendo, $E(X+Y) = 8{,}6$ que é o mesmo resultado de $E(X)+E(Y)$. Conforme Magalhães e Lima (2004, p. 145), “para as variáveis X e Y, vale sempre que $E(X+Y) = E(X)+E(Y)$”.

**P5** – Antes de enunciar a quinta propriedade, observe os três exemplos a seguir.

Tabela 3 - Tabela de dados da variável X,Y e X.Y, Exemplo 1

| Variável | Elemento | | | | |
|---|---|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ |
| X | 0 | 1 | 2 | 5 | 7 |
| Y | 3 | 5 | 7 | 9 | 4 |
| X.Y | 0 | 5 | 14 | 45 | 28 |

Conforme cálculos anteriores, E(X) = 3 e E(Y) = 5,6.

A variável $(X \cdot Y)$ assume os valores (0, 5, 14, 45, 28), cuja média é $E(X \cdot Y) = \frac{0+5+14+45+28}{5} = \frac{92}{5} = 18{,}4$.

Note que E(X) . E(Y) = 3 (5,6) = 16,8 e, portanto, $E(X \cdot Y) \neq E(X) \cdot E(Y)$

Tabela 4 - Tabela de dados da variável X, Y e X.Y, Exemplo 2

| Variável | Elemento | | | | |
|---|---|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ |
| X | 1 | 2 | 3 | 4 | 5 |
| Y | 2 | 0 | 2 | 0 | 2 |
| X.Y | 2 | 0 | 6 | 0 | 10 |

Nesse caso, tem-se que:

$E(X) = \frac{1+2+3+4+5}{5} = 3$ , $E(Y) = \frac{2+0+2+0+2}{5} = \frac{6}{5} = 1{,}2$ , logo E(X).E(Y) = 3,6. A variável $(X \cdot Y)$ assume os valores (2, 0, 6, 0, 10) cuja média é $E(X.Y) = \frac{2+0+6+0+10}{5} = \frac{18}{5} = 3{,}6$ e, portanto, $E(X \cdot Y) = E(X) \cdot E(Y)$.

Tabela 5 - Tabela de dados da variável X, Y e X.Y, Exemplo 3

| Variável | Elemento | | | | | |
|---|---|---|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ | $E_6$ |
| X | 1 | 1 | 2 | 2 | 3 | 3 |
| Y | 5 | 6 | 5 | 6 | 5 | 6 |
| X.Y | 5 | 6 | 10 | 12 | 15 | 18 |

Nesse caso, tem-se que:

$E(X) = \frac{1+1+2+2+3+3}{6} = 2$ ,$E(Y) = \frac{5+6+5+6+5+6}{6} = 5{,}5$, logo E(X).E(Y) = 11. A variável $(X \cdot Y)$ assume os valores (5, 6, 10, 12, 15, 18) cuja média é $E(X.Y) = \frac{5+6+10+12+15+18}{6} = \frac{66}{6} = 11$ e, portanto, $E(X \cdot Y) = E(X) \cdot E(Y)$.

Conclui-se que, diferentemente do que acontece com a média da soma ( ou diferença) de duas variáveis aleatórias, a média do produto nem sempre coincide com o produto das médias. Magalhães e Lima (2004, p. 146) acrescentam que "se X e Y independentes, então $E(X \cdot Y) = E(X) \cdot E(Y)$, no entanto, $E(X \cdot Y) = E(X) \cdot E(Y)$ não implica X e Y independentes[9]".

Para que a propriedade $E(X \cdot Y) = E(X) \cdot E(Y)$ seja válida é preciso que $\sum_{i=1}^{N} X_i . Y_i = \frac{\sum_{i=1}^{N} X_i . \sum_{i=1}^{N} Y_i}{N}$.

De fato:

$$E(X.Y) = \frac{1}{N} \sum_{i=1}^{N} X_i Y_i = \frac{\sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i}{N} = \frac{1}{N} \sum_{i=1}^{N} X_i . \frac{1}{N} \sum_{i=1}^{N} Y_i = E(X).E(Y)$$

Observe que nos Exemplos 2 e 3 tem-se que a propriedade é verdadeira, enquanto que no exemplo 1 ela não o é; acrescenta-se, ainda, que no Exemplo 3 as variáveis são independentes, o que não acontece no exemplo 2, ilustrando os comentários pré-citados de Magalhães e Lima (2004, p.146) referentes a essa questão.

Além das propriedades da média, o seu estimador não viesado também será utilizado na discussão do estimador para a variância. Considere uma amostra aleatória ($X_1$, $X_2$, ..., $X_n$) de uma variável aleatória X, com determinada distribuição de probabilidades, cuja média seja $E(X) = \mu$ (média). Pela definição de amostra aleatória, tem-se que cada $X_i$ (i = 1, 2, ..., n) tem a mesma distribuição de X, portanto, $E(X_i) = \mu$. Um estimador de $\mu$ é $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i$, que é não viesado. De fato,

$$E(\hat{\mu}) = E\left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} \sum_{i=1}^{n} \mu = \frac{1}{n} \cdot n \cdot \mu = \mu .$$

Conforme notação usual, adota-se $\overline{X}$ (média amostral) para representar o estimador de $\mu$ e $\overline{x}$ para a estimativa correspondente. Por exemplo, se os valores de uma amostra aleatória são (2, 2, 3, 4, 4), então, $\overline{x} = \sum_{i=1}^{5} \frac{x_i}{5} = \frac{2+2+3+4+4}{5} = 3$ é uma estimativa para a média populacional da qual a referida amostra foi extraída.

[9] O Apêndice 1 apresenta uma breve explicação sobre independência de variáveis.

As propriedades da média e o estimador não viesado da média são conceitos necessários para a discussão sobre o estimador não viesado da variância, que está apresentada no subcapítulo seguinte.

## 2.2 Variância

A variância de uma variável aleatória X é uma medida de dispersão dos valores da variável em torno da sua média. Ela é definida como a média aritmética dos quadrados dos desvios dos valores de X ($X_1$, $X_2$, ..., $X_N$) em relação à média dos próprios valores de X. Em uma população de tamanho N, a variância de uma variável aleatória X, denotada por Var(X), é dada por:

$\mathrm{Var}(X) = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$ , onde $\mu$ é a média da variável X.

A variância mede a distância em torno da média e não em torno de outra medida-resumo, devido ao fato de que a média minimiza o erro quadrático médio[10]. James (2004 p. 126) faz esta demonstração:

> "Proposição:
>
> Seja X integrável, $\mu = E(X)$. Então, $\mu$ minimiza $E(X-c)^2$ , $c \in \Re$, isto é, Var (X) = $E(X-\mu)^2 = \min_{c\in\Re} E(X-c)^2$
>
> Prova:
>
> $(X-c)^2 = (X-\mu+\mu-c)^2 = (X-\mu)^2 - 2(\mu-c)(X-\mu) + (\mu-c)^2$
>
> Pela linearidade da esperança [propriedade 4], temos:
>
> $E(X-c)^2 = E(X-\mu)^2 - 2(\mu-c)\cdot E(X-\mu) + (\mu-c)^2 = \mathrm{Var}X + (\mu-c)^2$
>
> Conclusão: $E(X-c)^2 \geq E(X-\mu)^2, \forall c \in \Re$ ".

Como apresentado na demonstração, a média minimiza o erro quadrático médio e, por esse motivo, a variância e conseqüentemente o desvio padrão (que está apresentado no sub capitulo seguinte) são medidas de dispersão em torno da média.

A variância, $\mathrm{Var}(X) = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$ , pode sofrer um tratamento algébrico e pode ser obtida da seguinte maneira: $\mathrm{Var}(X) = \frac{1}{N}\sum_{i=1}^{N}X_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N}X_i\right)^2$ .

10 Utilizar o erro quadrático médio é uma questão de escolha e, neste caso, a média é a medida que minimiza este erro.

Demonstração:

$$Var(X) = E\left[(X-\mu)^2\right] = E\left[X^2 - 2X\mu + \mu^2\right] = E\left(X^2\right) - E(2X\mu) + E\left(\mu^2\right) =$$
$$= E\left(X^2\right) - 2\mu E(X) + \mu^2 = E\left(X^2\right) - 2\mu^2 + \mu^2 = E\left(X^2\right) - \mu^2 = E\left(X^2\right) - [E(X)]^2$$

Logo,

$Var(X) = E\left(X^2\right) - [E(X)]^2 = \frac{1}{N}\sum_{i=1}^{N} X_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} X_i\right)^2$, ou seja, a variância de uma variável aleatória X, que assume os valores ($X_1$, $X_2$, ..., $X_N$), também pode ser obtida calculando-se a média dos quadrados dos valores da variável menos o quadrado da média de tais valores.

Observe os dois exemplos a seguir, com relação às vantagens de cada uma das duas fórmulas da variância:

Exemplo 1: Considere a variável aleatória X que assume os valores 0, 1, 1, 1, 1, 1, 1, 2.

$$E(X) = \frac{1}{8}\cdot\sum_{i=1}^{8} X_i = \frac{1}{8}\cdot 8 = 1$$

$$Var(X) = \frac{1}{8}\cdot\sum_{i=1}^{8}(X_i - 1)^2 =$$
$$= \frac{(0-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (2-1)^2}{8} = \frac{1}{4}$$

Utilizando a segunda fórmula apresentada para o cálculo da variância, tem-se:

$$Var(X) = \frac{1}{8}\sum_{i=1}^{8} X_i^2 - \left(\frac{1}{8}\sum_{i=1}^{8} X_i\right)^2 = \frac{0^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 2^2}{8} - 1^2 = \frac{10}{8} - 1 = \frac{1}{4}\ .$$

Nesse exemplo, as duas formas levaram ao mesmo resultado, conforme a indicação teórica.

Exemplo 2: Considere a variável aleatória X que assume os valores 2,3; 3,1; 3,7; 4,2; 5,8; 6,8.

$$E(X) = \frac{1}{6}\cdot\sum_{i=1}^{6} X_i = \frac{2,3 + 3,1 + 3,7 + 4,2 + 5,8 + 6,8}{6} = \frac{25,9}{6} \cong 4,32$$

$$Var(X) = \frac{1}{6}\cdot\sum_{i=1}^{6}(X_i - 4,32)^2 =$$
$$= \frac{(2,3-4,32)^2 + (3,1-4,32)^2 + (3,7-4,32)^2 + (4,2-4,32)^2 + (5,8-4,32)^2 + (6,8-4,32)^2}{6} =$$
$$\cong \frac{14,31}{6} \cong 2,39$$

Utilizando a segunda fórmula apresentada para o cálculo da variância, tem-se:

$$Var(X)=\frac{1}{6}\sum_{i=1}^{6}X_i^2-\left(\frac{1}{6}\sum_{i=1}^{6}X_i\right)^2=$$

$$=\frac{1}{6}\cdot(5{,}29+9{,}61+13{,}69+17{,}64+33{,}64+46{,}24)-(4{,}32)^2=\frac{126{,}11}{6}-18{,}66=2{,}36$$

Com a utilização desses dois exemplos é possível discutir as vantagens e desvantagens de cada fórmula. A fórmula $Var(X)=\frac{1}{N}\sum_{i=1}^{N}(X_i-\mu)^2$ é útil, pois modela e exprime a própria definição de variância e, portanto, pode ajudar o aluno a compreender seu significado. Em contrapartida, os cálculos parciais exigem arrendondamentos e, em conseqüência, o resultado final é menos preciso.

Já a fórmula $Var(X)=\frac{1}{N}\sum_{i=1}^{N}X_i^2-\left(\frac{1}{N}\sum_{i=1}^{N}X_i\right)^2$ não recupera o conceito de variância, pois não é um modelo explícito de seu conceito, mas sua grande vantagem é a simplificação dos cálculos e a melhor precisão do resultado final. Essa maneira de calcular a variância também permite uma simplificação na demonstração de suas propriedades, apresentadas a seguir.

Seja k uma constante real,

**P1**- $Var(k)=0$.

Demonstração:

$$Var(k)=E(k^2)-[E(k)]^2=k^2-k^2=0$$

Por exemplo, para os valores (5, 5, 5, 5, 5), a média é 5.

$$Var(X)=\frac{5^2+5^2+5^2+5^2+5^2}{5}-5^2=\frac{125}{5}-25=0$$

**P2 -** $Var(kX)=k^2\cdot Var(X)$.

Demonstração:

$$Var(kX)=E\left[(kX)^2\right]-[E(kX)]^2=E(k^2X^2)-[k\cdot E(X)]^2=$$

$$=k^2E(X^2)-k^2[E(X)]^2=k^2\left\{E(X^2)-[E(X)]^2\right\}=k^2Var(X)$$

Por exemplo, para os valores ( 0, 1, 2, 5 e 7), a média é 3 e $Var(X)=\frac{0^2+1^2+2^2+5^2+7^2}{5}-3^2=\frac{79}{5}-9=6{,}8$.

Considerando k = 7, obtém-se o novo conjunto numérico (0, 7, 14, 35, 49), com média 21 e $\mathrm{Var}(X) = \dfrac{0^2 + 7^2 + 14^2 + 35^2 + 49^2}{5} - 21^2 = \dfrac{3871}{5} - 441 = 333{,}2$.

Observando-se que 333,2 = (49).6,8, nota-se que $\mathrm{Var}(7X) = 7^2 \cdot \mathrm{Var}(X)$.

**P3 -** $\mathrm{Var}(X + k) = \mathrm{Var}(X)$, em que k é uma constante.

Demonstração:

$$\mathrm{Var}(X+k) = E\left[(X+k)^2\right] - \left[E(X+k)\right]^2 = E\left(X^2 + 2kX + k^2\right) - \left[E(X) + k\right]^2 =$$
$$= \left[E\left(X^2\right) + 2kE(X) + E\left(k^2\right)\right] - \left\{\left[E(X)\right]^2 + 2kE(X) + k^2\right\} = E\left(X^2\right) - \left[E(X)\right]^2 = \mathrm{Var}(X)$$

Por exemplo, para os valores ( 0, 1, 2, 5 e 7), a média é 3 e a variância é 6,8. Considerando $k = -7$, obtém-se um novo conjunto numérico (-7, -6, -5, -2, 0), cuja média é -4 e a variância é $\mathrm{Var}(X) = \dfrac{(-7)^2 + (-6)^2 + (-5)^2 + (-2)^2 + 0^2}{5} - (-4)^2 = \dfrac{114}{5} - 16 = 6{,}8$, exatamente a mesma do conjunto original. Assim, Var (X + k) = Var (X).

**P4 -** Sejam X e Y duas variáveis aleatórias que assumem os valores ($X_1$, $X_2$, ..., $X_n$) e ($Y_1$, $Y_2$, ..., $Y_n$), respectivamente. Se estas duas variáveis são independentes, então $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Demonstração:

$$\mathrm{Var}(X \pm Y) = E\left[(X \pm Y)^2\right] - \left[E(X \pm Y)\right]^2 = E\left(X^2 \pm 2XY + Y^2\right) - \left[E(X) \pm E(Y)\right]^2 =$$
$$= \left[E\left(X^2\right) \pm 2E(XY) + E\left(Y^2\right)\right] - \left\{\left[E(X)\right]^2 \pm 2E(X)E(Y) + \left[E(Y)\right]^2\right\} =$$
$$= E\left(X^2\right) \pm 2E(X)E(Y) + E\left(Y^2\right) - \left[E(X)\right]^2 \pm 2E(X)E(Y) - \left[E(Y)\right]^2 =$$
$$= E\left(X^2\right) - \left[E(X)\right]^2 + E\left(Y^2\right) - \left[E(Y)\right]^2 = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Para exemplificar essa propriedade, retome-se os Exemplos 1 e 3 da Propriedade 5 da média, em que o primeiro diz respeito a duas variáveis dependentes e o segundo (Exemplo 3) diz respeito a duas variáveis independentes.

Sabendo-se que a variável X ( 0, 1, 2, 5 e 7), que tem média 3 e variância 6,8 não é independente da variável Y (3, 5, 7, 9, 4), que tem média 5,6 e $\mathrm{Var}(Y) = \dfrac{3^2 + 5^2 + 7^2 + 9^2 + 4^2}{5} - 5{,}6^2 = \dfrac{180}{5} - 31{,}36 = 4{,}64$, tem-se que:

$\mathrm{Var}(X) + \mathrm{Var}(Y) = 6{,}8 + 4{,}64 = 11{,}44$.

Como apresentado na Tabela 2, (X + Y) = (3, 6, 9, 14, 11), e que a média de (X + Y) é 8,6.

Portanto, $\mathrm{Var}(X+Y)=\dfrac{3^2+6^2+9^2+14^2+11^2}{5}-8{,}6^2=\dfrac{443}{5}-73{,}96=14{,}64$.

De acordo com essa propriedade, como as variáveis X e Y não são independentes, $\mathrm{Var}(X+Y)\neq \mathrm{Var}(X)+\mathrm{Var}(Y)$.

Tomando-se o Exemplo 3, em que a variável X (1, 1, 2, 2, 3, 3), com média aritmética 2 e $\mathrm{Var}(X)=\dfrac{1^2+1^2+2^2+2^2+3^2+3^2}{6}-2^2=\dfrac{28}{6}-4=\dfrac{4}{6}$ é independente da variável Y (5, 6, 5, 6, 5, 6), com média 5,5 e $\mathrm{Var}(Y)=\dfrac{5^2+6^2+5^2+6^2+5^2+6^2}{6}-5{,}5^2=\dfrac{183}{6}-30{,}25=0{,}25$.

Então, $\mathrm{Var}(X)+\mathrm{Var}(Y)=\dfrac{4}{6}+0{,}25=\dfrac{5{,}5}{6}$.

A nova variável (X + Y) é (6, 7, 7, 8, 8, 9) que tem média 7,5 e variância $\mathrm{Var}(X+Y)=\dfrac{6^2+7^2+7^2+8^2+8^2+9^2}{6}-7{,}5^2=\dfrac{343}{6}-56{,}25=\dfrac{5{,}5}{6}$.

Como já se sabe que X e Y são independentes, então $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$.

Conforme notação usual, adota-se $\sigma^2$ para representar Var (X) e $\hat{s}^2$ para representar o estimador de $\sigma^2$ ou Var (X).

Voltando à amostra aleatória ($X_1$, $X_2$, ..., $X_n$) de uma variável aleatória X, citada anteriormente, com determinada distribuição de probabilidades, e sabendo-se que $\hat{\mu}=\overline{X}$ é um estimador não viesado de $\mu$ (média), o objetivo é verificar se $\hat{\sigma}^2=S^2{}'=\dfrac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\dfrac{1}{n}\sum_{i=1}^{n}X_i^2-\overline{X}^2$ (nas duas conotações apresentadas anteriormente) é um estimador não viesado para $\mathrm{Var}(X)=\sigma^2$ ,isto é, se $E(S^2{}')=\sigma^2$.

A esperança desse estimador é:

$$E(S'^2)=E\left(\frac{1}{n}\cdot\sum_{i=1}^{n}X_i^2-\overline{X}^2\right)=\frac{1}{n}\cdot\sum_{i=1}^{n}E(X_i^2)-E\left(\overline{X}^2\right)$$ **(1).**

Sabe-se que $Var(X_i) = E(X_i^2) - [E(X_i)]^2 \Rightarrow E(X_i^2) = Var(X_i) + [E(X_i)]^2$ **(2)** e que $Var(\overline{X}) = E\left(\overline{X}^2\right) - \left[E(\overline{X})\right]^2 \Rightarrow E\left(\overline{X}^2\right) = Var(\overline{X}) + \left[E(\overline{X})\right]^2$ **(3)**

No que diz respeito à variância da média, $Var(\overline{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$ e pelas propriedades P2 e P4 da variância tem-se que $Var(\overline{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} \cdot Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$ **(4)**.

Substituindo (2), (3) e (4) em (1), tem-se:

$$E(S'^2) = \frac{1}{n} \cdot \sum_{i=1}^{n} E(X_i^2) - E\left(\overline{X}^2\right) =$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \left\{Var(X_i) + [E(X_i)]^2\right\} - \left(\frac{\sigma^2}{n} + \mu^2\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\sigma^2 + \mu^2) - \frac{\sigma^2}{n} - \mu^2 =$$

$$= \frac{1}{n} \cdot n \cdot (\sigma^2 + \mu^2) - \frac{\sigma^2}{n} - \mu^2 =$$

$$= \sigma^2 - \frac{\sigma^2}{n} =$$

$$= \sigma^2\left(1 - \frac{1}{n}\right) = \sigma^2 \cdot \frac{n-1}{n}$$

Portanto, $S^{2\prime}$ é um estimador viesado de $\sigma^2$ pois $E(S^2{}') = \frac{n-1}{n} \cdot \sigma^2$. Como o objetivo é determinar um estimador não viesado de $\sigma^2$, define-se um novo estimador $S^2 = \frac{n}{n-1} S^2{}'$ e verifica-se se é não viesado.

$E(S^2) = E\left(\frac{n}{n-1} S^2{}'\right) = \frac{n}{n-1} \cdot E(S^2{}') = \frac{n}{n-1} \cdot \frac{\sigma^2(n-1)}{n} = \sigma^2$ permite concluir que esse novo estimador da variância é não viesado.

Assim, a variância da amostra é definida por

$$S^2 = \frac{n}{n-1} \cdot S^2{}' = \begin{cases} \frac{n}{n-1} \cdot \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \ \_\text{ou} \\ \frac{n}{n-1} \cdot \left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2\right) = \frac{1}{n-1} \cdot \left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right) \end{cases}$$

a fim de ser um **estimador não viesado** para a variância populacional ( $\sigma^2$ ).

E a estimativa não viesada da variância populacional[11] é dada por:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \quad \text{ou} \quad s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$$

Como exemplo, sejam os valores (0, 1, 2, 5 e 7) uma amostra aleatória da variável aleatória 'notas de uma turma de Matemática', cujos valores podem estar no intervalo [0, 10]. Esse conjunto numérico tem média 3 pontos e a estimativa da variância é:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 = \frac{(0-3)^2 + (1-3)^2 + (2-3)^2 + (5-3)^2 + (7-3)^2}{5-1} = \frac{34}{4} = 8{,}5$$

pontos ao quadrado.

Observa-se que a média guarda a mesma unidade de medida da variável observada (pontos de nota), enquanto que a variância utiliza essa unidade elevada ao quadrado. Esse fato dificulta sua interpretação e justifica a utilização do desvio padrão, que é a raiz quadrada positiva da variância e permite comparar a variação das observações na mesma unidade de medida da média aritmética.

Como lembra Costa Neto (2002, p. 27), embora a variância tenha esse inconveniente de interpretação, "ela é extremamente importante na teoria estatística".

## 2.3 Desvio Padrão

O desvio padrão de uma variável aleatória X é a raiz quadrada positiva da variância. O desvio padrão da população (parâmetro) é obtido da seguinte forma:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}} \quad \text{ou} \quad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} X_i^2 - \mu^2} \quad \text{ou simplesmente} \quad \sigma = \sqrt{\sigma^2}\,.$$

Quando se trata de uma amostra (estimativa), o desvio padrão é obtido da seguinte maneira:

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}} \quad \text{ou} \quad s = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)} \quad \text{ou ainda} \quad s = \sqrt{s^2}$$

[11] Em situações em que a média é conhecida, o estimador da variância fica dividido por n e não por n-1. Definindo $S^2 = \frac{1}{n}\left[\sum_{i=1}^{n}(X_i)^2 - n\mu^2\right]$, então, $E(S^2) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i^2\right) - E\left(\mu^2\right) =$

$$= \frac{1}{n}\sum_{i=1}^{n}\left\{Var(X_i) + \left[E(X_i)^2\right]\right\} - \mu^2 = = \frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2 + \mu^2\right) - \mu^2 = \frac{1}{n}n\left(\sigma^2 + \mu^2\right) - \mu^2 = \sigma^2\,.$$

O desvio padrão é útil para avaliar a homogeneidade das observações em relação à média aritmética. Quanto menor o valor do desvio padrão, mais homogêneas são as observações em relação à média, o que indica maior densidade das observações próximas à média aritmética.

Bussab e Morettin (2003) explicam que tanto o desvio médio (que está apresentado na seção 2.6 deste trabalho) quanto o desvio padrão indicam, em média, qual será o erro (desvio) cometido ao tentar substituir cada observação pela medida resumo da distribuição (que, no caso, é a média aritmética).

No exemplo das notas (0, 1, 2, 5, 7), para se obter o desvio padrão da população, basta calcular $\sqrt{\sigma^2} = \sqrt{6{,}8} \cong 2{,}61$. Então, os parâmetros da população de notas são $\mu = 3$ e $\sigma = 2{,}61$.

Se o conjunto de notas se refere a uma amostra aleatória, as estimativas não viesadas da média e do desvio padrão populacional são $\overline{x} = 3$ e $s = 2{,}92$ $\left(s = \sqrt{s^2} = \sqrt{8{,}5} \cong 2{,}92\right)$.

O livro de Bussab e Morettin (2003, p.40) apresenta a variância e o desvio padrão populacional no capítulo 3 intitulado Medidas-Resumo e citam que “a variância de uma amostra será calculada usando-se o denominador n -1, em vez de n” e que isto será explicado num capítulo posterior, capítulo 11 intitulado Estimação, quando discutem as propriedades dos estimadores. O livro de Martins (2001), apresenta a variância amostral no capítulo 2 intitulado Estatística Descritiva e, no capítulo 6 intitulado Distribuições Amostrais, explica que trata-se de um estimador da variância populacional, sem apresentar a demonstração.

Ambos os livros de Estatística tratam dos estimadores após o capítulo de Inferência Estatística, sendo que este último não é um assunto sugerido para o ensino médio. Fica, então, a critério do autor do livro didático de Matemática do Ensino Médio decidir a maneira como irá apresentar a variância (e conseqüentemente o desvio padrão) assim como fica a critério do professor de Matemática a decisão sobre a estratégia que irá utilizar para discuti-la com seus alunos. Daí, uma das dificuldades dos conteúdos variância e desvio padrão.

Outra dificuldade relacionada à variância (e conseqüentemente o desvio padrão), já salientada por Reading e Shaughnessy (2004), é a complexidade dos cálculos dessas medidas, que se tornam obstáculos para motivar os alunos em aulas de Matemática na escola. Para que um aluno do ensino médio efetue o cálculo de tais medidas estatísticas, ele deve mobilizar muitos conceitos matemáticos, tais como potência, radiciação, fração e o próprio conhecimento do símbolo de somatório.

Uma maneira de justificar a dificuldade de calcular a variância e o desvio padrão é a discussão sobre a aplicação dessas medidas, ou seja, a interpretação dos cálculos segundo o contexto apresentado. É possível que a interpretação do resultado de um cálculo matemático complexo em função de um contexto no qual os dados foram apresentados possa ser uma tarefa cognitivamente custosa para um aluno de ensino médio, mas pode propiciar não apenas uma aprendizagem significativa desses conceitos, mas também a construção do campo conceitual de variação.

Green apud Reading e Shaughnessy (2004, p.204) explica que "professores e alunos podem conhecer os procedimentos para calcular o desvio padrão; mas é possível que eles não consigam explicar o que significa, ou por quê ou quando é uma boa medida para a variação esperada".

Para tornar o conceito significativo para o aluno, a apresentação da análise do desvio padrão é fundamental. Hart (1984) e Loosen, Lioen e Lacante (1985) defendem a análise do desvio padrão a partir da porcentagem de observações que estão a ***k*** desvios da média.

Para a análise do desvio padrão, segundo a sugestão desses autores, é possível que um número pequeno de observações não permita ao aluno compreender a porcentagem de valores compreendidos no intervalo $(\bar{x} \pm ks)$.

Por este motivo, a partir deste ponto será utilizada uma amostra de tamanho maior, a partir do seguinte exemplo: uma amostra aleatória de 50 alunos de uma escola, cuja variável aleatória é a altura, medida em centímetros. Observe-se que esse mesmo tipo de mudança pode ser seguido pelo professor de Matemática do ensino médio na organização das atividades didáticas que pretende desenvolver com seus alunos, o que indica que o número de elementos da amostra é uma variável didática importante, de acordo com a Teoria das Situações Didáticas desenvolvidas por Brousseau (1997)[12].

A escolha didática de aumentar o tamanho da amostra pode ocasionar dificuldades em relação aos procedimentos para a obtenção do valor do desvio padrão, o que pode ser contornado a partir da utilização de uma calculadora científica ou o uso de uma planilha eletrônica ou ainda o cálculo do desvio padrão com dados agrupados. Em contrapartida, aumentar o tamanho da amostra possibilita a compreensão do conceito de desvio padrão.

A altura dos cinquenta alunos da amostra está apresentada na Tabela 6.

Tabela 6 - Rol das alturas (em cm) de 50 alunos.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 142 | 147 | 149 | 152 | 153 | 154 | 156 | 157 | 158 | 160 |
| 143 | 148 | 150 | 152 | 154 | 155 | 156 | 157 | 159 | 160 |
| 144 | 148 | 150 | 152 | 154 | 155 | 156 | 157 | 159 | 161 |
| 145 | 149 | 151 | 153 | 154 | 155 | 156 | 157 | 159 | 161 |
| 145 | 149 | 152 | 153 | 154 | 155 | 156 | 158 | 160 | 162 |

Fonte: Tiboni, 2002 p. 73

Para esses valores, a média é $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{7682}{50} = 153{,}64$ centímetros e o desvio padrão é $s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n-1}} = \sqrt{\frac{1213{,}52}{49}} = 4{,}97651628006125$ centímetros[13]. Devido à seqüência de cálculos realizados para a obtenção da estimativa do desvio padrão, geralmente o resultado numérico tem muitas casas decimais e necessita de aproximação. Logo, $s \cong 4{,}98$ centímetros.

A partir das estimativas da altura da população ($\overline{x} = 153{,}64$ e s = 4,98), escolhe-se o número de desvios-padrão e elabora-se o intervalo a partir da média mais ou menos *k* desvios-padrão. Para exemplificar, foi elaborado o intervalo de dois desvios-padrão da média: $[(153{,}64 - 2 \cdot 4{,}98),(153{,}64 + 2 \cdot 4{,}98)] = [143{,}68;163.60]$.

Em uma situação didática, o professor pode solicitar a seguinte tarefa: contar o número de alunos (observações) que estão entre 143,68 cm e 163,60 cm (intervalo de dois desvios-padrão da média). A partir da contagem, o professor pode solicitar que os alunos encontrem a porcentagem de alunos no intervalo.

---

[12] Segundo Artigue (1995), as variáveis microdidáticas dizem respeito às atividades que vão compor a organização didática feita pelo professor.

[13] Optou-se em utilizar duas casas decimais para apresentar o resultado do desvio padrão 4,97651628006125, arredondando-o para 4,98, por ser suficiente para os objetivos deste estudo.

No exemplo, é possível notar que há dois alunos cujas alturas estão fora do intervalo, ou seja, dois alunos cujas alturas são inferiores a 143,68 cm. Então, há quarenta e oito alunos (dentre cinqüenta) com alturas no intervalo $[143{,}68;163{,}60]$, o que corresponde a 96% dos alunos.

A partir da obtenção do intervalo de *k* desvios padrão da média, o aluno pode elaborar o gráfico para representar tal intervalo. Para o caso em que *n* = 2, cujo intervalo é [143,68; 163,60] a Figura 4 apresenta o gráfico.

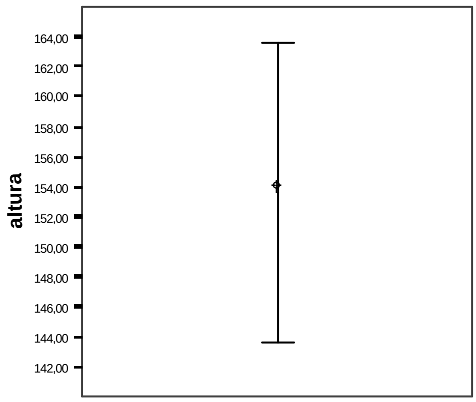



Figura 4 –Intervalo de dois desvios padrão da média da altura dos alunos

Na Figura 4, o ponto central do segmento de reta representa a média aritmética e os seus extremos os limites do intervalo de *k* desvios padrão da média.

O gráfico apresentado na Figura 4 é vantajoso quando se têm duas distribuições de observações, pois possibilita comparar a homogeneidade em torno da média, ou seja, permite visualizar a distribuição que tem o menor intervalo em torno da média.

Outra maneira de trabalhar a densidade dos valores em torno da média é a elaboração do gráfico de pontos (*dotplot*), que Bussab e Morettin (2003) denominam de gráfico de dispersão unidimensional. A Figura 5 apresenta um gráfico de pontos para o exemplo das alturas dos 50 alunos.

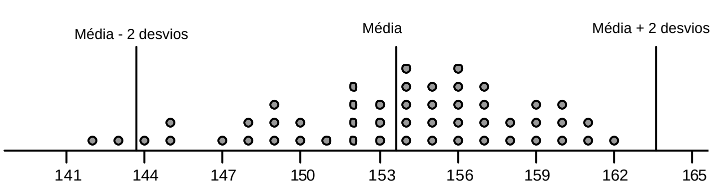



Figura 5 – Gráfico de dispersão da altura dos 50 alunos

Para elaborar um gráfico de pontos, tal como o da Figura 5, cada observação é representada por um ponto numa reta (com a escala dos valores da variável) e se há mais de uma observação com o mesmo valor, eles são “empilhados”. Para se observar a densidade dos pontos em torno da média, pode ser traçada um segmento de reta, perpendicular à reta dos valores da variável, no valor que representa a média aritmética e os limites do intervalo $\left(\bar{x} \pm 2s\right)$.

A vantagem da elaboração do gráfico de pontos é a possibilidade de visualizar as observações (o que não é possível no gráfico do intervalo de *n* desvios padrão da média), mas sua elaboração pode se tornar trabalhosa para variável contínua[14], se feita manualmente.

Uma alternativa é a elaboração do histograma[15], gráfico utilizado para representar variáveis contínuas. A Figura 6 apresenta o histograma para o exemplo da altura dos 50 alunos, em que foram traçados segmentos de reta representando a média aritmética e os limites do intervalo $\left(\bar{x} \pm 2s\right)$.

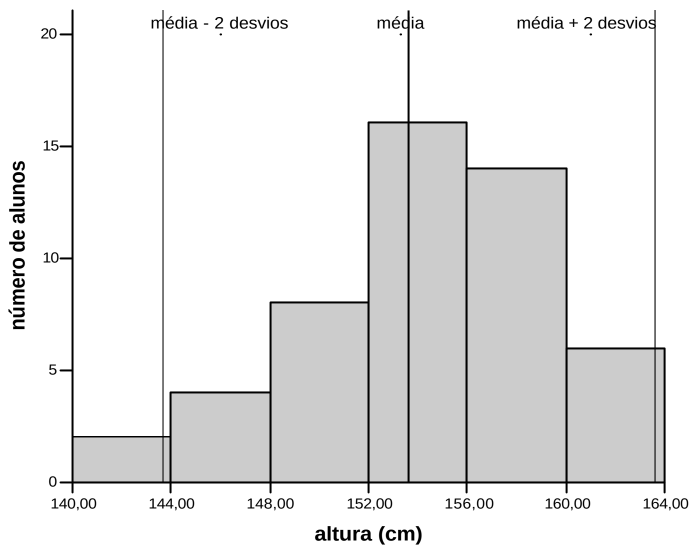



Figura 6 - Distribuição da freqüência dos alunos segundo a altura

[14] No exemplo, altura é uma variável contínua, mas foi medida de maneira que possibilitou trabalhá-la como uma variável discreta.

[15] Não é o objetivo deste trabalho discutir a elaboração do histograma. Uma discussão didática do assunto pode ser obtida em Cazorla e Santana (2006).

O histograma não permite visualizar claramente que apenas duas observações (dois alunos) não estão compreendidos no intervalo $(\bar{x} \pm 2s)$, como é possível notar no gráfico de pontos, mas será a alternativa em inúmeras situações cuja variável é contínua. Como já salientado por Meletiou e Lee (2002), mesmo os alunos de graduação apresentam dificuldades na compreensão desse tipo de gráfico, mas vale ressaltar sua importância na estimação do formato da distribuição.

A tarefa de contar o número de observações contida em um intervalo de $(\bar{x} \pm ks)$ pode ser exaustiva se esse número é muito grande (mil, um milhão, etc), justificando-se assim a utilização de estimativas da porcentagem de observações a menos de *n* desvios da média.

Uma maneira de estimar tal intervalo é o Teorema de Tchebichev[16] (como já defendido por Loosen, Lioen e Lacante, 1985). Este teorema pode ser utilizado com alunos do ensino médio, pois pode ser aplicado para qualquer conjunto numérico, independente de sua distribuição de probabilidades. O teorema diz:

> A proporção (ou fração) de qualquer conjunto de dados a menos de k desvios-padrão a contar da média é sempre ao menos $1-\frac{1}{k^2}$, onde k é um número positivo maior do que 1. Para k = 2 e k = 3, temos os seguintes resultados específicos:
> •Ao menos ¾ (ou 75%) de todos os valores estão no intervalo que vai de 2 desvios-padrão abaixo da média a 2 desvios-padrão acima da média ($\bar{x}-2s$ a $\bar{x}+2s$).
> •Ao menos 8/9 (ou 89%) de todos os valores estão no intervalo que vai de 3 desvios-padrão abaixo da média até 3 desvios-padrão acima da média ($\bar{x}-3s$ a $\bar{x}+3s$). (TRIOLA, 1999, p. 43)

Segundo o Teorema de Tchebichev, pelo menos 75% dos alunos têm altura entre 143,68 cm e 163,60 cm, o que representa uma estimativa da porcentagem de observações no intervalo $(\bar{x} \pm 2s)$.

É importante salientar que o Teorema de Tchebichev afirma que, pelo menos, 75% das observações estão no intervalo $(\bar{x} \pm 2s)$ e, nesse exemplo, foi obtido 96% dos alunos com alturas neste intervalo.

Essa diferença entre a porcentagem estimada e a porcentagem observada é devido a dois motivos. O primeiro é que, segundo o teorema, existem, pelo

[16] Mood, Graybill e Boes (1974) escrevem Chebyshev; Martins (2001) escreve Tchebycheff.

menos, 75% das observações no intervalo $(\bar{x} \pm 2s)$ e trabalhar com a desigualdade pode não ser tão simples. Quando se lê que pelo menos 75% das observações estão no intervalo, é preciso interpretar que se têm 75% ou mais das observações no intervalo.

A segunda explicação é que o Teorema de Tchebichev é aplicado a qualquer distribuição e, quando conhecida, a porcentagem pode ser melhor estimada, como no caso de distribuições simétricas, em que se pode utilizar a Distribuição Normal.

Quando se sabe que o conjunto de observações segue uma Distribuição Normal[17], sabe-se que cerca de 68,26% dos valores estão a menos de um desvio padrão da média (para cima e para baixo da média); cerca de 95,44% dos valores estão a menos de dois desvios padrão da média e cerca de 99,74% dos valores estão a menos de três desvios padrão da média, conforme pode ser observado na Figura 7.

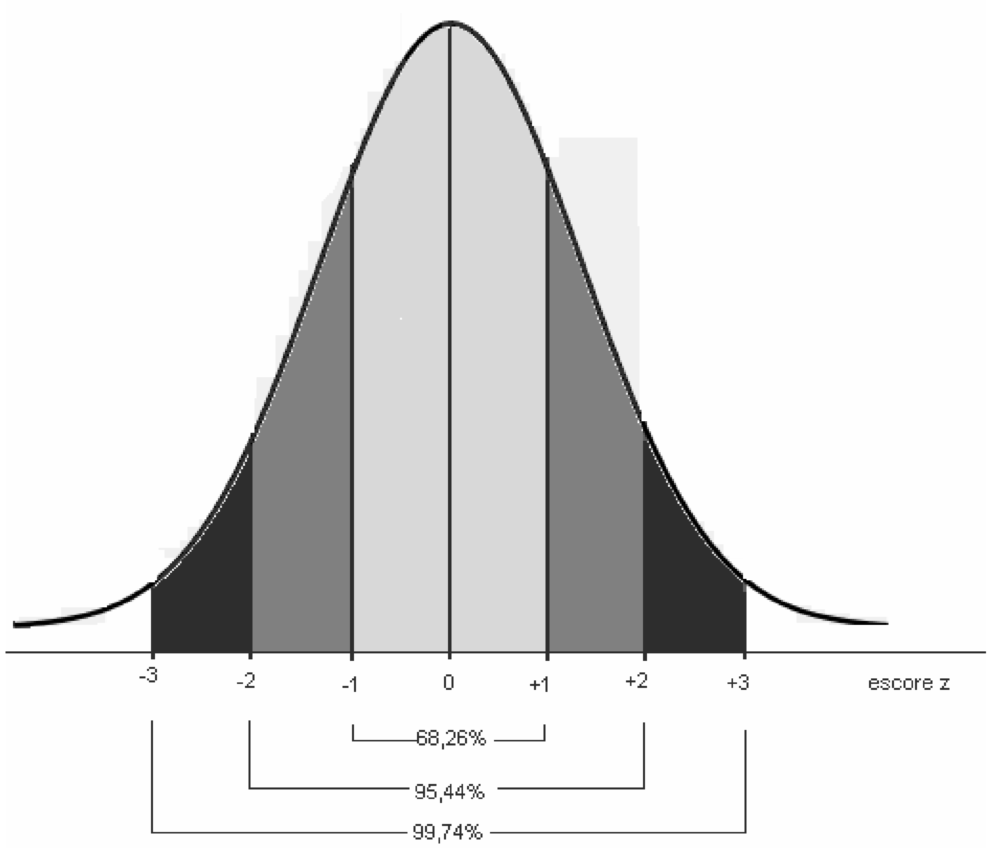



Figura 7 - Porcentagem de observações a menos de um, dois e três desvios padrão da média.

[17] Este trabalho não tem o objetivo de apresentar a Distribuição Normal de Probabilidades, cujo conteúdo está disponível em livros de Estatística. Mas, é importante ressaltar que é possível obter a porcentagem de observações a menos de k desvios padrão, sendo n qualquer número real.

Insere-se mais uma variável didática para ser trabalhada com o aluno do ensino médio: a verificação se a distribuição das observações segue uma distribuição normal.

O histograma, apresentado na Figura 7, pode ser utilizado para estimar, visualmente, o formato da distribuição. Como pode ser observado, a altura dos alunos não é simétrica em relação à média aritmética (condição da Distribuição Normal), mas se aproxima da simetria. Então, pode-se utilizar a porcentagem de observações sob a curva normal para estimar o intervalo das alturas que diferem da média em não mais que dois desvios padrão, já calculado anteriormente: [143,68 ; 163,60].

Utilizando a porcentagem de dados sob a curva normal, é possível dizer que, aproximadamente, 95% dos alunos têm entre 143,68 cm e 163,60 cm de altura.

Voltando à porcentagem observada de alunos com altura no intervalo $(\bar{x} \pm 2s)$, 96% dos alunos têm alturas entre a média mais ou menos dois desvios padrão. Isso indica que, para o exemplo, a estimativa da porcentagem de observações segundo a curva normal é mais próxima da realidade observada na amostra do que a análise segundo o Teorema de Tchebichev. Uma síntese da porcentagem de dados que diferem da média em não mais que dois desvios padrão está apresentada na Tabela 7.

Tabela 7 - Síntese da estimativa da porcentagem de dados no intervalo $(\bar{x} \pm 2s)$

| Porcentagem <u>observada</u> de alunos com alturas no intervalo $(\bar{x} \pm 2s)$ | Porcentagem <u>estimada</u> de alunos com alturas no intervalo $(\bar{x} \pm 2s)$ - Teorema de Tchebichev | Porcentagem <u>estimada</u> de alunos com alturas no intervalo $(\bar{x} \pm 2s)$ - Distribuição Normal de Probabilidade |
|---|---|---|
| 96% | Pelo menos 75% | 95,44% |

O conhecimento da estimativa da porcentagem de observações no intervalo $(\bar{x} \pm 2s)$, seja pelo Teorema de Tchebichev ou pela Distribuição Normal de Probabilidade, pode encorajar o professor de Matemática a desenvolver com seus alunos a tarefa de contagem do número de observações no intervalo de

$(\bar{x} \pm ks)$, de maneira que propicie condições de desenvolvimento da noção intuitiva da densidade de observações em torno da média aritmética.

É possível notar que, para o exemplo das alturas dos cinqüenta alunos, foram obtidas as medidas (média e desvio padrão); foi observada a porcentagem de observações no intervalo composto pela média mais ou menos dois desvios padrão da média; foi estimada a porcentagem de observações no intervalo por meio do Teorema de Tchebichev e pela Distribuição Normal; foram elaborados os gráficos de histograma, de pontos e do intervalo da média mais ou menos dois desvios padrão.

Todas essas representações podem permitir ao aprendiz construir o campo conceitual de variação em torno da média e tornar seu raciocínio sobre este tema mais avançado.

## 2.4 Cálculo do Desvio Padrão a partir da Distribuição de Freqüências

Quase todos os livros didáticos de Estatística apresentam o cálculo do desvio padrão (e das outras medidas de tendência central e dispersão) a partir da distribuição de freqüências.

O cálculo do desvio padrão a partir da distribuição de freqüências pode ser realizado de duas maneiras. Na primeira estratégia, é contado o número de casos de cada observação diferente, que é denominada neste trabalho de distribuição de freqüências simples. Na segunda estratégia, a variável observada é agrupada em classes e então é contado o número de casos em cada classe, que é denominada neste trabalho de distribuição de freqüências com dados agrupados. Ambas estratégias estão exemplificadas a seguir.

A Tabela 8 apresenta a distribuição de freqüências simples das alturas dos 50 alunos e os cálculos parciais para a obtenção do desvio padrão.

Tabela 8 - Distribuição de freqüência (simples) dos alunos de acordo com a altura

| Altura (cm) ($x_i$) | Número de alunos - $f_i$ (freqüência) | $(x_i - \bar{x})^2$ | $(x_i - \bar{x})^2 \cdot f_i$ |
|---|---|---|---|
| 142 | 1 | 135,4896 | 135,4896 |
| 143 | 1 | 113,2096 | 113,2096 |
| 144 | 1 | 92,9296 | 92,9296 |
| 145 | 2 | 74,6496 | 149,2992 |
| 147 | 1 | 44,0896 | 44,0896 |
| 148 | 2 | 31,8096 | 63,6192 |
| 149 | 3 | 21,5296 | 64,5888 |
| 150 | 2 | 13,2496 | 26,4992 |
| 151 | 1 | 6,9696 | 6,9696 |
| 152 | 4 | 2,6896 | 10,7584 |
| 153 | 3 | 0,4096 | 1,2288 |
| 154 | 5 | 0,1296 | 0,6480 |
| 155 | 4 | 1,8496 | 7,3984 |
| 156 | 5 | 5,5696 | 27,8480 |
| 157 | 4 | 11,2896 | 45,1584 |
| 158 | 2 | 19,0096 | 38,0192 |
| 159 | 3 | 28,7296 | 86,1888 |
| 160 | 3 | 40,4496 | 121,3488 |
| 161 | 2 | 54,1696 | 108,3392 |
| 162 | 1 | 69,8896 | 69,8896 |
| $\bar{x} = 153{,}64$ | | | $\sum (x_i - \bar{x})^2 \cdot f_i$ = 1.213,52 |

Na primeira coluna da Tabela 8 foram ordenadas todas as diferentes alturas da amostra de 50 alunos ($x_i$). A segunda coluna apresenta o número de alunos com cada altura diferente (freqüência - $f_i$ ). A terceira coluna apresenta o cálculo do quadrado da distância entre cada altura e a média da amostra, resultado esse multiplicado pelo número de alunos com a altura referente, apresentado na quarta coluna.

O desvio padrão para a variável altura é $s = \sqrt{\dfrac{\sum (x_i - \bar{x})^2 \cdot f_i}{\sum f_i - 1}} = \sqrt{\dfrac{1213{,}52}{50 - 1}} \cong 4{,}98$ cm, o mesmo valor apresentado anteriormente, calculado com o rol de observações.

Quando calculamos o desvio padrão a partir da distribuição de freqüências simples, não há nenhuma perda de informação, ou seja, o valor do desvio padrão obtido com os dados brutos (conjunto numérico original) é o mesmo obtido a partir da distribuição de freqüências simples.

Devido ao fato de existirem muitos valores diferentes de altura (o que é comum em variáveis contínuas), utiliza-se a estratégia de agrupar as observações, que se denomina de distribuição de freqüência com dados agrupados em classes.

Utilizando essa estratégia, as observações são agrupadas de acordo com algum critério[18]. Para o exemplo das alturas dos 50 alunos, as classes (grupos) foram elaboradas em intervalos de quatro centímetros (coluna 1 da Tabela 9). Após a determinação da amplitude das classes, é contado o número de alunos ( $f_i$ ) que pertencem a cada classe (coluna 3 da Tabela 9).

Tabela 9 - Distribuição de freqüência (com dados agrupados em classes) dos alunos de acordo com a altura (em cm)

| Altura – agrupada em classes | Ponto médio da classe $x_i$ | Número de alunos $f_i$ | $(x_i - \overline{x})^2$ | $(x_i - \overline{x})^2 \cdot f_i$ |
|---|---|---|---|---|
| [140, 144[ | 142 | 2 | 135,4896 | 270,9792 |
| [144, 148[ | 146 | 4 | 58,3696 | 233,4784 |
| [148, 152[ | 150 | 8 | 13,2496 | 105,9968 |
| [152, 156[ | 154 | 16 | 0,1296 | 2,0736 |
| [156, 160[ | 158 | 14 | 19,0096 | 266,1344 |
| [160, 164[ | 162 | 6 | 69,8896 | 419,3376 |
| Somatório | | | | 1.298,00 |

Para efeito do cálculo do desvio padrão, todos os alunos contidos em uma classe têm suas alturas representadas pelo ponto médio da classe. No exemplo, os dois alunos agrupados na primeira classe, cujas alturas eram 142 e 143, foram considerados com altura 142, que é o ponto médio do intervalo [140, 144[ , os

[18] Este assunto é tratado em livros de Estatística, no capítulo de Distribuição de Freqüências.

quatro alunos agrupados na segunda classe (144, 145, 145, 147) foram considerados com altura 146 e assim sucessivamente. Essa estratégia ocasiona uma perda de informação, devido à substituição do valor original pelo ponto médio da classe.

Com os cálculos apresentados na Tabela 9 calcula-se o desvio padrão com os dados agrupados em classes: $s = \sqrt{\dfrac{\sum\left(x_i - \bar{x}\right)^2 \cdot f_i}{\sum f_i - 1}} = \sqrt{\dfrac{1298}{50-1}} = 5{,}1468 \cong 5{,}15\ \text{cm}$, que é diferente do valor do desvio padrão obtido anteriormente ( 4,98 cm).

A apresentação do cálculo do desvio padrão a partir de uma distribuição de freqüência simples e a partir de dados agrupados teve dois objetivos: a) discussão sobre a aproximação da medida de variação e b) sua relação direta com o tipo de variável quantitativa (discreta ou contínua).

A distribuição de freqüência simples é a estratégia utilizada para variável discreta e com poucos valores diferentes na amostra. Por exemplo, o número de namorados que os alunos de uma quinta série já teve. O valor do desvio padrão obtido a partir da distribuição de freqüência simples é o mesmo que quando obtido a partir do rol de observações.

A distribuição de freqüência com dados agrupados é a estratégia utilizada para variável contínua, como por exemplo, altura, peso, etc. Os cálculos tornam-se menos exaustivos, porém o valor obtido é aproximado.

Os livros de Estatística justificam a utilização da distribuição de freqüências com dados agrupados como uma maneira de simplificar o cálculo do desvio padrão. Essa vantagem deixa de ter sentido quando se utiliza a ferramenta *desvio padrão* da calculadora ou da planilha eletrônica, que fornece automaticamente o valor final dessa medida de variação.

Porém, para o objetivo desta pesquisa, vale ressaltar que a vantagem do conhecimento do cálculo do desvio padrão com dados agrupados é a possibilidade de inferir a magnitude do desvio padrão (a densidade das observações em torno da média) a partir da visualização da representação gráfica de uma distribuição de freqüências com dados agrupados (histograma). Nesse sentido, existe o estudo de Delmas e Liu (2005) que está apresentado no Capítulo 4 deste trabalho.

## 2.5 Coeficiente de Variação

O coeficiente de variação, também denominado de coeficiente de dispersão, é uma medida relativa da variação em torno da média. É obtido pela fórmula: $\boxed{CV = \frac{\sigma}{\mu}}$ para os parâmetros populacionais e $\boxed{CV = \frac{s}{\bar{x}}}$ para as informações de uma amostra. Se este coeficiente for multiplicado por 100, permite obter uma percentagem de variação da medida.

Para o exemplo das notas (0, 1, 2, 5, 7), cujos parâmetros eram $\mu = 3$ e $\sigma = 2{,}61$, tem-se $CV = \frac{\sigma}{\mu} = \frac{2{,}61}{3} = 0{,}87$. Multiplicando por 100, tem-se 87% de variação das alturas em torno da média aritmética. É importante observar que não se trata de 87% das observações, mas da grandeza da observação. Se o desvio padrão fosse 3 (igual a média aritmética), o CV seria 100%, ou seja, haveria uma variação igual à medida de centro e, portanto, poderia se dizer que existe uma grande variação das alturas na amostra.

Segundo Costa Neto (2002 p. 28), “a vantagem é caracterizar a dispersão dos dados em termos relativos a seu valor médio. Assim, uma pequena dispersão absoluta pode ser, na verdade, considerável quando comparada com a ordem de grandeza dos valores da variável e vice-versa”

Para compreender a citação de Costa Neto, é possível utilizar os dois exemplos tratados neste trabalho. O conjunto numérico das notas (0, 1, 2, 5, 7) tem $\mu = 3$ e $\sigma = 2{,}61$ e CV = 87%. O conjunto numérico das alturas tem $\bar{x} = 153{,}64$, s = 4,98 e CV = 3,24%. Ao se comparar a variação dos dois conjuntos numéricos em termos absolutos, é possível dizer que as alturas têm maior variação (4,98 > 2,61). Porém, a variação de 4,98 deve ser comparada com sua medida de centro, média, que é 153,64, e o mesmo deve ser feito para o outro conjunto numérico. Assim, o desvio padrão de 4,98 representa 3,24% de variação enquanto que o desvio padrão de 2,61 representa 87% de variação. Logo, o conjunto numérico das notas apresenta variação muito maior que o conjunto numérico das alturas.

Esse exemplo explicita a vantagem da utilização do Coeficiente de Variação. Devido ao fato de o CV ser a razão entre o desvio padrão e a média aritmética, ele é uma medida de dispersão adimensional e, assim, permite comparar dois (ou mais) conjuntos de observações cujas unidades de medida não são as mesmas. Tal como nos dois exemplos, está conceitualmente errado comparar o desvio padrão das notas (que está medida em pontos) com o desvio padrão das alturas (que está medido em centímetros). Porém, é possível comparar o CV dos dois conjuntos e concluir que o conjunto numérico das notas apresenta variação maior que o conjunto numérico das alturas.

## 2.6 Desvio Médio

Uma outra medida de dispersão em relação à média é o desvio médio. Seja uma variável aleatória X ($X_1$, $X_2$, ..., $X_N$), o desvio médio é a média aritmética dos desvios absolutos em relação à média:

$$\boxed{dm(X) = \frac{\sum_{i=1}^{N} |X_i - \mu|}{N}}$$

No cálculo da variância, os desvios em relação à média são elevados ao quadrado. No caso do desvio médio, os desvios em relação à média são colocados em módulo. Ambas estratégias matemáticas são necessárias, pois a soma dos desvios em relação à média é sempre zero.

Diferentemente da variância, em que o símbolo utilizado é universal, o símbolo utilizado para o desvio médio pode mudar conforme a bibliografia utilizada. Bussab e Morettin (2003) utilizam o símbolo ***dm,*** Guimarães e Cabral (1997) utilizam ***DAM***.

Voltando ao exemplo das cinco notas (0, 1, 2, 5 e 7) , em que a média aritmética é 3, exemplifica-se a seguir o cálculo do desvio médio:

$$dm(X) = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} = \frac{|0-3| + |1-3| + |2-3| + |5-3| + |7-3|}{5} = \frac{12}{5} = 2{,}4\,,$$ significando que, em média, as notas ficam distantes da média em 2,4 pontos.

É importante ressaltar que, mesmo sendo utilizada a média aritmética para calcular o desvio médio, o que minimiza essa somatória é a mediana. Como lembram Guimarães e Cabral (1997), a mediana minimiza a soma dos desvios, em módulo, da mesma maneira que a média aritmética minimiza a soma dos quadrados dos desvios.

O desvio médio é pouco explorado nos livros de Estatística devido ao fato de haver restrições matemáticas em relação aos valores em módulo[19], o que restringe sua aplicação em outros contextos estatísticos e justifica a vasta utilização da variância (e conseqüentemente do desvio padrão).

## 2.7 Amplitude total

A amplitude total, simbolizada neste trabalho por $A_t$,é a medida de variação mais simples e é muito eficaz, pois permite observar a variação geral de uma distribuição. A amplitude total é definida como a diferença entre o valor máximo e o valor mínimo observados, ou seja,

$$\boxed{A_t = X_{max} - X_{min}}$$

Para o exemplo das notas: 0, 1, 2, 5, 7, como a menor nota é zero e a maior nota é sete, a amplitude total será $A_t = 7 - 0 = 7$, ou seja, existe uma diferença de sete pontos entre a menor a maior nota.

Diferentemente das medidas anteriormente discutidas, a amplitude total é uma medida de variação que não tem um ponto de referência, uma medida de tendência central. A variância, o desvio padrão, o coeficiente de variação e o desvio médio são medidas de variação em torno da média enquanto a amplitude total não representa uma variação em torno de alguma medida, mas simplesmente uma variação global.

A amplitude total é, talvez, a única medida de variação intuitiva, pois é muito natural a observação dos valores máximo e mínimo em uma distribuição. As crianças, desde muito pequenas, já observam o amigo mais alto e o amigo mais baixo, a maior quantidade de balas distribuídas e a menor quantidade, etc.

Embora seja uma medida simples, é de extrema importância na leitura de histograma, *dotplot*, etc., pois permite identificar a variação dos valores da variável, algumas vezes esquecida devido à complexidade do gráfico, tal como já salientado por Meletiou e Lee (2002).

Neste capítulo foram apresentadas as seguintes medidas de variação (dispersão): variância, desvio padrão, coeficiente de variação, desvio médio e amplitude. Foram apresentados os aspectos teóricos que determinam sua

---

[19] Estas restrições matemáticas podem ser exemplificadas como a não possibilidade de integrar/derivar a função modular.

fórmula, a interpretação do resultado numérico e as diferentes maneiras de representação gráfica.

Esta discussão teve como objetivo apresentar o objeto matemático deste estudo: medidas de variação. O capítulo seguinte apresenta uma análise de como o livro didático de Matemática do ensino médio faz a apresentação matemática e didática desse objeto e o quarto capítulo apresenta os aspectos deste objeto priorizados nos estudos recentes publicados em jornais e eventos científicos importantes na área.

## 3 Análise do Complexo Praxeológico do Objeto Variação/Variabilidade

O livro didático é um recurso desenvolvido com o objetivo de ser utilizado pelo estudante, mas sabe-se que também é um recurso de pesquisa utilizado pelo professor quando está planejando suas aulas, como salientado no próprio PCN.

> Não tendo oportunidade e condições para aprimorar sua formação e não dispondo de outros recursos para desenvolver as práticas da sala de aula, os professores apóiam-se quase exclusivamente nos livros didáticos, que, muitas vezes, são de qualidade insatisfatória. (BRASIL, 1998, p. 22).

Partindo-se do pressuposto de que grande quantidade de professores apóia-se nos livros didáticos, este capítulo tem o objetivo de identificar as tarefas, as técnicas e os aportes teóricos, ou seja, o complexo praxeológico apresentado pelos livros didáticos de Matemática do ensino médio para o desenvolvimento do objeto variação/variabilidade.

O conhecimento do complexo praxeológico de variação/variabilidade pode permitir a compreensão do(s) tipo(s) de raciocínio do professor de matemática frente à solução de problemas estatísticos que envolvem variação. E, por esse motivo, salienta-se que não é objetivo fazer nenhum juízo de valor das obras analisadas.

A análise praxeológica seguiu os pressupostos da Teoria Antropológica do Didático (TAD), desenvolvida por Chevallard (1995, 1996, 1999).

Apoiado na Antropologia Cognitiva, Chevallard (1996) utiliza os termos objeto, instituição e sujeito. Para ser um objeto[20], é preciso que haja uma relação deste com uma pessoa (X) ou com uma instituição (I), ou seja, deve haver uma $R(X,O)$ ou uma $R_I(O)$, respectivamente. Vale ressaltar que essa relação é explicada pelo autor como o conhecimento do objeto pela instituição ou pela pessoa.

Para Chevallard (1996, p. 148), um saber (S) é uma categoria particular de objeto que "pode ser aprendido, e pode ser ensinado; mas, não pode ser conhecido sem ter sido aprendido. Por outro lado, podem ser utilizados e, para existirem, têm de ser produzidos".

---

[20] Para Chevallard (1996), um objeto tem um sentido mais amplo que puramente matemático. Pode ser uma pessoa, uma escola, um saber.

A instituição pode ser uma sala de aula, uma escola, uma disciplina, um livro didático, etc. e “a cada instituição I está associado um conjunto de objetos $O_I$, chamado conjunto dos objetos institucionais (para I), que é o conjunto dos objetos O que I conhece” (CHEVALLARD, 1996, p. 129). Essa relação é dinâmica e, por esse motivo, sofre interferência temporal. A cada instante surgem (ou desaparecem) objetos institucionais e, portanto, o autor prefere representar esta relação por $O_I(t)$.

As instituições didáticas, categoria particular das instituições, dão origem a que o autor denomina de Teoria Antropológica do Didático (TAD).

> A intenção didática manifesta-se através da formação de instituições a que chamo, genericamente, sistemas didáticos. Um sistema didático (SD) comporta um ou vários sujeitos de I, que nele ocupam uma *posição de professor P*, um ou vários sujeitos de I que nele ocupam uma *posição de aluno a*, e finalmente um objeto O, pertencentes a $P_I$(a), que é o conjunto dos investimentos didáticos para I. (CHEVALLARD, 1999, p. 133).

Para que um sistema didático funcione, é necessária a existência de um meio, ou seja, que, em cada instante, exista uma relação institucional $R_I$ (p,O) (em que p = a,P) localmente estável (CHEVALLARD, 1999). Vale salientar também que um sistema didático se alimenta de outros sistemas, tal como o de ensino.

Segundo Chevallard (1999, p. 223), o postulado de base da TAD é que “se admite em efeito que toda atividade humana regularmente realizada pode ser descrita com um modelo único, que se resume com a palavra praxeologia”, inclusive as atividades matemáticas.

Rossini (2006, p. 30) explica que o termo praxeologia é a união de *praxis* (prática) e *logos* (razão) e explica que o termo lembra “uma prática humana, no interior de uma instituição, que está sempre acompanhada de um discurso mais ou menos desenvolvido, ou seja, de um logos que a justifica”.

Para a análise de uma *praxis*, o saber-fazer, a teoria revela o estudo das tarefas propostas pelos livros bem como as técnicas apresentadas para a solução das tarefas e para a análise do *logos*, tem-se o discurso teórico-tecnológico que justifica a *praxis*.

Chevallard (1995 p.85) explica que tarefa é uma atividade bem circunscrita. Ele recupera a explicação dada à palavra *tarefa* pelo dicionário histórico da língua francesa, que "é um trabalho determinado que somos obrigados a fazer, com uma noção de retribuição pelo trabalho". Chevallard (1999) explica que uma tarefa se expressa por meio de um verbo tal como limpar a casa, desenvolver uma expressão, dividir um número inteiro por outro inteiro, subir uma escada, etc. Uma tarefa t pode fazer parte de um tipo de tarefas *T*. No caso de variação, por exemplo, calcular o desvio padrão é um tipo de tarefa e calcular o desvio padrão com dados agrupados é uma tarefa pertencente a este tipo de tarefas *T*.

Para um indivíduo *X*, um tipo de tarefa *T* pode ser rotineiro ou problemático. "Se a tarefa *T* é rotineira para *X*, isto quer dizer que X possui e domina uma maneira de fazer, denominada técnica" e simbolizada por $\tau$ (CHEVALLARD, 1995, p. 86).

Chevallard (1995) explica que uma técnica $\tau$ é problemática para os iniciantes, mas que passa a ser rotineira à medida que o indivíduo *X* adquire a maneira de fazer a tarefa. E exemplifica com a criança pequena *(X)* que começa a andar *(T).* A técnica é problemática, mas com o tempo, passa a ser rotineira. Mas, uma técnica não tem êxito em mais de que uma parte da tarefa, parte que se denomina alcance da técnica. Por esse motivo, uma técnica pode ser superior a outra, de maneira parcial ou total.

No caso de variação, uma das tarefas pode ser calcular o desvio padrão com um rol de observações. Para tanto, existe uma técnica que é a aplicação da fórmula $\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{N}}$. Este é um bloco (*T*/$\tau$) que é o "saber-fazer".

A tecnologia, simbolizada por $\theta$, é um discurso compreensivo sobre a técnica, cujo objetivo é justificá-la racionalmente, assegurando-se de que permite realizar as tarefas do tipo *T* (CHEVALLARD, 1999).

Chevallard (1999, p. 226) explica que a técnica relativa ao tipo de tarefas *T* está sempre acompanhada de um vestígio de tecnologia. E apresenta o exemplo da aritmética:

> "Um mesmo pequeno discurso tem dupla função, técnica e tecnológica, que permite num momento encontrar o resultado pedido (técnica) e justificar que é correto o resultado esperado (tecnologia), como quando se diz : se 8 balas custam 10 Francos, 24 balas, ou seja, 3 vezes 8 balas, custam 3 vezes mais, 3 vezes 10 Francos."

O objetivo de uma tecnologia não é apenas justificar a técnica, mas também explicá-la.

Utilizando o exemplo da tarefa de calcular o desvio padrão, as tecnologias empregadas na técnica são as propriedades de operação em $\Re$ (somatório, operação de potência, ordem das operações, raiz quadrada), média aritmética e o conceito de desvios da média.

Por fim, no bloco que Chevallard (1995) denomina de saber, encontra-se a teoria, que é a tecnologia da tecnologia, a justificativa da justificativa. O autor explica que teoria é um nível superior de justificativa-explicação-produção. No caso da tarefa de calcular o desvio padrão para um rol de observações, as teorias que estão envolvidas são: medidas de tendência central e dispersão (estatística) e álgebra.

Diante do exposto, formaliza-se que uma praxeologia, ou organização praxeológica, está constituída de um bloco prático-técnico ($T/\ \tau$), referente ao saber-fazer e por um bloco tecnológico-teórico ($\theta\ /\Theta$) referente ao saber. Essa praxeologia pode ser pontual, quando relaciona um único tipo de tarefa T ao complexo ($T/\ \tau\ /\theta\ /\Theta$) [tarefa, técnica, tecnologia e teoria] (CHEVALLARD, 1999, p. 228).

> Geralmente, em uma instituição I, uma teoria $\Theta$ responde por várias tecnologias $\theta_j$, cada uma das quais, por sua vez, justifica e explica varias técnicas $\tau_{ij}$, correspondentes a outros tantos tipos de tarefas $T_{ij}$. As organizações pontuais vão combinar-se, em primeiro lugar, em organizações locais $[T_i / \tau_i / \theta / \Theta]$ centradas sobre uma tecnologia $\theta$ determinada e depois em organizações regionais $[T_{ij} / \tau_{ij} / \theta_j / \Theta]$ formadas ao redor de uma teoria $\Theta$.

E Chevallard (1999) explica que uma organização global é um complexo praxeológico dado pelo conjunto de organizações regionais.

A organização praxeológica de um determinado objeto divide-se em organização praxeológica matemática, ou simplesmente organização matemática de um tema de estudo $\theta$, que é o estudo da própria realidade matemática e a organização praxeológica didática, ou simplesmente organização didática desse tema $\theta$, que é a maneira como pode ser construída essa realidade matemática.

A seguir estão apresentadas a forma de seleção dos livros analisados, a análise da praxeologia didática e da praxeologia matemática.

### 3.1 A Seleção dos Livros Analisados

Foram selecionados para a análise, todos os livros do ensino médio que eram utilizados por, pelo menos, um dos professores participantes da pesquisa, sejam como livros-texto ou como material de suporte para a elaboração de suas aulas. A exceção se faz apenas a um livro indicado pelas professoras OB e RN, ao qual não foi possível ter acesso.

Todos os livros analisados se referiam ao exemplar do professor, sendo que alguns acompanhavam o Manual do Professor (ou Guia Pedagógico), também analisados com o objetivo de verificar se as técnicas e os discursos teórico-tecnológico se ampliavam.

A lista dos livros analisados está apresentada na Tabela 10. É importante destacar que os livros L1, L5 e L6 estão apresentados em volume único para o ensino médio e referem-se exatamente ao tipo de livro que os professores utilizavam. Sabe-se, porém, que o conteúdo dos livros de ensino médio que apresentam-se em volume único é mais abreviado que o dos livros que se apresentam numa coleção.

No que diz respeito à forma de utilização dos livros, é possível observar na Tabela 10 que apenas uma professora (LF) adotava livro em suas aulas. Os outros professores utilizavam o livro didático como recurso para elaborar as atividades a serem realizadas com seus alunos. Vale destacar o caso do professor AM e da professora CI que seguiam a apostila adotada pela escola privada em que lecionam.

Tabela 10 - Livros analisados neste trabalho

| Nomeado nesta análise de: | Referência do livro | Número de professores que: | |
|---|---|---|---|
| | | Adotam o livro | Utilizam o livro para elaborar as aulas |
| L1* | Giovanni, José Ruy, Bonjorno, José Roberto e Giovanni Junior, José Ruy. *Matemática Fundamental: uma nova abordagem: ensino médio: volume único.* São Paulo: FTD, 2002. | 0 | 2 |
| L2 | Paiva, Manoel Rodrigues. *Matemática. São Paulo: Moderna, 1995. vol 2.* | 0 | 3 |
| L3* | Smole, Kátia Cristina Stocco e Diniz, Maria Ignez de Souza Vieira. *Matemática: ensino médio. São Paulo: Saraiva, 2003. vol. 1,2,3* | 0 | 2 |
| L4 | Iezzi, Gelson, Dolce, Osvaldo, Degenszajn, David, Perigo, Roberto e Almeida, Nilze de. *Matemática: Ciência e Aplicações, vol. 3.* São Paulo: Atual, 2001 | 0 | 1 |
| L5* | Dante, Luiz Roberto. *Matemática: Contexto e Aplicações – volume único*. São Paulo: Ática, 2000 | 0 | 2 |
| L6 | Yossef, Antonio Nicolau, Soares, Elizabeth e Fernandez, Vicente Paz. *Matemática: volume único para o ensino médio*. São Paulo: Scipione, 2004. (Coleção De olho no mundo do trabalho) | 1 | 0 |
| L7* | Gentil, Nelson, Santos, Carlos Alberto Marcondes dos, Greco, Antonio Carlos, Belotto Filho, Antonio e Greco, Sergio Emilio. *Matemática para o 2º grau. Volume 2*. São Paulo: Ática, 1996. | 0 | 1 |

*O livro contém o Manual do professor (ou Guia Pedagógico)

Antes de iniciar a análise praxeológica dos livros, observou-se muita diferença na quantidade de páginas dedicadas à Estatística e, especificamente à variação/variabilidade, que pode ser observado na Tabela 11.

Devido ao fato da não possibilidade de acesso a todos os exemplares das coleções, não foi possível verificar a proporção do livro (ou da coleção) reservada ao assunto, restringindo-se a comparação apenas à observação de que as medidas de variação foram apresentadas em quatro ou cinco páginas nos livros L1, L5, L6 e L7, enquanto o mesmo conteúdo foi tratado com, no mínimo, o dobro de páginas nos livros L2, L3 e L4 , que certamente propiciou ao(s) autor(es) a oportunidade de explorar mais este conteúdo.

Tabela 11 - Número de páginas dedicadas à Estatística e à variação/variabilidade

| Livro | Número de páginas de Estatística | Número de páginas de Variabilidade |
|---|---|---|
| L1 | 32 | 4 |
| L2 | 34 | 11 |
| L3 | 86 | 14 |
| L4 | 66 | 18 |
| L5 | 25 | 5 |
| L6 | 25 | 5 |
| L7 | 24 | 4 |

### 3.2 Análise da Organização Didática do Objeto Variação/Variabilidade

Segundo Chevallard (1999, p. 244), “as praxeologias didáticas ou organizações didáticas são respostas a questões do tipo como estudar a questão $q = \tau_T$?” e ainda (p. 246) “é o conjunto dos tipos de tarefas, técnicas, de tecnologias, etc., mobilizadas para o estudo concreto em uma instituição concreta”.

Entende-se que organização didática é um conjunto de opções dos autores dos livros para trabalhar um determinado conteúdo, que neste estudo é variação/variabilidade. Como explica Chevallard (1999, p. 246) “o enfoque praxeológico contempla aspectos da organização do estudo, geralmente vistos como relevantes de escolhas pedagógicas, políticas, exteriores ao campo de questionamentos da didática da Matemática”

Para Chevallard (1999), um momento didático é uma situação em que o desenvolvimento do estudo deverá ser cumprido em seis momentos didáticos: 1) o encontro com a tarefa; 2) a exploração do tipo de tarefa e das técnicas para resolver tal tarefa; 3) constituição do bloco tecnológico-teórico; 4) pôr à prova a técnica, ou seja, explorar o alcance da técnica; 5) institucionalização, ou seja, precisar o que é exatamente a organização matemática elaborada, distinguindo claramente, por um lado, os elementos que não foram integrados e, por outro lado, os elementos que entraram de maneira definitiva na organização matemática considerada e 6) a avaliação, reflexão do que se aprendeu.

A criação de situações didáticas adequadas é uma resposta às questões sobre a realização dos diferentes momentos didáticos: "como realizar concretamente o primeiro encontro com tal organização matemática? Com qual tipo de tarefas? Como conduzir o estudo exploratório de um tipo de tarefas dado? Como conduzir a institucionalização? Como realizar a avaliação?" (CHEVALLARD, 1999, p. 255).

Com base nessas questões sobre momentos didáticos, foi realizada a análise da organização didática do objeto variação/variabilidade e a Figura 8 apresenta uma síntese dessa análise, com o objetivo de facilitar a leitura.

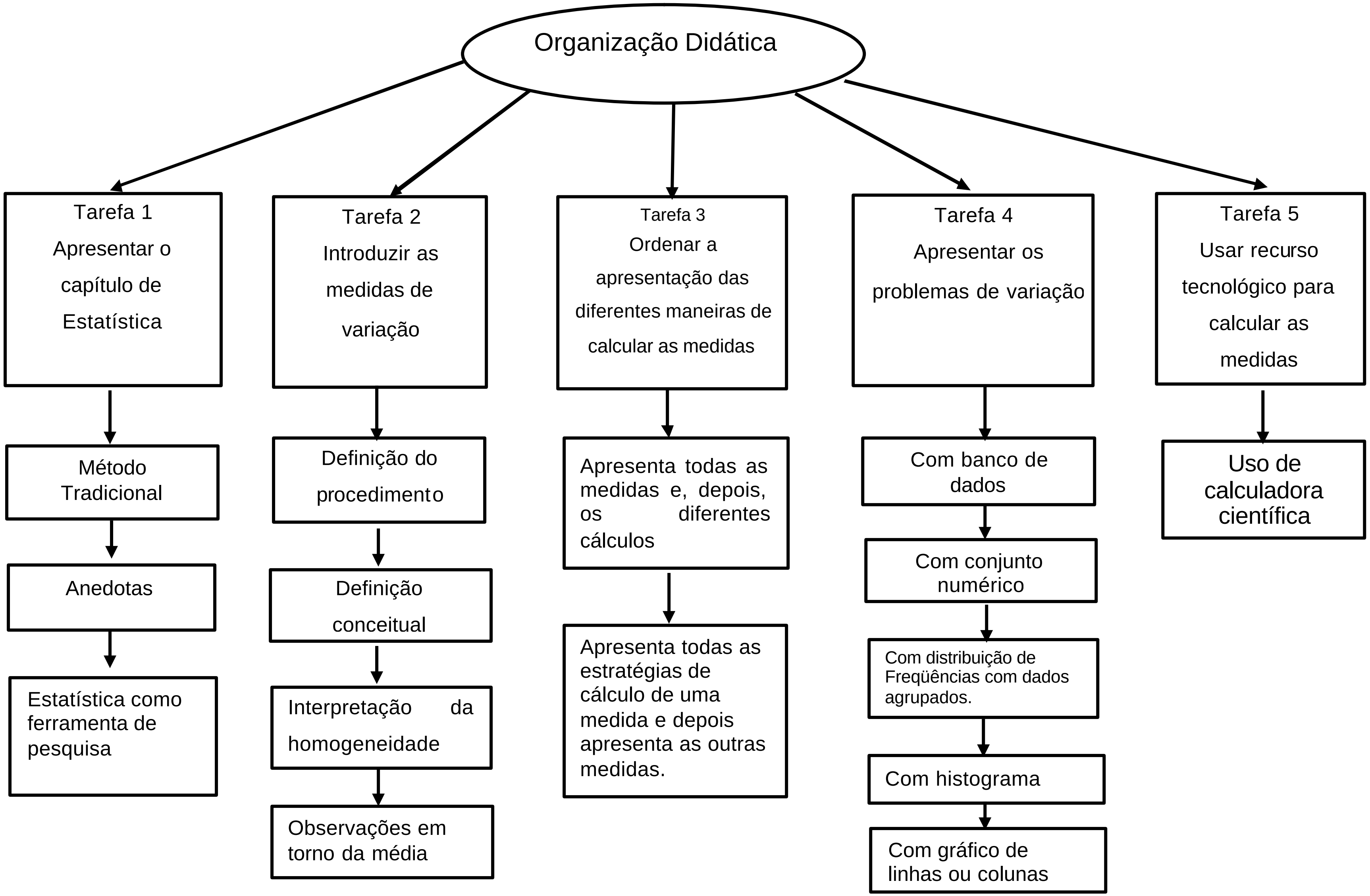



Figura 8 - Esquema da análise da organização didática dos livros de Matemática

Como pode ser observado na Figura 8, foram identificadas cinco tarefas didáticas no capítulo de Estatística que têm relação direta com o objeto variação/variabilidade e os retângulos que se encontram abaixo de cada tarefa sintetizam as diferentes técnicas observadas para cada tarefa.

Antes de iniciar a apresentação da análise didática, é importante identificar as medidas de variação constantes nos livros didáticos. O Quadro 3 relaciona essas medidas e aponta o livro didático que a apresenta. Além disso, também pode ser observado no quadro, a simbologia adotada pelo livro.

Quadro 3 – Identificação das medidas de variação e respectivo símbolo apresentados nos livros didáticos de Matemática

| Medida de variação | Livros didáticos | | | | | | |
|---|---|---|---|---|---|---|---|
| | L1 | L2 | L3 | L4 | L5 | L6 | L7 |
| amplitude | --- | Não usa símbolo | Não usa símbolo | --- | --- | R | Não usa símbolo |
| desvio relativo | --- | $D_r$ | --- | --- | --- | --- | --- |
| desvio absoluto | --- | $D_a$ | --- | --- | --- | --- | --- |
| desvio médio | $d_m$ | $D_{ma}$ | --- | DM | --- | --- | $d_m$ |
| variância | $V_a$ | $\sigma^2$ | V | $\sigma^2$ ou Var(X) | V | $s^2$ | --- |
| desvio padrão | s | $\sigma$ | S | $\sigma$ | DP | d | $d_p$ |

O Quadro 3 revela alguns fatos que merecem discussão. Primeiro, todos os livros apresentam uma medida de variação em torno da média e não comentam que existem outras medidas de variação que medem a dispersão em torno de outra medida de tendência central, tal como o intervalo interquartílico. Isso pode levar o leitor a pensar que só existem medidas de variação em torno da média.

Um outro fato relevante é que a amplitude, uma medida útil e intuitiva, só é apresentada por quatro livros, sendo que um deles (L2) não a apresenta no subcapítulo Medidas de Variação, mas no início do capítulo de Estatística.

O livro L2 utiliza duas medidas antecessoras ao desvio médio. Uma delas, o desvio relativo - $D_r$ (xi) - (distância de cada valor em relação a média) é trabalhada no livro para verificar se um valor qualquer do conjunto numérico está acima, abaixo ou é igual à média. A outra medida, o desvio absoluto – $D_a$(xi) -

(módulo da distância de cada valor em relação à média), é calculada mas o autor não explica a necessidade de se usar o módulo, ou valor absoluto.

A estratégia utilizada pelo autor do livro L2 é muito interessante, pois familiariza o leitor com a interpretação de uma medida de variação.

O desvio médio é apresentado por alguns livros enquanto a variância e o desvio padrão são adotados por todos, atribuindo a importância que essas medidas têm na Estatística. Embora o livro L7 não apresente um símbolo para a variância e nem formalize sua existência, ela aparece como o cálculo parcial para a obtenção do desvio padrão.

O fato que mais chamou a atenção no Quadro 3 é a diversidade de símbolos adotados pelos livros. Só para o desvio padrão, foram encontrados cinco símbolos (em sete livros): S, s, $\sigma$, DP, $d_p$ e d. Supondo que o professor de Matemática recorra a mais de um livro para elaborar sua aula e preparar uma seqüência didática para seus alunos, é possível que ele se confunda e desista dessa tarefa. Ehrenberg (1976) já alertava para o fato da não padronização da simbologia em Estatística, o que dificultava a solução de exercícios por parte de alunos de graduação.

Vale ressaltar que todos os livros analisados não fazem nenhuma menção ao parâmetro e estimador, o que já era esperado, dado o nível de escolaridade. E todos apresentam a fórmula da variância e do desvio padrão do estimador viesado do parâmetro.

**Tarefa 1**: Apresentar o capítulo de Estatística

Como apresentado por Chevallard (1999), a maneira como será realizado concretamente o primeiro encontro com a organização matemática é um aspecto didático que merece apreciação e que está colocado aqui como Tarefa 1. Foram encontradas três diferentes maneiras com as quais o capítulo Estatística se inicia, que foram denominadas de técnica.

**Técnica 1.1**: Seguindo a seqüência: definição conceitual, exercício resolvido, exercícios a serem resolvidos pelos alunos.

O autor escolhe apresentar primeiro as definições e alguns exercícios resolvidos para explicar as técnicas. Após essa apresentação inicial, são propostos exercícios que repetem as técnicas já desenvolvidas e, eventualmente, alguns exercícios com técnicas não apresentadas.

**Técnica 1.2**: Sensibilização com anedotas sobre Estatística e retorno à Técnica 1.1.

Nesta opção, o autor tenta motivar o aluno a estudar o assunto em pauta, apresentando possíveis enganos cometidos na interpretação de resultados estatísticos. É comum a anedota: um homem de um metro e oitenta se afogou num lago que tinha profundidade média de um metro e meio, apresentada em Giovanni, Bonjorno e Giovanni Jr (2002 ).

**Técnica 1.3**: Sensibilização sobre o papel da Estatística como ferramenta de pesquisa e retorno à Técnica 1.1.

Esta técnica permite mostrar ao leitor do livro a utilidade da Estatística e a proximidade com o mundo real. Smole e Diniz (2003a) apresentam o perfil de compradores de imóveis numa cidade brasileira, que tinha sido veiculado num jornal de grande circulação e faz uma interpretação dos gráficos. Utilizando esta introdução, o leitor tem condições de ver a aplicabilidade do conteúdo que vai estudar e motivar-se para aprender. Após a introdução, o desenvolvimento do conteúdo se faz de acordo com a Técnica 1.1.

**Discurso Teórico-Tecnológico das Técnicas 1.1, 1.2 e 1.3 -** Esta apresentação do conteúdo é uma opção pelo método de ensino tradicional, em que o aluno espera a explicação e repete o tipo de técnica apresentada para resolver cada tipo de tarefa. É importante observar que esta abordagem é oposta à Teoria das Situações, em que se propõe ao aluno uma seqüência de exercícios (situação) para que ele mesmo encontre a técnica para resolver o exercício e tenha condições de generalizar o discurso teórico-tecnológico.

O discurso utilizado pelos livros também está em discordância com as orientações dos Parâmetros Curriculares Nacionais, que orientam o trabalho com a resolução de problemas, que permite o desenvolvimento do raciocínio, neste caso, o raciocínio estatístico.

Foi observado como cada livro iniciava o capítulo Estatística e o Quadro 4 apresenta as técnicas utilizadas em cada livro.

Quadro 4 - Distribuição das técnicas identificadas na Tarefa Didática 1

| Técnica | L1 | L2 | L3 | L4 | L5 | L6 | L7 |
|---|---|---|---|---|---|---|---|
| Técnica 1.1 | X | X | X | X | X | X | X |
| Técnica 1.2 | X | | X | | | | |
| Técnica 1.3 | X | X | X | X | X | X | X |

Como pode ser observado no Quadro 4, todos os livros apresentam o capítulo de Estatística enaltecendo seu papel de ferramenta de pesquisa, embora alguns apenas restrinjam-se a escrever sobre isso enquanto outros apresentam resultados de pesquisa, o que torna mais concreta para o estudante[21] a utilização do conteúdo.

Todos os livros apresentam o conteúdo estatístico num discurso tradicional.

**Tarefa 2**: Introduzir as medidas de variação

Como já salientado no Capítulo 2 deste trabalho, o cálculo de algumas medidas de variação é extenso e, algumas vezes, complexo para a idade escolar. Por esse motivo, a tarefa de introduzir tais medidas é importante, para que possa motivar o leitor a buscar o conteúdo em questão.

**Técnica 2.1**: Diretamente com a definição de procedimento das medidas de variação.

Usando esta técnica, o autor do livro didático começa o capítulo de medidas de variação com a apresentação da fórmula, sem que sejam explicadas para que servem ou como são utilizadas em uma análise de dados. Esse tipo de introdução pode levar o leitor a pensar que a aplicação da fórmula é suficiente

[21] O termo estudante está sendo utilizado neste capítulo com a conotação de quem estará se utilizando da obra em referência para aprender, seja o professor ou o aluno.

para resolver o problema de variação e, dessa maneira, não o estimula a compreender o conceito, condição necessária para o desenvolvimento do raciocínio de variação e do pensamento estatístico.

**Técnica 2.2**: Diretamente com a definição conceitual das medidas de variação.

O autor começa o capítulo de medidas de variação com a apresentação da definição conceitual. Por exemplo, se estiver apresentando o desvio padrão, o autor explica que é a raiz quadrada positiva da média aritmética dos quadrados dos desvios em relação à média dos próprios valores.

**Discurso Teórico-Tecnológico***:* As técnicas 2.1 e 2.2 apresentam o mesmo discurso tradicional de ensino. Iniciam o capítulo com a definição, seja do procedimento ou do conceito, apresentam exercícios resolvidos e solicitam ao leitor que realize exercícios semelhantes à técnica apresentada.

Mesmo com a apresentação da definição conceitual, o livro deixa a cargo do estudante a tarefa de compreender o significado, a aplicação e a conexão com outros conteúdos estatísticos. Como lembra Gal (2002), as medidas de variação são conteúdos necessários para o letramento estatístico e não é possível garantir que aprender fatos, regras e procedimentos possa gerar um nível adequado desse letramento.

**Técnica 2.3 e Discurso Teórico-Tecnológico**: Com a interpretação de homogeneidade

No momento em que o autor explica que o resultado obtido do cálculo das medidas de variação é um indicativo da homogeneidade (ou não) da amostra, ele apresenta uma interpretação para essas medidas, a partir da qual, o leitor pode ser levado a olhar o todo, olhar o conjunto de observações e começar a fazer uma análise, antes mesmo de calcular.

Essa técnica pode evitar resultados semelhantes ao obtido por Ben-Zvi (2002), em que seus alunos só conseguiam observar pontos isolados ao invés de olhar toda a distribuição como o qual o estudante pode perceber que quanto maior for a variância (ou o desvio padrão ou o desvio médio), mais heterogênea é a amostra.

Um exemplo dessa técnica é retirado de Iezzi et al. (2001, p. 54), que apresentam as notas de quatro turmas cujas médias são iguais e explicam que só a média é insuficiente para a análise, pois "esse valor esconde informações em relação à homogeneidade ou heterogeneidade do desempenho dos alunos de uma mesma turma".

**Técnica 2.4 e Discurso Teórico-Tecnológico*:*** Com a idéia de afastamento ou aproximação dos dados em torno da média

Iniciar o capítulo com essa interpretação das medidas de variação é muito importante, pois o leitor pode perceber que as medidas de variação têm um ponto de referência, a medida de tendência central, e pode evitar que pense na diferença inerente às observações.

Essa interpretação é mais consistente que da homogeneidade, pois naquele caso, o leitor não tinha condições de observar que a medida de variação (tal como o desvio médio, desvio padrão e a variância) está apresentando a homogeneidade em relação a uma medida de tendência central.

Um exemplo sobre esta técnica foi retirado de Smole e Diniz (2003b), que apresentam a idade dos alunos de três turmas com mesma média, mas com distribuições diferentes e explicam que, por esse motivo, é necessário um indicador para informar a maneira como os dados se distribuem em volta da média.

As quatro técnicas identificadas para introduzir as medidas de variação foram encontradas nos livros analisados e estão identificadas no Quadro 5.

Quadro 5 - Distribuição das técnicas identificadas na Tarefa Didática 2

| Técnica | L1 | L2 | L3 | L4 | L5 | L6 | L7 |
|---|---|---|---|---|---|---|---|
| Técnica 2.1 | | | | | | | X |
| Técnica 2.2 | X | | | | | | |
| Técnica 2.3 | | | | X | X | X | |
| Técnica 2.4 | | X | X | X | X | X | |

É interessante observar que a maior parte dos livros está apresentando inicialmente uma interpretação das medidas de variação, porém em intensidades diferentes. Os livros L3, L4, L5 e L6 apresentam, pelo menos, um exemplo no início do capítulo de medidas de variação, em que a média de dois ou mais grupos são iguais e as observações têm distribuições muito diferentes. Alguns ainda associam o fato à homogeneidade (ou não) da distribuição, mas nenhum livro se refere à densidade dos dados em torno da média.

**Tarefa 3**: Ordenar a apresentação das diferentes maneiras de calcular as medidas de variação

**Técnica 3.1 e Discurso *Teórico-Tecnológico:*** O autor apresenta, primeiramente, todas as medidas de variação com o conjunto numérico e depois as apresenta novamente com dados agrupados.

Esta técnica pode contribuir para que o estudante perceba a relação existente entre as medidas de variação e possa entender a análise pertinente em cada caso, sem que se preocupe com o tipo de procedimento técnico a ser adotado. O estudante pode entender o que significa e pode generalizar o uso da técnica 3.1 para o uso de outra técnica para resolver uma tarefa do tipo: calcule o desvio padrão a partir do histograma. Ou seja, mesmo que o livro adote um método tradicional de apresentar o conteúdo estatístico (Tarefa 1), a maneira com que o autor vai apresentar as medidas de variação pode colaborar para que o indivíduo trabalhe numa linha construtivista da aprendizagem.

**Técnica 3.2 e Discurso *Teórico-Tecnológico*:** Primeiramente, apresenta todas as maneiras de se calcular uma medida de variação (conjunto numérico, distribuição de freqüência simples e agrupada em classes) e, em seguida, apresenta as demais medidas, na mesma ordem.

Esta técnica pode levar o estudante a se concentrar para aprender a calcular, sem pensar no significado daquela medida. O aprendiz pode interpretar que é um jogo de calcular, brincando de verificar se o resultado "bateu". Ou seja, há o favorecimento pelo desenvolvimento do procedimento em detrimento do desenvolvimento do pensamento estatístico e do raciocínio sobre variação.

O Quadro 6 mostra as técnicas utilizadas pelos livros analisados para apresentar as diferentes maneiras de se calcular as medidas.

Quadro 6 - Distribuição das técnicas identificadas na Tarefa Didática 3

| Técnica | L1 | L2 | L3 | L4 | L5 | L6 | L7 |
|---|---|---|---|---|---|---|---|
| Técnica 3.1 | | X | | X | X | | |
| Técnica 3.2 | X | | | | | X | X |

É possível observar no Quadro 6 que há uma divisão sobre a estratégia utilizada pelos livros. Quanto ao livro L3, não foi possível realizar essa análise, pois os autores não priorizaram o cálculo, apresentando um exemplo em que há muitos valores repetidos na amostra e utilizaram a estratégia de distribuição de freqüências simples, sem nomeá-la nem distingui-la de outras estratégias.

**Tarefa 4: Apresentar os problemas de variação**

Tanto os livros didáticos quanto os PCNs reforçam o papel da Estatística como ferramenta de pesquisa, mas a forma como os problemas de variação são apresentados podem não permitir a conexão das medidas de variação com as técnicas de análise de dados de uma pesquisa. Foram encontradas cinco diferentes técnicas de apresentação dos problemas de variação.

**Técnica 4.1 e Discurso Teórico-Tecnológico:** Utilizando um banco de dados

O autor utiliza um banco de dados, com variáveis qualitativas e quantitativas, que permite ao leitor rever o conteúdo estatístico já aprendido e organizar os dados para desenvolver os exercícios sobre variação. É uma forma de apresentação que prepara o aprendiz para trabalhar com dados reais, pois simula uma análise de pesquisa e permite o desenvolvimento do pensamento estatístico. Permite a aprendizagem em aspiral e não linear, ou seja, à medida que o aprendiz assimila o novo conteúdo, pode fazer relações com conteúdos vistos anteriormente, dando significado para ambos.

**Técnica 4.2 e Discurso Teórico-Tecnológico:** Com um conjunto numérico

O autor apresenta o problema com um conjunto numérico de alguma variável, ordenado ou não. Ou seja, o autor utiliza apenas uma variável quantitativa e adota alguns valores para essa variável. Esta técnica é interessante para a introdução do conceito, pois permite compreender o significado sem que o cálculo seja o protagonista do conceito. Porém, aliar esta técnica de apresentação do problema com outras (como a Técnica 4.1) pode permitir ao aprendiz compreender o significado das medidas de variação numa tarefa de análise de dados de uma pesquisa.

**Técnica 4.3 e Discurso Teórico-Tecnológico:** Com uma distribuição de freqüências com dados agrupados

Este tipo de apresentação exige que o leitor utilize a fórmula para dados agrupados, que requer menos cálculos, porém com maior dificuldade de interpretação, pois precisa lembrar que existem *n* observações em cada classe. É uma técnica que requer a mobilização de conceitos matemáticos, mas pode não contribuir para a aquisição do conceito de variabilidade, pois pode estimular apenas o aspecto algorítmico do estudo de variabilidade.

**Técnica 4.4 e Discurso Teórico-Tecnológico:** Com histogramas

O leitor precisa identificar que o histograma é uma representação de uma distribuição de freqüências com dados agrupados, em que cada coluna refere-se a uma classe cujos limites estão identificados no eixo das abscissas e cuja freqüência está identificada no eixo das ordenadas. Após essa identificação, ele vai utilizar a técnica 4.3. Esta técnica é muito interessante se for aliada à visualização da medida de variação e de tendência central no gráfico, pois pode permitir ao aluno relacionar a medida de dispersão com o formato da distribuição, como já trabalhado por Bakker (2004), Meletiou (2000) e principalmente por Delmas e Liu (2005).

**Técnica 4.5 e Discurso Teórico-Tecnológico:** Com gráfico de linhas ou colunas (em que cada coluna representa uma única observação)

Para resolver esse problema, o estudante precisa visualizar que cada coluna (ou cada ponto no gráfico de linha) é uma observação de uma variável. Da mesma maneira que a Técnica 4.4, esta forma de apresentação do problema é muito interessante se, após a obtenção da medida de variação, houver a relação da medida com o formato do gráfico. O mesmo ocorre com o gráfico de linha, em que o estudante deve observar que cada ponto refere-se a uma observação e que, se as medidas de tendência central e dispersão forem agregadas ao gráfico, pode haver a conexão entre a representação da distribuição e sua respectiva medida de variação.

O Quadro 7 sintetiza as técnicas encontradas nos livros didáticos para apresentar os problemas de variação.

Quadro 7 - Distribuição das técnicas identificadas na Tarefa Didática 4

| Técnica | L1 | L2 | L3 | L4 | L5 | L6 | L7 |
|---|---|---|---|---|---|---|---|
| Técnica 4.1 | | | X | X | | | |
| Técnica 4.2 | X | X | X | X | X | X | |
| Técnica 4.3 | X | X | X | X | X | X | X |
| Técnica 4.4 | | X | | X | X | | |
| Técnica 4.5 | | | | X | | | |

**Tarefa 5:** Usar recurso tecnológico para calcular as medidas de variação

À medida que um aluno aprende calcular uma medida de variação, por exemplo, o desvio padrão, uma variável didática que geralmente é introduzida para analisá-la é o aumento do número de observações. Esse fato faz com que os cálculos tornem-se exaustivos e, assim, o uso de recurso tecnológico é uma alternativa didática interessante e defendida por inúmeros autores.

Certamente o uso do recurso deve ser feito com cuidado, para que a aprendizagem do manejo da ferramenta não se torne mais importante do que a aprendizagem do conceito.

**Técnica 5.1 e Discurso *Teórico-Tecnológico*:** Uso das funções estatísticas da calculadora científica

A utilização das funções estatísticas da calculadora substitui a aplicação da fórmula. Se o aprendiz quer obter a média e o desvio padrão de um conjunto numérico, ele deve entrar com todos os dados e solicitar a função $\bar{x}$ e s, obtendo diretamente o resultado final.

Somente o livro L3 faz esse tipo de apresentação ao término do capítulo de Estatística, explicando a utilização das funções estatísticas e resolvendo um exercício. Ou seja, o uso das funções estatísticas da calculadora não é estratégia de trabalho, mas é apresentado como uma das possibilidades para se obter as medidas.

Estimular o uso de uma ferramenta auxiliar é importante, pois diminui a ênfase nos cálculos e permite ao aluno focar sua atenção na interpretação dos resultados obtidos. Porém, não está sendo defendida aqui a retirada do ensino dos cálculos, mas, uma vez que o aluno aprendeu a calcular, ele pode usar uma ferramenta tecnológica para realizar esses cálculos. Existe uma vasta literatura sobre o papel da tecnologia no ensino de Estatística e as mais recentes alertam para o cuidado na utilização da tecnologia.

Os estudos mais recentes têm investigado o efeito do uso de *softwares* educacionais como o *Fathom* e o *Tinkerplots* na aprendizagem dos conceitos, que têm objetivos distintos do uso de uma ferramenta tecnológica, mas que também apresentam os resultados das medidas diretamente.

A análise da organização didática do capítulo Estatística permitiu observar que, nos livros verificados, há uma forte relação entre a Estatística e a pesquisa, que não é reforçada pelo uso de banco de dados na solução de problemas.

Especificamente sobre o desenvolvimento do conteúdo sobre variação, os livros adotam o método tradicional de ensino, contrariando as teorias mais recentes em Educação Matemática e a própria sugestão dos PCNs, mas é importante ressaltar que a maior parte dos livros analisados apresentam uma estratégia de interpretação das medidas de variação logo no início do capítulo, que pode permitir ao estudante dar início ao processo de desenvolvimento do raciocínio sobre variação/variabilidade.

### 3.3 Análise da Organização Matemática do Objeto Variação/Variabilidade

Segundo Chevallard (1999), uma organização praxeológica matemática, ou simplesmente organização matemática de um tema de estudo $\theta$ é o estudo da própria realidade matemática.

Neste trabalho, o tema de estudo se refere às medidas de variação e ao conceito de variabilidade e a organização matemática desse objeto foi realizada buscando responder às seguintes perguntas: a) durante a apresentação do capítulo Estatística há o interesse dos autores de desenvolver idéias intuitivas de variação, ou seja, desenvolver o conceito de variabilidade? b) em que situações se propõe a análise da medida de variação? c) como é desenvolvida a tarefa de calcular as medidas de variação? d) que tipo de análise da medida de variação é proporcionada? e por fim, e) há uma tendência para aspectos mais formais das medidas, como o trabalho com as propriedades?

A Figura 9 sintetiza as tarefas identificadas na organização matemática do objeto variação/variabilidade e os retângulos abaixo de cada tarefa referem-se às técnicas apresentadas para a solução de cada tarefa.

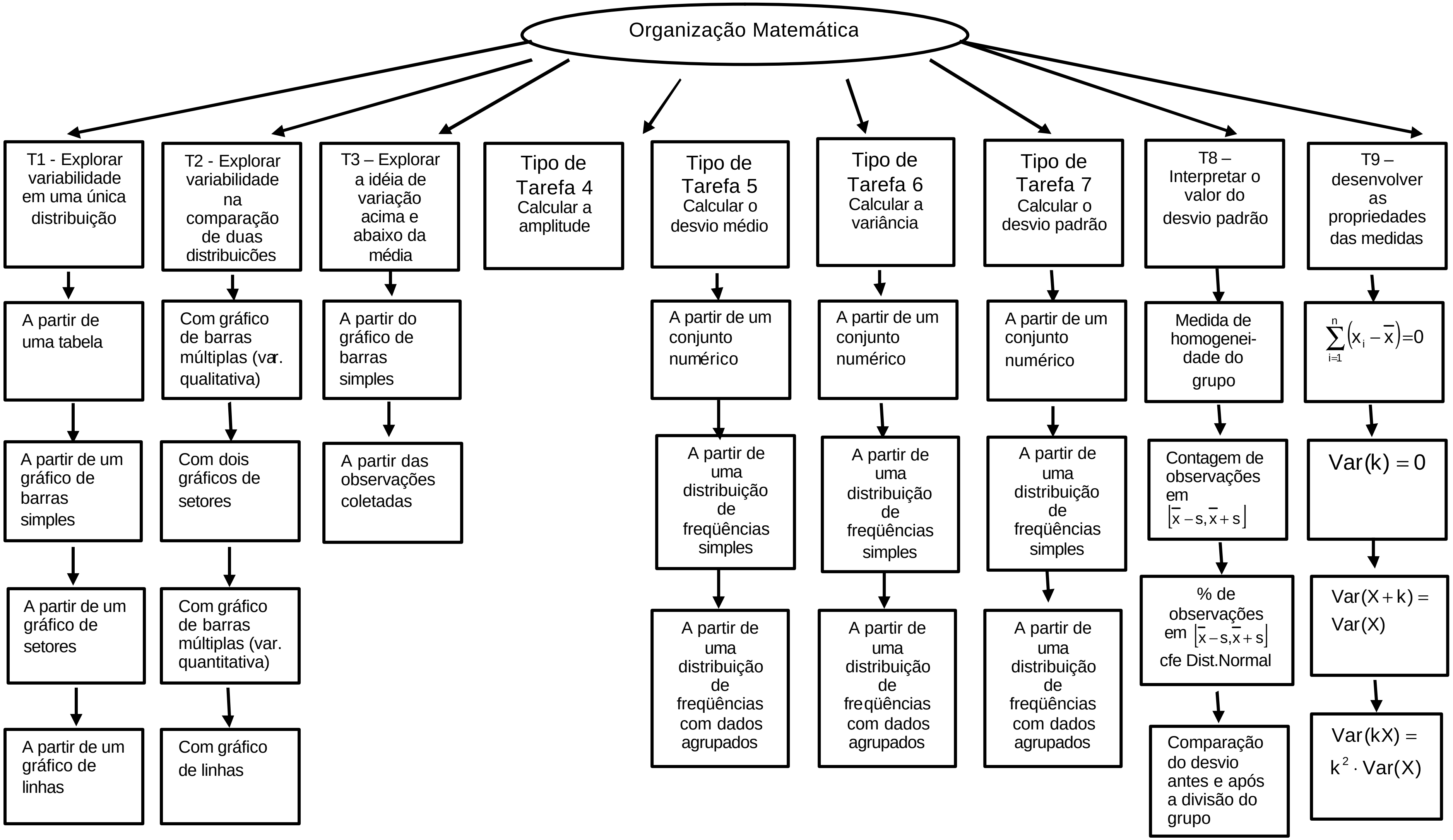



Figura 9 - Esquema da análise da organização matemática dos livros de Matemática

Como pode ser observado na Figura 9, foram encontradas nove tarefas matemáticas e cada retângulo imediatamente abaixo da tarefa representa a técnica empregada pelo livro para solucioná-la. A análise da organização matemática do objeto variação/variabilidade está descrita a seguir.

**Tarefa 1**: Explorar a variabilidade em uma única distribuição

A onipresença da variabilidade é um dos componentes do pensamento estatístico. Para tanto, observar e reconhecer a existência da variação é um exercício que pode ter início em atividades simples de Estatística, como a representação gráfica.

Em uma distribuição, a identificação da categoria que apresenta maior freqüência, ou seja, a categoria modal[22], permite observar o que é mais comum na distribuição, sem negligenciar a existência de observações diferentes, um passo importante para o desenvolvimento do raciocínio sobre variação.

Ben-Zvi (2004) observou que existe uma tendência natural de generalizar sem considerar a variação. Ele solicitou que os alunos observassem uma distribuição de freqüências (apresentada na Figura 15 deste capítulo) e apresentassem uma análise do tamanho dos nomes de alunos israelenses e americanos. Inicialmente os alunos diziam que os americanos têm nomes grandes e os israelenses nomes pequenos. Ou seja, o raciocínio deles era: <u>todos tem nomes pequenos</u>, e somente depois da intervenção do professor eles passaram a raciocinar: <u>a maioria dos israelenses tem nomes pequenos.</u>

Embora os livros analisados utilizassem inúmeras maneiras de apresentar uma distribuição, poucas vezes foi observada a tarefa de analisá-la.

**Técnica 1.1 e Discurso Teórico-Tecnológico:** Explorar a variabilidade em uma única distribuição por meio de uma tabela.

A leitura de uma tabela requer a mobilização de conceitos estatísticos básicos como variável e freqüência. A Tabela 12 apresenta uma distribuição de freqüência simples, em que a variável é qualitativa. Porém, em casos em que a variável é quantitativa discreta, por exemplo, número de irmãos, é possível que,

[22] Moda é uma medida de tendência central que é obtida a partir da identificação da observação (ou da categoria) que aparece com mais freqüência.

se esses conhecimentos não estiverem disponíveis completamente, haja confusão na interpretação da tabela.

Tabela 12 – Distribuição da freqüência dos alunos de acordo com o esporte preferido

| Esporte preferido | Número de alunos | % de alunos |
|---|---|---|
| Futebol | 50 | 62,5 |
| Vôlei | 20 | 25,5 |
| Natação | 10 | 12,5 |
| Total | 80 | 100,0 |

Analisar essa distribuição de freqüências permite observar que a maioria prefere futebol, mas nem todos os alunos. Ou seja, admite-se a variação, condição necessária para a existência da variável[23].

Como pode ser observada na Tabela 13 (a seguir), esta estratégia de analisar a distribuição de freqüências não é uma técnica muito utilizada pelos livros, embora todos eles utilizem inúmeros exemplos de distribuição de freqüências.

Para alguns autores dos livros analisados, interpretar a distribuição restringe-se à verificação da porcentagem de observações em cada categoria, com tarefas do tipo: qual a porcentagem de alunos que preferem natação? No entender desta autora, este tipo de tarefa pode permitir desenvolver o raciocínio sobre proporcionalidade, mas dificilmente promove o desenvolvimento do raciocínio sobre variação.

**Técnica 1.2 e Discurso Teórico-Tecnológico:** Explorar a variabilidade em uma única distribuição por meio de um gráfico de barras simples

O gráfico de barras é muito utilizado e explorado nos livros didáticos, principalmente com exemplos veiculados pela mídia. Este tipo de gráfico pode representar uma variável qualitativa, como o exemplo apresentado na Tabela 12 ou uma variável quantitativa discreta.

[23] Exceção se faz à variável degenerada no ponto, em que ela existe e a variância é zero.

Explorar a variabilidade no gráfico de barras simples não se resume à sua elaboração, mas à sua análise, que deve ter todos os componentes da Técnica 1.1, somados à interpretação de um plano cartesiano. O indivíduo que se propõe a analisar o gráfico de barras precisa identificar o eixo em que a variável está apresentada e o eixo que contém a freqüência de cada categoria de resposta da variável. Além disso, a leitura da escala no eixo que contém a freqüência é um fator muito importante. É comum ser ludibriado por um gráfico cujo eixo não tem início na origem (0,0), podendo distorcer a análise dos fatos.

Para o exemplo da Tabela 12, a Figura 10 apresenta seu gráfico de barras.

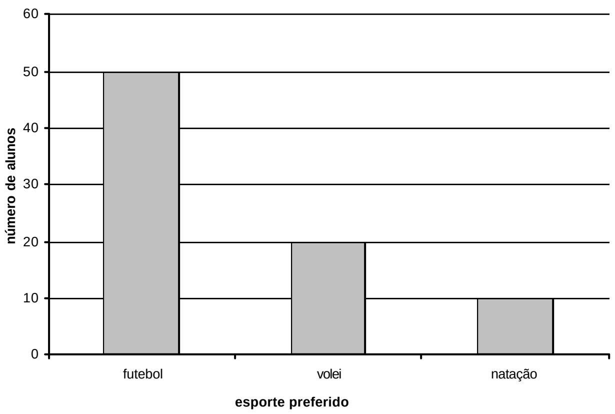



Figura 10 – Distribuição dos alunos segundo o esporte preferido

**Técnica 1.3 e Discurso Teórico-Tecnológico:** Explorar a variabilidade em uma única distribuição por meio de um gráfico de setores.

O gráfico de setores, popularmente conhecido como “gráfico de pizza”, é um recurso utilizado em meios de comunicação de massa devido à sua facilidade de interpretação.

Embora utilize conhecimentos geométricos como a construção de ângulos na circunferência, fazendo uso de transferidor, a sua leitura requer apenas a observação do maior setor, da maior “fatia da pizza”, que pode ser feita sem qualquer conhecimento geométrico, a partir de uma interpretação intuitiva.

Todos os livros analisados apresentam inúmeros exemplos desse tipo de gráfico, mas poucos exploram sua análise, como pode ser observado na Tabela 13.

A Figura 11 mostra que o setor cinza, que representa a categoria futebol, é muito maior que os demais, indicando a categoria da *maioria*.

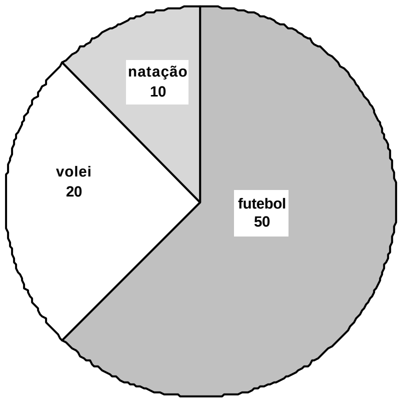



Figura 11 – Distribuição do número de alunos segundo o esporte preferido.

**Técnica 1.4 e Discurso Teórico-Tecnológico*:*** Explorar a variabilidade de um conjunto numérico representado em um gráfico de linhas

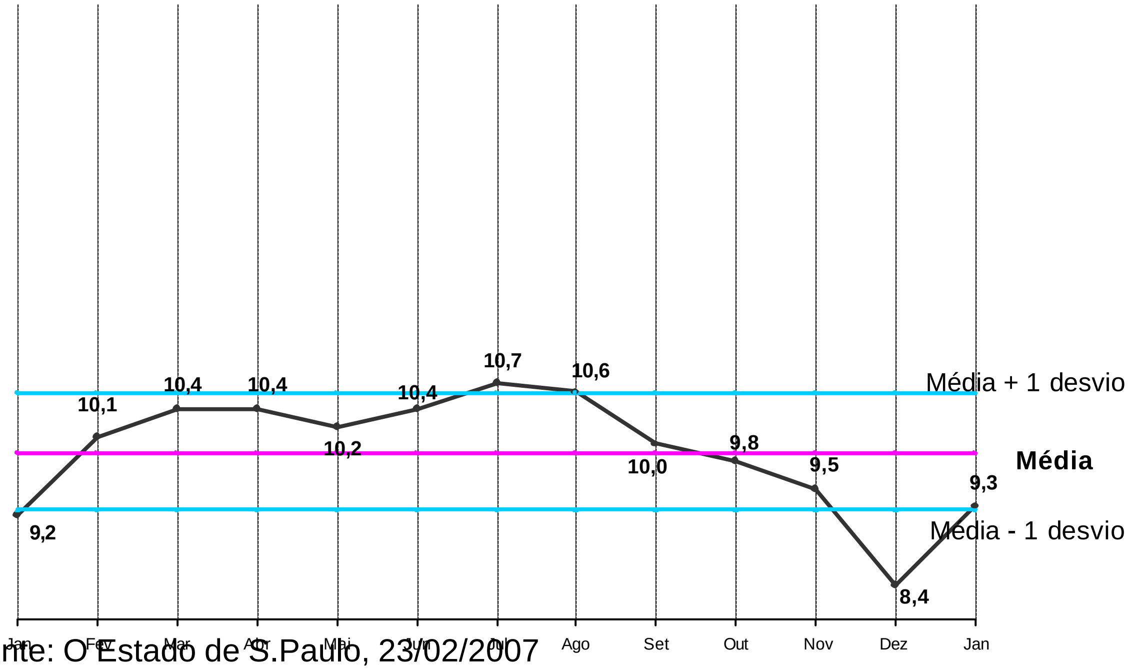
Diferentemente das técnicas 1.1, 1.2 e 1.3, neste caso não é elaborada uma distribuição de freqüências simples, mas apenas a disposição das observações da amostra em um gráfico de linhas, cujo objetivo é acompanhar uma variável ao longo do tempo. Ou seja, em cada ponto do gráfico há apenas uma observação, representada por um par ordenado (tempo; medida da variável), e não um conjunto de observações. Como ilustrado na Figura 12, o eixo das abscissas apresenta o tempo, em meses, e o eixo das ordenadas está implícito.

O objetivo do gráfico de linha é verificar se há uma regularidade da medida ao longo do tempo ou se há uma oscilação grande. Porém, essa comparação se faz entre as observações, que Loosen, Lioen e Lacante (1985) denominaram de *unalikeability* e que pouco contribui para o raciocínio sobre a variação em torno de uma medida de tendência central.

Em contrapartida, se o livro didático ou o professor de Matemática motivar o estudante a estimar a medida de tendência central e a medida de dispersão que representariam a distribuição, o raciocínio sobre variação em torno da medida de centro pode começar a se desenvolver.

Após as estimações das medidas, o professor pode solicitar que os alunos as calculem e as representem no mesmo sistema cartesiano, que permite uma visualização da distribuição e de sua representação numérica.

A Figura 12 apresenta um gráfico retirado do jornal O Estado de S.Paulo e as linhas que representam a média e o desvio padrão foram inseridas pela autora desta pesquisa.




Fonte: O Estado de S.Paulo, 23/02/2007

Figura 12 – Taxa de desemprego de Janeiro de 2006 a Janeiro de 2007 (em %)

É possível analisar na Figura 12 que há uma pequena oscilação na taxa de desemprego no Brasil, que justifica o fato de quase todas as observações estarem compreendidas no intervalo de um desvio padrão da média. Esta pode ser uma estratégia de utilizar o gráfico de linhas para desenvolver o raciocínio sobre variação.

A análise encontrada nos livros resume-se à identificação da regularidade (ou não) das observações e o livro L3 é o único que explica como fazer essa análise: “nele se visualizam o ponto de máximo e mínimo, entre os quais a variável tem um determinado comportamento” (SMOLE e DINIZ, 2003b, p. 11).

Na Tabela 13 só foram contabilizados os gráficos que apresentavam algum tipo de análise que pudesse permitir ao professor de Matemática utilizar o exemplo (ou exercício) para estimular o raciocínio sobre a variação. Por esse motivo, o número de exemplos e exercícios é tão ínfimo.

Tabela 13 – Número de exemplos e exercícios das técnicas identificadas na Tarefa Matemática 1

| Livros | Número de | Técnicas | | | |
|---|---|---|---|---|---|
| | | 1.1 | 1.2 | 1.3 | 1.4 |
| L1 | Exemplos | --- | 1 | 1 | --- |
| | Exercícios | --- | 1 | 1 | --- |
| L2 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L3 | Exemplos | --- | --- | 1 | --- |
| | Exercícios | --- | --- | --- | --- |
| L4 | Exemplos | 1 | 3 | 3 | 1 |
| | Exercícios | --- | 1 | --- | 1 |
| L5 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L6 | Exemplos | 1 | --- | --- | --- |
| | Exercícios | 1 | --- | --- | --- |
| L7 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |

O livro L4 apresenta as técnicas 1.2, 1.3 e 1.4 juntas, no momento que faz uso da apresentação dos resultados de uma pesquisa sobre o perfil dos moradores de rua. Isso é muito interessante para o raciocínio sobre variação, pois alia a variação ao pensamento estatístico, a Estatística como ferramenta de pesquisa.

Esse mesmo livro utilizou três exercícios para solicitar o cálculo da média e do desvio padrão (Tarefa Matemática 8) de observações apresentadas num gráfico de linha, mas não explorou a possibilidade de representar tais medidas no mesmo sistema cartesiano.

**Tarefa 2:** Explorar a variabilidade na comparação de duas distribuições

Utilizar a comparação de duas distribuições é uma estratégia muito mais vantajosa para estimular o raciocínio sobre variação, pois os recursos mentais utilizados pelo indivíduo requerem uma representação de cada grupo, que pode ser uma medida de tendência central acompanhada de uma medida de variação.

**Técnica 2.1 e Discurso Teórico-Tecnológico** *:* A partir do gráfico de barras múltiplas (variável qualitativa)

Na apresentação do capítulo de Estatística, Smole e Diniz (2003) apresentaram vários gráficos estatísticos extraídos de meios de comunicação de massa e um deles está reproduzido na Figura 13, que tratava do estilo de vida da população de hipertensos e diabéticos.

Embora as autoras não estivessem tratando das medidas de variação, elas explicaram que o gráfico de barras múltiplas é empregado para “facilitar a comparação” entre dois fenômenos.

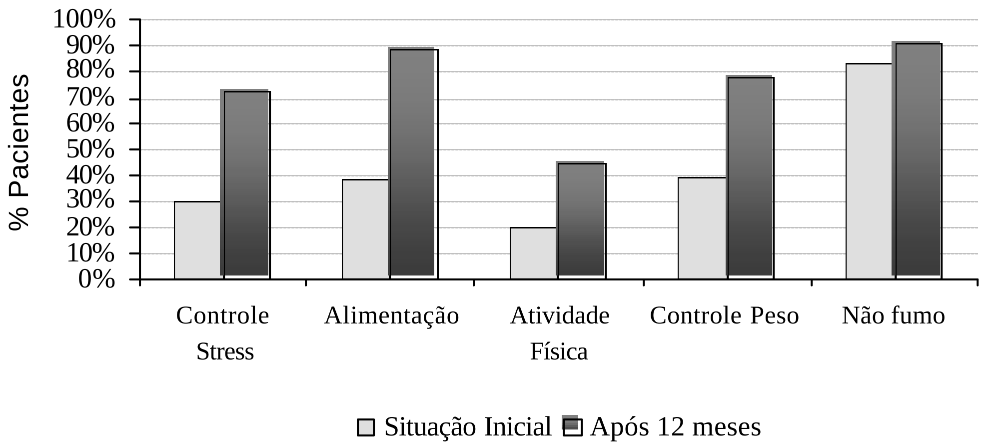



Figura 13 - Gráfico apresentado por Smole e Diniz (2003a, p. 43)

Provavelmente o gráfico foi construído para verificar se havia diferença no estilo de vida dos pacientes no início do tratamento e após um ano. Logo, para fazer essa comparação, é necessário observar cada característica dos pacientes (stress, alimentação, etc.) e identificar a porcentagem na situação inicial e após 12 meses. Fazendo essa análise, é possível notar que houve variação da situação inicial para a situação final, com um aumento na porcentagem de pacientes em todas as categorias.

Outra análise que poderia ser feita é explorar se uma categoria variou mais do que a outra. E o aluno trabalharia com a diferença entre as porcentagens, notando que a porcentagem de pacientes que não fumava na situação inicial era

82% e que no final de um ano passou a 91% (um aumento de 9%) enquanto, nas outras categorias, a porcentagem de pacientes praticamente dobrou.

Ainda, observando a maioria dos pacientes nas categorias, poderia ser traçado um perfil dos pacientes na situação inicial e final. É possível notar que, na situação inicial, a maioria dos pacientes não fumava e que, na situação final, a maioria dos pacientes continuava não fumando e também controlava o stress, a alimentação e o peso.

Ou seja, o fato de fazer uma análise desse gráfico, mesmo sem a utilização formal dos conceitos, faz com que o aluno comece a observar o todo, verificar onde está a maioria, observar a variação de uma categoria para a outra, etc.

As autoras apresentaram o gráfico de maneira ilustrativa, mas não fizeram nenhuma análise, deixando para o professor de Matemática essa atividade.

**Técnica 2.2 e Discurso Teórico-Tecnológico*:*** Por meio de dois gráficos de setores

Como apresentado na Figura 14, os dois gráficos de setores permitem observar se há diferença no tamanho da categoria mais representativa em cada grupo. Neste caso, a maioria dos alunos do 1A prefere basquete e a maioria dos alunos do 1B prefere futebol.

Como salientado na Técnica 1.3, para realizar esta análise basta que o estudante verifique o maior setor circular em um gráfico e compare-o com o outro.

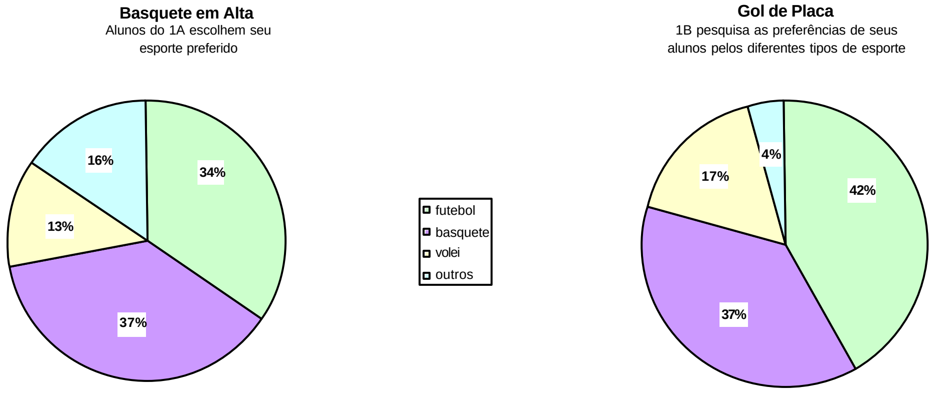



Figura 14 – Porcentagem de alunos do 1A e 1B segundo o esporte preferido, apresentado por Smole e Diniz (2003a, p.57)

**Técnica 2.3 e Discurso Teórico-Tecnológico***:* A partir do gráfico de barras múltiplas (variável quantitativa discreta)

A idéia de variação pode ser explorada na análise de um gráfico de colunas múltiplas, tal como feito por Ben-Zvi (2004), em que é possível identificar as características da maioria das observações em cada grupo.

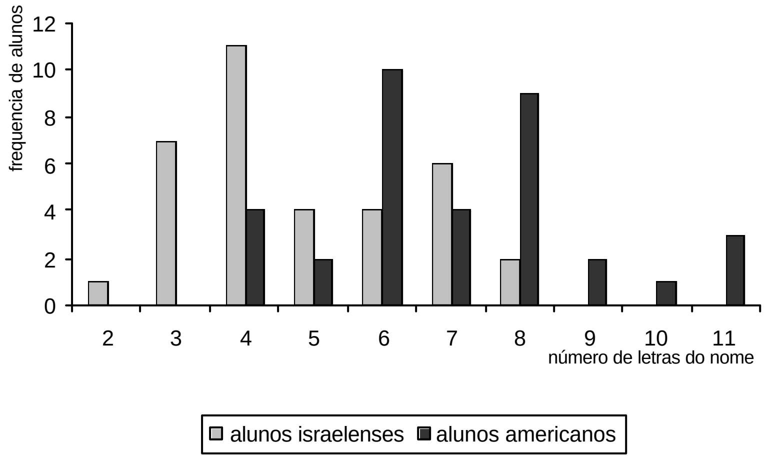



Figura 15 - Gráfico utilizado por Ben-Zvi (2004)

Para analisar a variação dos dados, é preciso ler ambos os eixos do gráfico. No eixo das abscissas, é necessário identificar a amplitude dos valores da variável (At). No exemplo, a amplitude do tamanho dos nomes americanos é 11 - 4 = 7 e a amplitude do tamanho dos nomes israelenses é 8 – 2 = 6.

A simples observação da amplitude total não é suficiente, pois observa-se que elas são semelhantes para os dois grupos. A medida precisa ser associada à observação dos valores máximo e mínimo de cada grupo. O número de letras dos nomes varia entre dois e onze, sendo que só existem nomes israelenses com duas e três letras e que também só existem nomes americanos com nove, dez e onze letras.

É preciso, também, ler o eixo das ordenadas, que identifica a quantidade de alunos em cada categoria de tamanho de nome (freqüência). E o que aparece no gráfico é a existência de muitos alunos israelenses com 4 letras no nome (moda) e muitos alunos americanos com 6 e 8 letras no nome.

Diante da leitura dos eixos, o aluno poderá identificar a característica da maioria dos sujeitos. No caso, é possível notar que a maioria dos nomes dos alunos israelenses tem menos letras que a maioria dos nomes dos alunos americanos. Esse tipo de raciocínio leva à identificação da densidade de freqüência e pode permitir ao aluno estimar (imaginar) a média e a medida de dispersão em torno da média, que então, podem ser calculadas e comparadas com a estimativa realizada.

**Técnica 2.4 e Discurso Teórico-Tecnológico*:*** A partir do gráfico de linhas (variável quantitativa discreta)

Na introdução do capítulo de Estatística, Smole e Diniz (2003a) apresentam um gráfico de linhas ilustrado, em que havia um avião no final de cada linha. Uma aproximação do gráfico está apresentada na Figura 16.

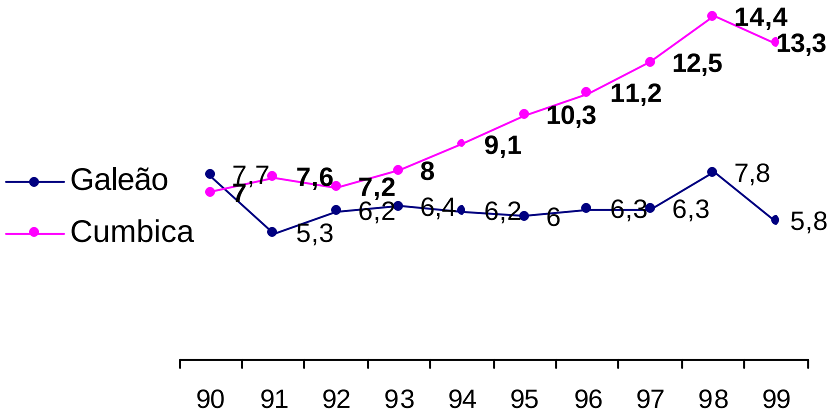



Figura 16 - Gráfico apresentado por Smole e Diniz (2003a,p. 44)

Cada linha da Figura 16 representa o número de passageiros (em milhões) em um aeroporto, no período de 1990 até 1999. Ao lado de cada ponto do gráfico está apresentada a segunda coordenada, que se refere à medida do eixo das ordenadas, permitindo a omissão deste eixo, técnica pouco recomendada por professores de Matemática, mas muito utilizada na mídia.

É interessante observar que a representação decimal do número de passageiros não diz respeito a uma variável contínua, pois 7,7 é uma forma abreviada da variável quantitativa discreta (7,7 significa 7.700.000).

Tal como discutido na Técnica 1.4, o objetivo principal é verificar se existe variação ao longo do tempo (nesse caso, anos). É possível observar que o aeroporto de Cumbica teve uma variação crescente do número de passageiros ao longo dos anos e o aeroporto do Galeão manteve-se estável (com exceção do ano de 98, cujo aumento foi explicado pelo incêndio acontecido no outro aeroporto do Rio de Janeiro). Assim, poder-se-ia supor uma medida de variação bem menor para o número de passageiros no aeroporto do Galeão, se comparado ao aeroporto de Cumbica.

As autoras apresentaram o gráfico de maneira ilustrativa, sem explorar nenhuma análise. Explorar um gráfico, mesmo que intuitivamente, auxilia o desenvolvimento do pensamento estatístico (como foi feito por Ben-Zvi, 2002).

A Tabela 14 apresenta o número de exemplos e exercícios sobre a tarefa de explorar a variação (variabilidade) em diferentes representações de duas ou mais distribuições.

Tabela 14 – Número de exemplos e exercícios das técnicas identificadas na tarefa Matemática 2.

| Livros | Número de | Técnica | | | |
|---|---|---|---|---|---|
| | | 2.1 | 2.2 | 2.3 | 2.4 |
| L1 | Exemplos | --- | --- | --- | 1 |
| | Exercícios | --- | 1 | --- | --- |
| L2 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | 1 | --- | --- |
| L3 | Exemplos | 1 | 1 | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L4 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L5 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L6 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L7 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |

É possível observar nessa tabela que não houve utilização da técnica 2.3 por nenhum livro, mas escolheu-se deixá-la relatada neste trabalho, pois é reforçada pelos estudos recentes na área. Porém, a restrita utilização das técnicas da Tarefa 2 não é privilégio da técnica 2.3, pois há apenas cinco situações em que a comparação de duas distribuições aconteceu.

**Tarefa 3:** Explorar a idéia de variação acima e abaixo da média aritmética

Em geral, a idéia de variação não só pode ser estimulada na representação gráfica, mas principalmente na apresentação das medidas de tendência central. Utilizando apenas a média aritmética, foi objetivo verificar se os autores dos livros estimulavam algum raciocínio sobre variação quando apresentavam exemplos e exercícios sobre a medida. Foram encontradas duas técnicas, descritas a seguir.

**Técnica 3.1 e o Discurso Teórico-Tecnológico*:*** A partir do gráfico de barras simples

Este gráfico foi adaptado de Loosen, Lioen e Lacante (1985) e tem como objetivo mostrar as observações de uma variável em forma gráfica.

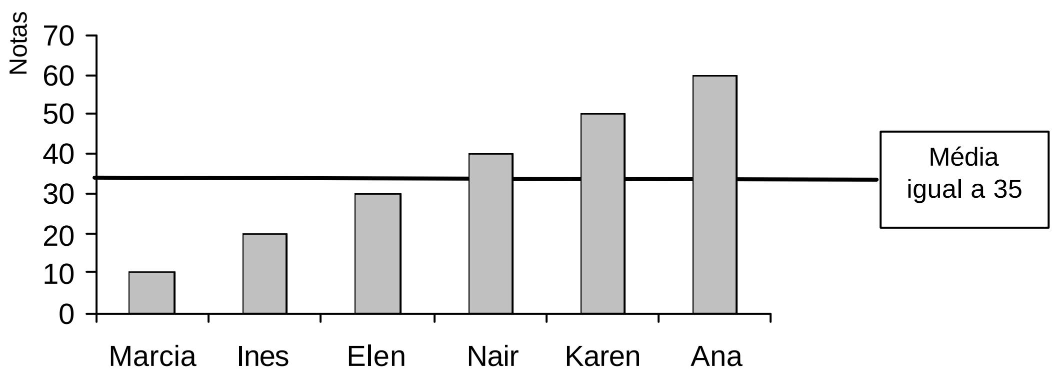



Figura 17 - Gráfico adaptado de Loosen, Lioen e Lacante (1985)

O gráfico ajuda o aluno a visualizar que Márcia, Inês e Elen tem notas abaixo da média do seu grupo e que Nair, Karen e Ana estão acima da média, desde que a linha da média esteja traçada. Esse tipo de representação permite identificar que nem todos os valores são iguais ao valor da média e que pode acontecer que nenhum valor seja igual à média.

Um obstáculo didático já identificado por Loosen, Lioen e Lacante (1985) é o fato de que o estudante sempre associa o gráfico de barras com uma distribuição de freqüência simples (e ele não está errado). Portanto, a utilização do gráfico de barras, em que cada barra representa apenas uma observação requer uma leitura mais atenta dos eixos: o eixo das abscissas contém as observações e o eixo das ordenadas contém o valor da variável para cada observação. Se a linha da média não estiver explícita, o aluno precisa calcular a média e traçá-la (ou imaginá-la).

Sem a representação da média, os alunos podem ser induzidos a observar a variação existente entre os dados (*unalikeability*) e não a variação em torno da média.

**Técnica 3.2 e o Discurso Teórico-Tecnológico***:* a partir das observações coletadas

Esta técnica é semelhante à anterior, porém as observações estão apresentadas na maneira original, tal como foram obtidas.

Para realizar a tarefa de identificar os valores que estão acima e abaixo da média, é necessário calcular a média aritmética e compará-la com cada observação. Para exemplificar, utiliza-se o exemplo de Giovanni, Bonjorno e Giovanni Jr (2002), apresentado na Tabela 15, em que os autores pedem os meses em que o número de nascimentos ficou acima da média.

Tabela 15 – Número de nascimentos por mês, em uma maternidade.

| MÊS | Jan. | Fev. | Mar. | Abr. | Maio | Jun. | Jul. | Ago. | Set. | Out. | Nov. | Dez. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NASCIMENTO | 38 | 25 | 42 | 30 | 29 | 47 | 18 | 36 | 38 | 43 | 49 | 37 |

Embora seja uma técnica simples, esse estímulo de calcular a média e olhar novamente a distribuição pode permitir ao estudante perceber a variação em torno da média e não a variação entre os meses.

Esta técnica, embora não visual, é essencial para desenvolver um dos aspectos mais importantes do raciocínio sobre variação: a comparação dos valores com a média e faz surgir a necessidade da medida de variação.

A Tabela 16 aponta que ambas as técnicas, essenciais para desenvolver a necessidade da medida de variação, são apresentadas como exemplo em três livros e como exercícios apenas em L3. Porém, se o professor não trabalhar os exercícios, a análise pode não ser explorada.

É interessante notar que o L3, quando vai trabalhar o conceito de média, já induz o aluno a pensar na variação em torno dela. “Utilizamos a média para observar o valor em torno do qual os dados se distribuem. Ela é tanto mais representativa quanto menor for a variação dos dados” (SMOLE e DINIZ, 2003a, p. 67).

É importante destacar que o livro L2 utiliza a técnica 3.2 quando inicia o capítulo de medidas de variação, calculando o desvio relativo, já discutido anteriormente.

Tabela 16 - Número de exemplos e exercícios das técnicas identificadas na tarefa Matemática 3.

| Livros | Número de | Técnica | |
|---|---|---|---|
| | | 3.1 | 3.2 |
| L1 | Exemplos | 1 | 1 |
| | Exercícios | --- | --- |
| L2 | Exemplos | --- | 1 |
| | Exercícios | --- | --- |
| L3 | Exemplos | --- | --- |
| | Exercícios | --- | 2 |
| L4 | Exemplos | --- | 1 |
| | Exercícios | --- | --- |
| L5 | Exemplos | --- | --- |
| | Exercícios | --- | --- |
| L6 | Exemplos | --- | --- |
| | Exercícios | --- | --- |
| L7 | Exemplos | --- | --- |
| | Exercícios | --- | --- |

**Tarefa 4: Calcular a amplitude**

Como já discutido no Capítulo 2, a amplitude é uma medida intuitiva de variação e muito útil na leitura de gráficos, tabelas, banco de dados, etc...

Todos os livros analisados utilizavam a amplitude desde o capítulo de gráficos para variável contínua (histograma), mesmo sem nomeá-la.

No capítulo de variação desses livros, esta medida teve pouca ênfase. Foram observados três exercícios em L3 e um exemplo em L6.

Outro aspecto a ser observado é que os livros L3, L4 e L5 solicitam, em um exercício, a observação dos valores máximos e mínimos, mas não solicitam o cálculo da amplitude nem a constatação de uma grande diferença entre o máximo e o mínimo.

**Tipo de tarefa 5: Calcular o desvio médio**

Calcular o desvio médio, assim como a variância e o desvio padrão foi considerado um tipo de tarefa *T*, pois as tarefas incluídas neste tipo diferenciam-se na maneira com que as observações estão apresentadas.

O cálculo do desvio médio não é uma tarefa requerida em todos os livros de Matemática analisados, diferentemente das outras medidas.

**Tarefa 5.1**: Calcular o desvio médio com o conjunto numérico

**Técnica 5.1 e o Discurso Teórico-Tecnológico**

Como discutido no Capítulo 2, o desvio médio é obtido a partir da fórmula $dm = \frac{\sum |X_i - \mu|}{N}$, cuja seqüência de cálculos está enunciada a seguir:

a) Calcular a média: somar todos os valores do conjunto e dividir pela quantidade: $\left(\mu = \frac{\sum X_i}{N}\right)$;
b) Calcular as distâncias de cada valor em relação à média;
c) Obter o módulo de cada valor obtido na etapa anterior;
d) Somar os módulos;
e) Dividir o resultado do somatório por n (número total de observações);

Essa seqüência de cálculos permite ao estudante obter o resultado numérico que representa o desvio médio. Alguns livros relembram o uso do somatório, explicando que, dados três valores, a soma é representada por $\sum_{i=1}^{3} x_i = x_1 + x_2 + x_3$.

**Tarefa 5.2** – calcular o desvio médio a partir de uma distribuição de freqüência simples.

**Técnica 5.2 e o Discurso Teórico-Tecnológico**

Para se obter o desvio médio a partir de uma distribuição de freqüências simples, utiliza-se a fórmula $dm = \frac{\sum |X_i - \mu| \cdot f_i}{\sum f_i}$, cuja seqüência de procedimentos está descrita a seguir.

a) Calcular a média aritmética $\left(\mu = \frac{\sum X_i \cdot f_i}{\sum f_i}\right)$;
b) Calcular as distâncias de cada valor em relação à média;
c) Obter o módulo de cada valor obtido na etapa anterior;
d) Multiplicar essas distâncias, em módulo, pelo número de casos ( $f_i$ );
e) Somar os resultados da multiplicação;
f) Dividir o resultado do somatório por n (número total de observações);

**Tarefa 5.3** – calcular o desvio médio a partir de uma distribuição de freqüência com dados agrupados.

**Técnica 5.3 e o Discurso Teórico-Tecnológico**

Utilizando a fórmula $dm = \frac{\sum |X_i - \mu| \cdot f_i}{\sum f_i}$, o estudante precisa saber que cada valor do conjunto de dados repete-se $f_i$ vezes e que o valor de $X_i$ será o ponto médio de cada classe da distribuição de freqüências. A seqüência de procedimentos está apresentada a seguir:

a) Encontrar o ponto médio de cada classe, a partir dos limites da classe, utilizando a seguinte fórmula $\frac{(\lim_{superior} + \lim_{inferior})}{2}$;
b) Calcular a média aritmética $\left(\mu = \frac{\sum X_i \cdot f_i}{\sum f_i}\right)$, em que $X_i$ é o ponto médio da classe;
c) Calcular as distâncias de cada valor (ponto médio) em relação à média;
d) Obter o módulo de cada valor obtido na etapa anterior;
e) Multiplicar essas distâncias, em módulo, pelo número de casos ( $f_i$ );
f) Somar os resultados da multiplicação;
g) Dividir o resultado do somatório por n (número total de observações);

É possível observar que a teoria que dá suporte para o tipo de Tarefa 5 encontra-se na álgebra e na aritmética e não na Estatística. A Tabela 17 apresenta o número de exemplos e exercícios do tipo de tarefa 5 encontrado nos livros analisados.

É possível notar que somente L1, L2 e L7 trabalham o desvio médio, uma opção dos autores, em que L1 apresenta uma definição de procedimento e L2 apresenta uma interpretação dessa medida de variação que pode possibilitar ao estudante compreender o motivo de sua necessidade.

> O desvio médio absoluto é uma medida associada à amostra como um todo; quando no exemplo anterior dizemos que $D_{ma}$ = 1,2 ano estamos afirmando que, **em média**, os elementos da amostra se afastam 1,2 ano da média aritmética, para cima ou para baixo. (PAIVA, 1995 p. 273).

Tabela 17 - Número de exemplos e exercícios das técnicas identificadas na tarefa Matemática 5.

| Livros | Número de | Técnica | | |
|---|---|---|---|---|
| | | 5.1 | 5.2 | 5.3 |
| L1 | Exemplos | 1 | 1 | 1 |
| | Exercícios | 2 | 2 | 3 |
| L2 | Exemplos | 1 | 1 | 2 |
| | Exercícios | 2 | 1 | 3 |
| L3 | Exemplos | --- | --- | --- |
| | Exercícios | --- | --- | --- |
| L4 | Exemplos | --- | --- | --- |
| | Exercícios | 1 | --- | --- |
| L5 | Exemplos | --- | --- | --- |
| | Exercícios | --- | --- | --- |
| L6 | Exemplos | --- | --- | --- |
| | Exercícios | --- | --- | --- |
| L7 | Exemplos | 1 | --- | 1 |
| | Exercícios | --- | --- | 6 |

**Tipo de Tarefa 6: Calcular a variância**

**Tarefa 6.1:** Calcular a variância a partir do conjunto numérico

**Técnica 6.1 e o Discurso Teórico-Tecnológico**

Como apresentado no Capítulo 2, a variância é obtida pela fórmula $\sigma^2 = \frac{\sum(X_i - \mu)^2}{N}$, utilizada por todos os livros, exceto L7, que utilizou a fórmula alternativa.

Os cálculos são idênticos à Técnica 5.1, exceto no item referente ao cálculo do módulo, que deve ser substituído por “elevar ao quadrado as distâncias em relação à média”.

**Tarefa 6.2** Calcular a variância a partir de uma distribuição de freqüências simples

**Técnica 6.2 e o Discurso Teórico-Tecnológico**

Utilizando a fórmula $\sigma^2 = \frac{\sum(X_i - \mu)^2 \cdot f_i}{\sum f_i}$, o estudante precisa saber que cada valor do conjunto de dados repete-se $f_i$ vezes. O procedimento é idêntico à Técnica 5.2, exceto no item referente ao cálculo do módulo, que deve ser substituído por “elevar ao quadrado as distâncias em relação à média”.

**Tarefa 6.3** Calcular a variância a partir de uma distribuição de freqüências com dados agrupados.

**Técnica 6.3 e o Discurso Teórico-Tecnológico**

Utilizando a fórmula $\sigma^2 = \frac{\sum(X_i - \mu)^2 \cdot f_i}{\sum f_i}$, o aluno precisa saber que cada valor do conjunto de dados repete-se $f_i$ vezes e que o valor de $X_i$ será o ponto médio de cada classe da distribuição de freqüências. O procedimento é igual à Técnica 5.3, substituindo o cálculo do módulo por “elevar ao quadrado as distâncias em relação à média”.

A Tabela 18 apresenta o número de exemplos e exercícios identificados nos livros, permite verificar que é uma medida apresentada pelos autores e solicitada em exercícios.

É interessante observar que os livros L2, L4 e L5 explicam que a razão pela qual a variância não é muito utilizada é a incompatibilidade da sua unidade de medida (ao quadrado) com a unidade de medida da média, o que pode permitir ao estudante compreender a necessidade do desvio padrão.

Tabela 18 - Número de exemplos e exercícios das técnicas identificadas na tarefa Matemática 6.

| Livros | Número de | Técnica | | |
|---|---|---|---|---|
| | | 6.1 | 6.2 | 6.3 |
| L1 | Exemplos | 1 | --- | 1 |
| | Exercícios | 1 | 1 | 2 |
| L2 | Exemplos | 1 | 1 | --- |
| | Exercícios | 2 | 1 | 3 |
| L3 | Exemplos | --- | 2 | 1 |
| | Exercícios | 1 | 1 | 1 |
| L4 | Exemplos | 1 | --- | --- |
| | Exercícios | 3 | --- | --- |
| L5 | Exemplos | 2 | --- | 1 |
| | Exercícios | --- | 1 | --- |
| L6 | Exemplos | 1 | --- | 2 |
| | Exercícios | 2 | 1 | 1 |
| L7 | Exemplos | --- | --- | --- |
| | Exercícios | --- | --- | --- |

**Tipo de Tarefa : Calcular o desvio padrão**

**Tarefa 7.1:** Calcular o desvio padrão a partir do conjunto numérico

**Técnica 7.1 e o Discurso Teórico-Tecnológico**

Se o valor da variância já foi obtido, a obtenção do desvio padrão reduz-se à extração da raiz quadrada da variância. Caso contrário, primeiro calcula-se a variância (Técnica 6.1) e depois extrai-se a raiz quadrada.

**Tarefa 7.2** Calcular o desvio padrão a partir de uma distribuição de freqüências simples

**Técnica 7.2 e o Discurso Teórico-Tecnológico**

Se o aluno já calculou a variância, o desvio padrão será a raiz quadrada positiva da variância. Caso contrário, primeiro calcula-se a variância (Técnica 6.2) e depois extrai-se a raiz quadrada.

**Tarefa 7.3** Calcular o desvio padrão a partir de uma distribuição de freqüências com dados agrupados

**Técnica 7.3 e o Discurso Teórico-Tecnológico**

Se o aluno já calculou a variância, o desvio padrão será a raiz quadrada positiva da variância. Caso contrário, primeiro calcula-se a variância (Técnica 6.3) e depois extrai-se a raiz quadrada.

Tabela 19 - Número de exemplos e exercícios das técnicas identificadas na tarefa Matemática 7.

| Livros | Número de | Técnica | | |
|---|---|---|---|---|
| | | 7.1 | 7.2 | 7.3 |
| L1 | Exemplos | 1 | --- | 1 |
| | Exercícios | 3 | 1 | 2 |
| L2 | Exemplos | 2 | 1 | 2 |
| | Exercícios | 3 | 1 | 3 |
| L3 | Exemplos | --- | 2 | 1 |
| | Exercícios | 7 | 4 | 2 |
| L4 | Exemplos | 2 | --- | 1 |
| | Exercícios | 13 | 2 | 7 |
| L5 | Exemplos | 2 | --- | 1 |
| | Exercícios | 3 | --- | 1 |
| L6 | Exemplos | --- | --- | 1 |
| | Exercícios | 4 | 1 | 2 |
| L7 | Exemplos | --- | --- | 1 |
| | Exercícios | --- | --- | 6 |

Observando o complexo de tarefas relativas aos cálculos das medidas de variação em torno da média, é possível notar que todos os livros apresentam exemplos e muitos exercícios (exceção é feita para o desvio médio). Ou seja, calcular a medida é uma tarefa abordada, com diferentes técnicas, tanto em exemplos como em exercícios. O fato é saber se o estudante consegue compreender o resultado obtido na Tarefa 7, o que está discutido na Tarefa 8.

**Tarefa 8:** Interpretar o valor do desvio padrão

Segundo Konold e Pollatsek (2002, p. 260), a maior proposta de calcular estatísticas tais como a média ou a mediana é representar um sinal num processo ruidoso, "entretanto, esta idéia não aparece nos currículos e nos documentos padrões".

Utilizando a metáfora dos autores, que pode ser uma interpretação do desvio padrão, ele é uma medida do ruído deste sinal (que é a média).

A seguir estão descritas as maneiras (técnicas) com que os livros analisam o desvio padrão.

**Técnica 8.1 e o Discurso Teórico-Tecnológico:** Medida da homogeneidade do grupo

Um conjunto numérico é mais regular, ou seja, mais homogêneo, quanto menor for o valor do seu desvio padrão e Dante (2000, p. 288) formaliza essa interpretação quando expressa que "quanto mais próximo de 0 é o desvio padrão, mais homogênea é a distribuição dos valores da variável".

O complemento dessa análise torna-a mais completa: quanto menor for o valor do desvio padrão, mais homogêneas são as observações <u>em relação à média</u>, o que não foi encontrado em nenhum livro analisado.

Todos os livros analisados (exceto L7) iniciam a explicação do desvio padrão na comparação de duas distribuições com mesma média aritmética, solicitando que seja observado o grupo mais regular, mais homogêneo, ou seja, aquele que apresenta o menor desvio padrão.

Essa análise é adequada quando se trata de grupos com mesma média, mas precisa de um complemento quando os grupos apresentam médias diferentes. A comparação da homogeneidade de dois conjuntos cujas médias são diferentes requer a utilização de uma medida relativa de variação, conhecida como coeficiente de variação $\left( CV = \frac{s}{\overline{x}} \cdot 100 \right)$, já discutida no Capítulo 2 deste trabalho.

Foram encontrados dois exemplos em que se comparavam a homogeneidade de grupos com médias diferentes e os autores não faziam qualquer referência ao CV. Em nenhum dos casos o resultado estava errado, haja vista que foram apresentados grupos com características muito diferentes, mas a técnica para se chegar ao resultado não foi discutida.

Vale destacar que o estudo de Delmas e Liu (2005) tinha como objetivo desenvolver uma idéia intuitiva do desvio padrão e os autores trabalharam com a comparação do valor absoluto do desvio padrão, ou seja, sem levar em consideração a variação relativa.

Os livros L3 e L4 são os que mais exploraram esse tipo de análise, embora o livro L3 apresentasse a resposta do exercício como sendo pessoal, ou seja, deixava para o professor de Matemática o desenvolvimento da técnica.

**Técnica 8.2 e o Discurso Teórico-Tecnológico*:*** contagem do número de observações no intervalo $\left[\overline{x} - s, \overline{x} + s\right]$

Como apresentado no Capítulo 2, solicitar a tarefa de contar o número de observações num intervalo de *n* desvios padrão da média pode levar o estudante a identificar que há uma tendência nesta porcentagem e a inserção de técnicas como o Teorema de Tchebichev ou a Distribuição Normal pode se fazer de maneira mais natural.

Para realizar essa tarefa, o estudante precisa calcular inicialmente a média, o desvio padrão e o intervalo $\left[\overline{x} - s, \overline{x} + s\right]$, a partir da soma e da subtração do valor do desvio padrão à média aritmética.

A partir da obtenção do intervalo, o estudante precisa contar o número de observações que está no intervalo. Uma variável didática essencial para a realização dessa tarefa é a escolha do formato da distribuição. Realizar essa tarefa para mais de uma variável com distribuição normal e também realizar essa tarefa para mais de uma variável assimétrica pode permitir ao estudante verificar que existe uma tendência de porcentagens de observações no intervalo $\left[\overline{x} - s, \overline{x} + s\right]$ de acordo com o formato da distribuição.

Somente o livro L3 insere formalmente esta técnica, com o objetivo de introduzir a Distribuição Normal de Probabilidades.

**Técnica 8.3 e o Discurso Teórico-Tecnológico :** Porcentagem de observações *no intervalo* $[\mu - \sigma, \mu + \sigma]$de acordo com a Distribuição Normal

Se uma variável segue uma distribuição normal, aproximadamente 68% das observações estão compreendidas no intervalo $[\mu - \sigma, \mu + \sigma]$, aproximadamente 95% das observações estão compreendidas no intervalo $[\mu - 2\sigma, \mu + 2\sigma]$ e aproximadamente 99% das observações estão no intervalo $[\mu - 3\sigma, \mu + 3\sigma]$, embora possa ser obtida a porcentagem de observações no intervalo de qualquer valor real de desvio padrão.

Para aplicar a técnica, calcula-se a média aritmética, o desvio padrão e o intervalo $[\mu - k\sigma, \mu + k\sigma]$. Depois se estima a porcentagem de observações no intervalo de *n* desvios padrão da média, sem que seja necessário aplicar a técnica 8.2.

A Distribuição Normal foi tratada pelos livros L3 e L6, apresentando a porcentagem de dados no intervalo de um, dois e três desvios padrão da média. O livro L7 faz uma pequena nota sobre o intervalo de um desvio padrão da média, sem apresentar nenhum exemplo.

O livro L3 apresenta a Distribuição Normal no capítulo Probabilidade e Estatística, presente no terceiro volume da coleção, que também aborda as definições e propriedades de probabilidades, probabilidade freqüentista, a lei dos grandes números e a associação entre a probabilidade e as medidas descritivas denominadas média e desvio padrão. Talvez pelo fato de que o assunto seja tratado no capítulo de Probabilidade, os exercícios propostos priorizam análises como: *a probabilidade de pertencer ao intervalo é x%* , mas também apresentam a análise: *x% de observações estão no intervalo [a, b].*

O discurso teórico-tecnológico desse bloco saber-fazer abrange não só a álgebra e a aritmética como também a Teoria de Probabilidades e medidas descritivas, sendo que as duas últimas pertencem à Estatística.

**Técnica 8.4 e o Discurso Teórico-Tecnológico:** Comparação do valor do desvio padrão antes e após a divisão da distribuição em dois grupos homogêneos.

Apenas o L4 apresenta esta técnica. Com base em um banco de dados, eles utilizam a seguinte seqüência de tarefas:

a) Calcular a média e o desvio padrão de todas as observações de uma variável;
b) Dividir as observações de acordo com a orientação dos autores;
c) após a divisão das observações, calcular a média e o desvio padrão de cada grupo;
d) comparar a média e o desvio padrão antes e após a divisão do grupo.

Os autores de L4 utilizam o exemplo da população urbana de sete continentes e solicitam que sejam divididas da seguinte maneira: Bloco I (África, América Central e Ásia) e o Bloco II (América do Norte, América do Sul, Europa e Oceania). São calculadas as medidas de todos juntos e de cada grupo e é possível notar nitidamente que a divisão dos grupos permitiu a obtenção de distribuições mais homogêneas e que, no caso, cada média representa melhor seu grupo do que a média geral e a variação é muito menor.

Nos exercícios, os autores não só solicitaram atividade exatamente como está descrita acima, como também solicitaram a retirada de valores muito diferentes dos demais e pediram para calcular novamente as medidas e compará-las antes e após a retirada.

Tabela 20 - Número de exemplos e exercícios das técnicas identificadas na tarefa Matemática 8.

| Livros | Número de | Técnica | | | |
|---|---|---|---|---|---|
| | | 8.1 * | 8.2 | 8.3 | 8.4 |
| L1 | Exemplos | 1* | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L2 | Exemplos | 1* | --- | --- | --- |
| | Exercícios | 1* | --- | --- | --- |
| L3 | Exemplos | 1* | 2 | 1 | --- |
| | Exercícios | 2* + 1 | 1 | 5 | --- |
| L4 | Exemplos | 1* +1 | --- | --- | 1 |
| | Exercícios | 4* | 1 | --- | 3 |
| L5 | Exemplos | 1* + 1 | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L6 | Exemplos | 1* | --- | 2 | --- |
| | Exercícios | --- | --- | --- | --- |
| L7 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | 2 | --- |

O símbolo * representa os exercícios e/ou exemplos em que os grupos têm a mesma média.

**Tarefa 9:** desenvolver as propriedades das medidas

As propriedades da variância estão descritas e demonstradas no Capítulo 2 deste trabalho e podem ser estendidas para o desvio padrão.

Nenhum dos livros didáticos apresentou formalmente as propriedades. O livro L5 fez um exemplo da técnica 9.2 e os outros livros deixaram exercícios para serem resolvidos acerca dessa e das outras propriedades, sem que houvesse o desenvolvimento da técnica no capítulo de respostas dos exercícios.

Dessa maneira, os livros deixam a cargo do professor de Matemática a busca pela técnica e pela didática para resolvê-la.

**Técnica 9.1 e o Discurso Teórico-Tecnológico**: Soma dos desvios em relação à média é zero

Como propriedade da média, tem-se que a soma dos desvios em relação à média é zero. Esta propriedade já foi trabalhada por Strauss e Bichler (1988) e é a justificativa da utilização do módulo das diferenças, quando se calcula o desvio médio e, também, da utilização do quadrado das diferenças quando se calcula a variância.

O livro L2 apresenta, nas respostas dos exercícios a demonstração desta propriedade: $\sum_{i=1}^{n}\left(x_i - \bar{x}\right) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = n\bar{x} - n\bar{x} = 0$, enquanto que o livro L4 apenas relata que a soma é zero.

Para que o estudante apresente uma demonstração como essa, ele precisa saber que a média aritmética é: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \Rightarrow \sum_{i=1}^{n} x_i = n\bar{x}$ e que a soma de uma constante (a média) é *n* vezes essa constante e, a partir de então, com seu conhecimento algébrico, desenvolver a demonstração.

**Técnica 9.2 e o Discurso Teórico-Tecnológico**: variância de uma constante é zero ( $Var(k) = 0$ )

Esta técnica é muito importante, pois pode despertar o entendimento de que só existe necessidade de uma medida de variação quando os valores são diferentes. Por esse motivo, vários autores dizem que a variação é o coração da Estatística.

O único livro que apresenta um exemplo desta propriedade é L5, quando inicia a explicação de variância a partir de três grupos de mesma média e variância zero, variância pequena e variância grande.

A técnica utilizada por esse livro é intuitiva, a partir de um exemplo, enquanto o livro L4 não apresenta nenhuma técnica, apenas o resultado final.

**Técnica 9.3 e o Discurso Teórico-Tecnológico** Somando-se k em cada observação, o valor da variância não se altera

Esta propriedade está apresentada no Capítulo 2 e simbolizada por $Var(X + k) = Var(X)$.

Provavelmente, o estudante resolveria o exercício criando um conjunto numérico qualquer, calculando a média e o desvio padrão. Em seguida, o estudante escolheria uma constante qualquer, somaria tal constante em cada um dos valores do conjunto e calcularia a média e o desvio padrão do novo conjunto.

Os dois livros didáticos só apresentaram o resultado sem a explicação da técnica utilizada, ficando a cargo do professor de Matemática desenvolvê-la.

**Técnica 9.4 e o Discurso Teórico-Tecnológico**: Multiplicando cada observação por k, o valor da variância será multiplicado por $k^2$

Esta técnica está apresentada no Capítulo 2 para a variância e o livro L4, o único a fazer alguma menção sobre esta propriedade, apresentou um exercício sobre a variância e um sobre o desvio padrão.

Da mesma maneira que os outros exercícios sobre as propriedades, o livro didático apresentou o resultado final sem expor a técnica para resolvê-la.

A Tabela 21 aponta os livros que utilizaram as propriedades, sejam como exercícios e/ou como exemplos.

Tabela 21 - Número de exemplos e exercícios das técnicas identificadas na tarefa Matemática 9.

| Livros | Número de | Técnica | | | |
|---|---|---|---|---|---|
| | | 9.1 | 9.2 | 9.3 | 9.4 |
| L1 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L2 | Exemplos | --- | --- | --- | --- |
| | Exercícios | 1 | --- | --- | --- |
| L3 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L4 | Exemplos | --- | --- | --- | --- |
| | Exercícios | 1 | 1 | 1 | 2 |
| L5 | Exemplos | --- | 1 | --- | --- |
| | Exercícios | --- | --- | --- | --- |
| L6 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | 1 | --- |
| L7 | Exemplos | --- | --- | --- | --- |
| | Exercícios | --- | --- | --- | --- |

### 3.4 Pontos relevantes da análise da organização praxeológica dos livros didáticos

A amostra dos livros é muito pequena e, por este motivo, não se pretende aqui fazer nenhuma generalização, mas sim levantar alguns pontos coincidentes na maioria desses livros e que, por terem sido material de apoio dos professores participantes da pesquisa, podem auxiliar na análise dos resultados.

Esses livros adotam um método tradicional de apresentar o conteúdo estatístico, embora tenham se preocupado com a relação da Estatística com o mundo real, quando utilizam gráficos e resultados de pesquisa publicados na mídia.

O grande número de gráficos não foi utilizado para a exploração da noção intuitiva de variação, a variabilidade, perdendo a oportunidade de relacionar os diferentes conceitos estatísticos.

No que tange às medidas de variação, há uma ênfase nos exemplos e exercícios sobre a tarefa de calcular, inclusive com diferentes estratégias. Quanto à interpretação, especificamente do desvio padrão, quase todos os livros explicam que é uma medida de homogeneidade do grupo e apenas dois livros utilizam a porcentagem de dados sob a distribuição normal.

Segundo Loosen, Lioen e Lacante (1985), os livros didáticos enfatizam a heterogeneidade entre as observações e não a heterogeneidade em relação à tendência central.

Neste capítulo, foi possível observar o que é priorizado pelos livros de Matemática utilizados pelos professores participantes da pesquisa, acerca do objeto variação/variabilidade. O capítulo seguinte tem o objetivo de apresentar o que é priorizado pelas pesquisas em Educação Estatística acerca do mesmo objeto.

## 4 Os Resultados das Pesquisas acerca do Ensino-aprendizagem do Conceito de Variação e Variabilidade

O objetivo deste capítulo é verificar o que já foi pesquisado sobre ensino-aprendizagem de variação e variabilidade, buscando destacar a série escolar dos participantes, as tarefas realizadas, a metodologia empregada e os resultados obtidos.

Para facilitar a relação entre os estudos já realizados e os resultados desta pesquisa, foram elaborados três blocos de trabalhos. O primeiro bloco contém os subcapítulos 4.1 até 4.3, que dizem respeito aos estudos realizados sobre o conceito de variabilidade com participantes do Ensino Fundamental e Médio, da graduação e professores de Matemática (em formação ou em atuação), respectivamente.

O segundo bloco contém o subcapítulo 4.4, que apresenta os estudos feitos com as medidas de variação.

O terceiro bloco é composto pelos subcapítulos 4.5 e 4.6, sendo que o primeiro faz uma síntese dos aspectos de raciocínio de variabilidade e variação encontrado nos estudos publicados e o segundo faz uma reflexão entre o que os autores estão denominando de aspectos do raciocínio de variação e os níveis de raciocínio de variação desenvolvidos por Garfield (2002).

### 4.1 O conceito de Variabilidade apresentado por alunos do ensino fundamental e médio

Watson e Kelly (2002) realizaram uma pesquisa na Austrália para verificar a possibilidade de iniciar o ensino de variabilidade para crianças da <u>3ª série do ensino fundamental</u> enfatizando acaso, dados e variação. Participaram da pesquisa setenta e dois alunos de três escolas primárias, ensinados durante oito semanas (dez aulas) por um especialista em Matemática para crianças. As atividades de ensino realizadas em cada aula estão descritas a seguir, de maneira sintetizada, para permitir uma inferência das questões avaliadas no pré e pós-teste, haja vista que não foram discriminadas no artigo.

Na primeira aula, foi solicitado aos alunos contar o número de chocolates em uma embalagem (semelhante ao M&Ms brasileiro). Em duplas, os alunos contaram o número de chocolates de cada caixa, classificando-os por cor e fizeram um gráfico de barras para as cores. Com isso, os alunos perceberam que havia variação das cores no conteúdo das caixas. Foi feita uma tabela para cada classe e feita a comparação por classes. Os alunos puderam perceber que existiam mais chocolates verdes e sugeriram que talvez fossem mais baratos.

Na segunda aula, o objetivo era representar os dados de maneiras diferentes e descrever o formato geral da distribuição. Foi perguntado a cada um dos alunos quantas pessoas havia em sua família, sendo que o conceito de família tinha sido definido previamente pelos próprios alunos. Foi feito um gráfico de colunas do número de pessoas da família. Em uma classe havia uma família de nove pessoas e os alunos discutiram sobre a variação no tamanho das famílias e foi introduzido o conceito de valores discrepantes. Foi pedido aos alunos para falar sobre o formato geral do gráfico. As respostas foram: “uma montanha”, “uma montanha russa” e “uma pedra grande”.

Nas próximas duas aulas, as atividades eram relacionadas à aleatoriedade. A primeira lidava com eventos equiprováveis por meio da *atividade da roleta com meio/meio* (melhor definida no estudo de Torok, 2000, apresentado a seguir) e a segunda com eventos não equiprováveis usando dois dados e somando os resultados. Os alunos analisaram os gráficos feitos para os resultados com um dado (tabela, montanha achatada) e com dois dados (uma grande montanha, tal qual o gráfico das famílias). As idéias dos alunos sobre aleatoriedade foram transcritas pelas autoras: “As coisas não são perfeitas, você sabe!” e “Se você obteve os mesmos números (com um dado), provavelmente você roubou!” (WATSON e KELLY, 2002, p. 2).

A quinta e sexta aulas lidavam com amostragem. A sétima e oitava aulas tinham como objetivo introduzir um método experimental que geraria dados e lidaria com o uso de gráfico de pontos de dois conjuntos de dados para poder fazer comparação. A atividade pedida era verificar quanto tempo cada aluno conseguia ficar apoiado em um único pé, com os olhos fechados. Eles anotaram o tempo para o pé direito e para o pé esquerdo.

As duas últimas aulas tinham como objetivo proporcionar aos alunos a criação de suas próprias investigações. A atividade era medir a distância obtida pelo lápis que foi soprado numa superfície plana. Cada sala de aula decidiu sobre a investigação que conduziria, tais como os praticantes de esporte vão soprar mais longe do que os não praticantes ou meninos vão soprar mais longe do que as meninas.

Foram aplicados um pré e um pós-teste aos alunos, que incluíam questões sobre aleatoriedade, leitura de gráficos e tabelas e três temas relacionados à variação: variação em situações aleatórias, variação nos dados e gráficos e variação em situações amostrais. As autoras compararam a soma total de pontos no pré e pós-teste, além de comparar a soma de pontos das questões por tema: aleatoriedade e dados, variação em situações aleatórias, variação nos dados e variação na amostragem. Em todas as comparações, os alunos foram melhores no pós-teste do que no pré-teste, indicando que é possível trabalhar com variabilidade com alunos de 3ª série de ensino fundamental.

Para auxiliar o entendimento de variação natural nos dados e o entendimento de erro, Lehrer e Schauble (2002) realizaram uma pesquisa nos Estados Unidos com vinte e dois alunos de 4ª série do ensino fundamental. Os autores realizaram dois estudos em que cada um tinha duração de oito semanas, com encontros diários de uma hora com os alunos e no final do semestre realizaram entrevistas individuais. No primeiro estudo, os alunos de 4ª série foram envolvidos com um conjunto de tarefas e ferramentas para ajudá-los a considerar a existência do erro e suas várias fontes, tal como apresentado na Figura 18.

**Distribuição: uma fonte para entender o erro**

Tarefa 1: Os alunos conduziram um experimento sobre o modelo de aviões com a ponta arredondada ou reta. O objetivo da atividade era verificar se as diferenças na distribuição da altura atingida pelo avião eram devidas ao acaso ou devida ao seu modelo.

Tarefa 2: Para trabalhar medidas de centro e dispersão da distribuição, foi pedido aos alunos para medir a altura do mastro da bandeira da escola usando um instrumento feito a mão. A intenção era que os alunos concordassem que o mastro da bandeira tinha uma verdadeira altura e que a medida de centro era uma estimativa desse valor.

Tarefa 3: Com uma régua os alunos mediram um lápis número 2. Foi pedido para comparar a precisão relativa da altura do mastro da bandeira e do comprimento do lápis.

Tarefa 4: Os alunos mediram novamente o mastro da bandeira, agora com um instrumento plástico.

Figura 18 - Primeira atividade desenvolvida por Lehrer e Schauble (2002, p. 2-3)

Para a Tarefa 2, os alunos encontraram medidas que variaram de 6,2 a 15,5 metros, com uma mediana de 9,7 metros. Os alunos argumentaram que os valores centrais pareciam mais verdadeiros e que o verdadeiro valor provavelmente estaria em um dos grupos que tinha uma maior número de observações.

Na Tarefa 3, os alunos calcularam as distâncias dos valores em relação à mediana e concluíram que os valores na atividade 2 eram muito mais "espalhados". O professor introduziu a noção de "variação típica" com a mediana da distribuição das diferenças em torno da mediana[24].

Na Tarefa 4 os alunos perceberam que existia muito menos variação do que na Tarefa 2 e com isso foi possível trabalhar a variação devido ao erro, que pode ser conseqüência de erro humano (movimento da mão) ou da precisão do instrumento.

Voltando à Tarefa 1, os alunos lançaram os aviões com pontas arredondadas e retas e o professor fez uma distribuição para cada tipo de avião dividida em 3 partes: abaixo da mediana menos a medida de variação, um intervalo entre a mediana menos a medida de variação e a mediana mais a medida de variação e o terceiro grupo composto dos valores acima do valor obtido da mediana mais a medida de variação. Os alunos ficaram surpresos que 86% das medidas das alturas do avião com ponta reta caíram no intervalo mais baixo da distribuição.

No segundo estudo, os mesmos alunos já estavam na 5ª série e foram envolvidos com situações em que existisse variação natural, tal como apresentado na Figura 19.

---

[24] Segundo Bussab e Moretti (2003), o desvio mediano é a mediana dos desvios (absolutos) em relação à mediana e obtido por $dam = md_{1\le j\le n}\left|x_j - md_{1\le i\le n}(x_i)\right|$. Para exemplificar, seja o conjunto numérico (0, 1, 2, 5 e 7), a mediana é 2 e os desvios absolutos em relação à mediana são, respectivamente, (2, 1, 0, 3, 5). Estes desvios em ordem crescente são (0, 1, 2, 3, 5) e, portanto, dam = 2, que é a mediana dos desvios absolutos em relação à mediana.

| |
|---|
| **Distribuição: uma fonte para entender a variação natural**<br>Os alunos acompanharam diariamente, durante 23 dias, o desenvolvimento de uma planta típica da região, cujo ciclo de vida durava 40 dias. Eles observaram a altura, a largura, o número de folhas, sementes, etc. No último dia do experimento, o professor pediu para que encontrassem a altura típica da planta nesta etapa de seu ciclo de vida. |

Figura 19 - Segunda atividade desenvolvida por Lehrer e Schauble (2002, p. 4-5)

Na tarefa da altura das plantas, os alunos puderam perceber que plantas de mesmo tempo de vida tinham alturas diferentes (variabilidade) e que também poderiam ter sido medidas de maneira incorreta (erro). Para alguns alunos ainda ficou difícil entender como um valor típico representava todos os casos, quando na realidade existiam casos diferentes. Por outro lado, parecia evidente para os alunos que havia uma altura verdadeira da planta e que tal altura poderia ser estimada pela medida de centro de um grupo de medidas.

É interessante observar que Lehrer e Schauble (2002) trabalharam com a mediana como medida de centro e a mediana das distâncias em torno da mediana como medida de dispersão. Ou seja, os autores não começaram com a média aritmética e o desvio padrão, mas sim com a mediana e o desvio absoluto mediano. Provavelmente trabalharam a média quando os alunos já estavam na quinta série. E trabalharam dois aspectos de variação: devido ao erro e devido a uma situação natural.

Ben-Zvi (2002) fez um estudo com oitenta alunos de 7ª série do ensino fundamental em Israel. Ele desenvolveu uma disciplina denominada Análise Exploratória de Dados em dez semanas, utilizando planilha eletrônica. O objetivo era inferir o grau de entendimento e disposição dos alunos para formular questões e hipóteses de pesquisa usando representação de dados, em particular, o uso de observações individuais ou globais dos dados. Após duas semanas de término da disciplina, ele solicitou que os alunos fizessem uma análise da imigração em Israel (observações reais obtidas do *Statistical Abstract of Israel* de 1995). Para tanto, os alunos (em duplas) tiveram noventa minutos no laboratório de informática para responder a quatro questões apresentadas na Figura 20.

| |
|---|
| 1) Elabore uma questão de pesquisa e uma hipótese sobre a imigração de Israel |
| 2) O número de imigrantes em 1949 era 240.000 e em 1995 era 76.000. Baseado nos dados, sugira uma questão e uma hipótese para esta informação. |
| 3) Usando o computador para apresentar gráficos e verificar tendências, o que você analisa do gráfico sobre a imigração de Israel? |
| 4) E explique o significado de oscilação da imigração. |

Figura 20 - Atividade apresentada por Ben-Zvi (2002, p. 2)

Para a primeira questão, 57,5% dos alunos responderam de maneira global (foco em padrões gerais, tais como mudanças ao longo do tempo ou tendências), enquanto os outros alunos responderam localmente (focando em um valor individual), ou usando informações do contexto (buscavam explicações e melhor entendimento do contexto). Para a segunda questão, as respostas de maneira global foram 67,5%, aumento explicado pelo autor como a maneira de colocar a questão (para ele, a segunda questão era mais direta do que a primeira).

Para responder as questões 3 e 4, os alunos fizeram gráficos de linha e barras, o que proporcionou mais respostas globais (87,5%). Cinco categorias de respostas globais foram elaboradas por Ben-Zvi (2002, p.4): (a) dividiram os dados em três períodos de imigração: alta, baixa e alta novamente; (b) dividiram os dados em quatro períodos: alta, baixa, alta e moderada; (c) indicaram variabilidade nos dados mas não especificaram um padrão, por exemplo, o número de imigrantes não é constante e não tem tendência; (d) interpretação cíclica e (e) interpretação de padrões gerais.

A interpretação cíclica era pedida especificamente na questão 4 e somente 45% dos alunos apresentaram respostas globais como, por exemplo, a oscilação da imigração indica que há altos e baixos no período, mostrando que os alunos tiveram habilidade para reconhecer e descrever ciclos com base na variabilidade dos dados.

Ben-Zvi (2002) explica que os alunos não apresentaram apenas respostas globais ou locais, mas muitos apresentaram ambas e ele sugere que duas trajetórias de desenvolvimento possam ocorrer simultaneamente: vertical (crescimento na sofisticação dentro do tipo de resposta – global ou local) e horizontal (crescendo de local para global).

O mesmo autor (BEN-ZVI, 2004) realizou um outro estudo com dois alunos de 7ª série do ensino fundamental para verificar o raciocínio sobre variabilidade quando são feitas comparações entre dois grupos num ambiente informatizado (usando Microsoft Excel). Estes dois alunos foram considerados acima da média dos outros alunos e já conheciam os gráficos de barras e setores e média aritmética.

Segundo o autor, a comparação de grupos é um tipo de problema em que o aluno não sabe, inicialmente, como lidar e o desafio pode permanecer mesmo depois de extensos períodos de ensino. Ben-Zvi (2004) identifica alguns conhecimentos necessários para realizar uma comparação entre dois grupos: entendimento de distribuição, representatividade e variabilidade nos dados. O autor também lembra de algumas dificuldades que já foram detectadas anteriormente como visão local (ao invés de uma visão global), lidar com grupos de tamanhos diferentes (que requer raciocínio proporcional) e o não uso de medidas de tendência central para representar os grupos.

> Konold et al. argumentam que a relutância dos alunos a usar medidas de tendência central para comparar dois grupos sugere que eles não têm desenvolvido a noção de que as medidas de tendência central são medidas características de um grupo, que podem ser usadas para representá-lo. (apud BEN-ZVI, 2004, p. 44).

O autor trabalhou com um conjunto de dados sobre o tamanho do sobrenome de alunos israelenses e americanos, tal como mostrado na Figura 21.

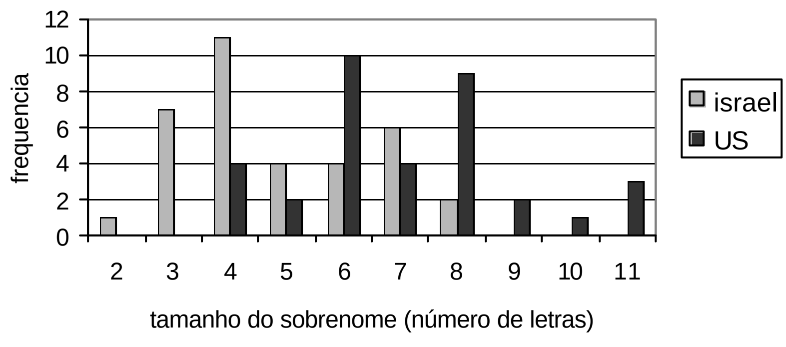



Figura 21 - Gráfico feito pelos alunos de Ben-Zvi (2004), para representar o tamanho dos nomes

Os dados foram fornecidos aos alunos em uma tabela com todas as observações de cada grupo. Essa atividade durou três sessões de noventa minutos cada. O objetivo da atividade era proporcionar oportunidades de observar, reconhecer, lidar intuitivamente com e descrever a variabilidade dentro e entre as distribuições. Foi feita uma micro-análise interpretativa dos dados e o autor encontrou sete fases de desenvolvimento do raciocínio sobre variabilidade: 1) foco em informações irrelevantes ou locais, 2) descrição informal de variabilidade no rol de dados (entender a comparação da variável em questão), 3) formular uma hipótese estatística que leva em conta a variabilidade (a maioria de); 4) explicar a variabilidade em um tabela de distribuição de freqüência (começaram observando valores extremos e depois os valores de maior freqüência); 5)uso de medidas de centro e dispersão para comparar grupos (usaram valores máximos e mínimos, média, mediana, moda e amplitude); 6) modelar variabilidade lidando com os *outliers* e 7) observação e distinção da variabilidade dentro e entre as distribuições, a partir do gráfico.

Em síntese, os resultados encontrados pelo autor sugerem que os alunos raciocinam sobre variabilidade começando a observar os valores extremos, embora ainda existisse uma confusão do que seria *outliers* (considerando que os nomes de Israel eram pequenos, os alunos achavam que um *outlier* seria um nome grande e vice-versa para os nomes americanos) para poder se encaminhar para os valores centrais.

> Uma vez que os valores extremos foram reconhecidos, os alunos encaminhavam-se para comparar as freqüências dos valores vizinhos, respectivamente o último e o primeiro valor comum da distribuição ( 8 letras e 4 letras) (...) Eles também observaram que o nome de 4 letras era a moda em Israel. Estes comentários podem representar os primeiros passos em relação ao entendimento de densidade de uma distribuição. (BEN-ZVI, 2004, p. 52).

Ben-Zvi (2004) ainda explica que, quando compararam os valores das medidas, os alunos generalizaram que todas as medidas estatísticas seriam menores para o conjunto de nomes israelenses, o que indica que eles não tinham o significado das medidas como números representativos e nem tinham clara a distinção entre as medidas de centro e dispersão (se a média é maior, isto não implica em um desvio padrão maior). E os dois alunos apresentaram a descrição final da variabilidade entre os grupos baseado na comparação das freqüências de dois subgrupos, ignorando os desvios em relação à medida de tendência central.

Quando Pfannkuch (2005 p. 85) analisou esse trabalho de Ben-Zvi (2004), questionou se o uso da tabela ajudou o pensamento dos estudantes, haja vista que, para muitos, a tabela serve apenas para organizar os dados para fazer o gráfico.

> Era a tabela de porcentagem um precursor necessário para o pensamento dos alunos com o gráfico? Os estudantes integraram a tabela de porcentagem e o gráfico para elaborar seu pensamento ou o pensamento permaneceu separado? Como o professor deveria integrar as medidas estatísticas com o raciocínio gráfico?. ( PFANNKUCH, 2005, p. 85)

São perguntas que a autora coloca e para as quais não oferece respostas. São reflexões que ela considera necessárias para trabalhos futuros. Ou seja, a autora questiona sobre uma possível maneira de relacionar os conceitos estatísticos e como isso pode colaborar no desenvolvimento do raciocínio sobre variabilidade.

Bakker (2004) realizou uma experiência com trinta alunos holandeses de 8ª série do ensino fundamental de uma escola pública que não tinham nenhum conhecimento estatístico, para verificar o processo de aprendizagem dos alunos sobre conceitos-chave para análise de dados, especialmente, variabilidade (usando a metáfora de Konold e Pollatsek, 2002, o autor explica que objetivava verificar a aprendizagem dos alunos acerca do ruído existente numa distribuição), amostragem, dados e distribuição. O autor trabalhou com duas atividades, conforme Figuras 22 e 23.

**Atividade: aumentando o número de observações numa variável**

1) Supor e representar o peso de 10 pessoas que entrarão num balão;
2) Comparar sua representação com dados reais apresentados pelo professor;
3) Predizer e representar o peso dos alunos de uma sala de oitava série e de três outras salas de oitava série.
4) Comparar com os dados reais. Descrever as diferenças utilizando palavras como maioria, *outliers*, dispersão e medidas de tendência central. O autor explica que solicitou o uso destes termos para que as respostas não fossem tão superficiais e eles começassem a utilizar as noções estatísticas em seu raciocínio.
5) Predizer e representar os pesos de todas as pessoas da cidade.

Figura 22 - Primeira atividade desenvolvida por Bakker (2004)

A primeira atividade tinha como objetivo permitir ao aluno raciocinar sobre o formato da distribuição em relação aos aspectos de amostragem. Na primeira atividade, os alunos fizeram gráficos de pontos, gráfico de barras e o gráfico semelhante ao de bastão (em que só ficava o ponto que estava no término de cada bastão) todos exemplos que já haviam sido tratados em uma lição anterior.

Na segunda atividade, o objetivo do professor era verificar se os alunos percebiam que o tamanho da amostra era muito pequeno para permitir conclusões. E os alunos usaram termos como juntos, separados ou muito separados para descrever características das observações. O autor nota que todos são predicados e não substantivos e, segundo a teoria semiótica usada por ele, variação pode se tornar um conceito disponível, quando for pronunciada e raciocinada como um substantivo. E que a passagem do predicado “os pontos estão espalhados” para “a variação é grande” é um passo importante na formação do conceito.

Depois que os alunos perceberam que amostras de tamanho dez (n = 10) eram muito pequenas e que amostras maiores são mais confiáveis, os estudantes fizeram um gráfico para predizer o peso dos vinte e sete alunos de sua série e um outro gráfico para predizer o peso de sessenta e sete alunos, estudantes de três salas de oitava série de uma outra escola e deveriam comparar com os dados reais.

O autor explica que os alunos não utilizaram os termos estatísticos de maneira precisa e exemplifica: média significava alguma medida de tendência central (mediana ou moda) ou o grupo típico do meio; dispersão significava o quão espalhado os dados encontravam-se e amostra significava somente um conjunto de pessoas e não necessariamente os dados como sendo representativos de uma população.

Embora ainda existisse confusão nos conceitos, os alunos começaram a utilizar os substantivos em vez dos predicados. No caso de variação, quando o aluno diz que a variação é grande (em vez de dizer que os dados estão espalhados), ele tende a compreender a variação como “uma entidade que pode ter uma característica particular que pode ser medida (por exemplo: a amplitude, o intervalo interquartílico e o desvio padrão)” (BAKKER, 2004, p. 73).

Quando foi pedido aos alunos para elaborar uma representação do peso de todas as pessoas da cidade, o autor queria verificar se os alunos fariam um formato contínuo.

> A conjectura era que a transição de uma pluralidade discreta dos valores dos dados para uma entidade contínua de uma distribuição era importante para promover uma noção de distribuição como uma entidade

> com a qual os alunos poderiam modelar os dados e descrever propriedades gerais do conjunto de dados. (BAKKER, 2004, p.73).

Os alunos produziram gráficos semelhantes à curva de sino (olhando o grupo que está no meio), pirâmide (provavelmente inspirado no gráfico de linha) e semicírculo (provavelmente partindo do gráfico de pontos), cujos termos foram usados pelos próprios alunos. Todos os gráficos eram aproximadamente simétricos e o autor salienta que o gráfico de peso (com os dados reais) era assimétrico à direita.

Até este momento o autor tinha tratado de medidas de dispersão e centro como aspectos essenciais de distribuição, mas salienta que assimetria também é uma característica importante de uma distribuição e, por esse motivo, foi alvo da atividade seguinte de Bakker, como apresentado na Figura 23.

| |
|---|
| O autor colocou na lousa cinco possíveis gráficos do peso das pessoas (curva de sino, pirâmide e semi-círculo - os três que eles haviam feito - e dois assimétricos, para direita e para esquerda) e pediu para que os alunos identificassem quais não se assemelhariam à distribuição dos pesos. |

Figura 23 - Segunda atividade desenvolvida por Bakker (2004).

Os estudantes identificaram todos os gráficos como não representativos dos pesos de todas as pessoas, exceto para o formato da curva normal. Na discussão, os alunos foram percebendo que poderia haver pessoas com mais peso, o que justificaria não ser o formato da distribuição normal. Usaram palavras como *outliers* e média (embora confundindo-a com a moda) para justificar a mudança de escolha.

Nesse estudo de Bakker (2004), destaca-se a relação feita entre a estrutura da frase elaborada pelo aluno e a construção do conceito de variação. Em relação aos objetivos propostos nas atividades, Pfannkuch (2005, p. 85) fez os seguintes questionamentos:

> os estudantes perceberam que havia mais variação nas amostras pequenas? Que padrões os estudantes estavam imaginando na variabilidade de amostras pequenas? Como eles raciocinaram entre distribuição da amostra e a distribuição da população?

Ficou implícito no trabalho de Bakker (2004) que os alunos compreenderam que existia mais variação nas amostras pequenas, a partir do avanço na estrutura das frases elaboradas, mas isso não foi explicitamente colocado.

Torok (2000) fez um estudo com alunos australianos acerca da variabilidade em situações de acaso. Ele estava especialmente interessado em verificar a aprendizagem sobre: reconhecimento da existência de variação, reconhecimento e nomeação de algumas fontes de variação e o uso de métodos formais e informais para descrever variação, tais como gráficos e medidas de variação. Foram sujeitos da pesquisa alunos de <u>uma sala de 7ª. Série e duas salas de 8ª. Série</u>, em que na 7ª série foi enfatizada a descrição de espaços amostrais de eventos aleatórios e nas outras duas classes houve foco na coleta e descrição de dados. O ensino em cada sala foi feito em sete sessões de quarenta minutos cada.

**Atividade 1:** *Roleta* dividida ao meio. Uma metade em branco e outra metade em preto.
1) Se você rodar uma vez, qual é a chance de cair no lado preto?
2) Se você rodar 50 vezes, quantas vezes você espera que caia no lado preto? Por quê você pensa isto?
3) Se você fosse fazer isto novamente, você esperaria obter o mesmo número de vezes? Por quê você pensa isto?
4) Com qual número você ficaria surpreso?
5) Suponha que você rodasse 50 vezes a roleta e repetisse este procedimento mais cinco vezes. Qual das seqüências abaixo você acha que é mais provável acontecer?
10, 0, 1, 25, 30, 49
10, 20, 30, 40, 50, 60
25, 26, 24, 30, 20, 28
6) O que a palavra variação quer dizer para você? Você poderia usá-la em uma frase?

Figura 24 - Atividade proposta por Torok (2000).

Poucos alunos deram como resposta 25 ou aproximadamente 25 para a resposta da segunda questão e, dois desses exemplos são apresentados: “porque você tem 50% de chance disto ocorrer” (aluno da 7ª. Série) ou “porque metade da roleta é preta então metade das vezes vai cair no preto” (aluno da 8ª. Série).

Para a quinta questão, alguns alunos não sugeriram nenhum agrupamento dos dados em torno de um valor esperado (10, 0, 1, 25, 30, 49) enquanto que outros alunos fizeram uma lista de seis números ao redor do seu valor esperado (25, 26, 24, 30, 20, 28) ou ainda responderam que esperariam um número em torno de 25.

Poucos alunos conseguiram elaborar uma frase com a palavra variação. “Variação assemelha-se a uma pequena mudança” ou “Há uma variação naqueles números” ou ainda “Significa uma variedade para mim. Eu gosto de uma variedade de cores no meu quarto” (TOROK, 2000, p. 27).

Os alunos fizeram o experimento com 50 rodadas da roleta, colorida nas proporções meio preto e meio branco, um quarto e três quartos e colocaram seus resultados em gráficos de pontos. Na discussão dos gráficos, foi enfatizado que nem todos os resultados estavam localizados nos valores esperados, apesar de as medianas estarem muito próximas do que se esperava. Foi percebido, também, que as distribuições tinham um formato de colina e eram aproximadamente simétricas em torno do valor esperado. Para a roleta de meio/meio, era óbvio que os dados estavam agrupados em torno do valor do meio entre zero e cinqüenta. O professor apresentou aos alunos três outros gráficos simulando dados coletados por outras três classes e perguntou aos alunos qual (ou quais) não tinha sido realmente coletado (um gráfico perfeitamente simétrico em torno do número 25, um gráfico com valores espalhados de 5 até 50 e um gráfico com os pontos concentrados entre 20 e 30). Os estudantes disseram que somente o último gráfico tinha sido realmente coletado, pois o primeiro está muito organizado e perfeito e o segundo porque está muito espalhado e isto é improvável.

A atividade de ensino foi concluída com uma avaliação sobre a tarefa das balas, conforme Figura 25.

Suponha que você tem um pote com 100 balas, em que 50 são vermelhas, 20 amarelas e 30 verdes. As balas foram misturadas.

A) Você vai retirar 10 balas.
1) Quantas balas vermelhas você espera obter?
2) Quantas balas vermelhas surpreenderiam você?Por que você pensa isto?

B) Suponha que você pegue, de uma vez só, 10 balas, conta o número de vermelhas e volta todas as balas ao pote, misturando-as. Você repete este procedimento cinco vezes, sempre voltando as balas ao pote e misturando-as.

O que você acha mais provável de ocorrer para os números de balas vermelhas anotadas? Por que você acha isso?

C) Olhe essas possibilidades que alguns alunos anotaram para os números que eles pensaram prováveis.
a) 5, 9, 7, 6, 8, 7
b) 3, 7, 5, 8, 5, 4
c) 5, 5, 5, 5, 5, 5
d) 2, 3, 4, 3, 4, 4
e) 7, 7, 7, 7, 7, 7
f) 3, 0, 9, 2, 8, 5
g) 10, 10, 10, 10, 10, 10

Qual dessas seqüências você acha que melhor descreve o que aconteceu? Qual descreve pior o que poderia ter acontecido? Por que você pensa isto?

D) Suponha que 6 alunos fizeram o experimento. De que número até que número você acha que saiu? Por que você acha isto?

E) Agora, faça você a experiência e comente.

Figura 25 - Atividade usada para avaliar o conhecimento dos alunos no estudo de Torok (2000).

Foi pedido aos alunos que discutissem em duplas alguma alteração para o problema original e era esperado que fizessem alterações no tamanho da amostra, na proporção populacional, no tamanho da população ou no número de amostras tomadas. Fizeram a atividade usando o *software SimLollies*[25]. Todos os alunos coletaram resultados que mostravam variação no número de vermelhos para cada punhado, porém, quando foram questionados se obteriam resultados iguais ou diferentes se coletassem mais dados, alguns estudantes disseram que

[25] Um software que simula o problema das balas.

obteriam números idênticos. Ou seja, os alunos começavam a entender o conceito de variação numa situação aleatória, mas algumas afirmações ainda negavam a existência de variação.

Segundo o autor, a despeito disso, o experimento ajudou os alunos a desenvolver o entendimento de situações que contêm variação e reconhecer a ausência de variação onde geralmente é esperada (por exemplo, um aluno investigou o efeito do aumento do tamanho populacional de 50/20/30 para 500/200/300 no problema das balas e ficou surpreso com o resultado, pois esperava que a proporção de vermelhos fosse bem maior).

Reading e Shaughnessy (2004) fizeram um estudo com doze crianças, sendo uma de quarta série, duas de quinta série, uma de sexta série, três da oitava série e três do terceiro ano do ensino médio. O objetivo dos autores era compreender o raciocínio de variação em situação de acaso. A atividade aplicada era semelhante à utilizada por Torok (2000) (apresentada na Figura 25) e as adaptações bem como os resultados obtidos estão apresentados na Tabela 22.

Tabela 22 - Resultados obtidos por Reading e Shaughnessy (2004).

| Questão | Síntese da questão analisada | Tipos de respostas apresentadas | | |
|---|---|---|---|---|
| | | Causa de variação | Descrição de variação | Ambas respostas |
| Questão A | A sua sugestão de vermelhas aconteceria sempre? | 10 | 1 | 1 |
| Questão B | Lista de número de vermelhas em seis seleções de 10 balas | 5 | 1 | 3 |
| Questão C | Escolha de uma possível lista de resultados | 3 | 8 | 1 |
| Questão D | Amplitude de valores | 5 | 3 | 1 |

Os autores desenvolveram uma análise hierárquica para descrever o raciocínio em relação à variação segundo duas perspectivas: como os alunos descrevem a variação (resultado apresentado) e como eles atribuem causa para a variação (argumento utilizado para explicar o resultado apresentado).

A análise hierárquica desenvolvida pelos autores para a descrição de variação foi dividida em quatro níveis: D1: concentra-se em valores no meio do conjunto ou valores extremos; D2: concentra-se em ambos valores internos e extremos; D3: discute desvios de um valor e D4: discute desvios de um valor central. As respostas referentes à descrição de variação diziam respeito a como os números estavam dispersos e o que estava acontecendo com os números contidos em um intervalo.

> Quando estes dois aspectos da descrição de variação se apresentam juntos, os desvios começam a se tornar uma questão de destaque; e quando estes desvios estão em relação a um valor específico, geralmente um valor central, ele [o desvio] se tornará naturalmente o foco da descrição de uma distribuição. (READING e SHAUGHNESSY, 2004, p. 220)

Para as respostas causais de variação, a análise hierárquica apresentada por Reading e Shaughnessy (2004) está sintetizada no Quadro 8 .

Quadro 8 - Análise hierárquica das respostas causais de variação elaborada por Reading e Shaughnessy (2004, p. 217-220)

| Categoria | Exemplos de respostas consideradas na categoria |
|---|---|
| C1 – identificação de causas irrelevantes de variação | Explicações sobre onde e como foram dispostas as balas no pote (causas físicas). |
| C2 – discute freqüência de cores como causa de variação | Como a proporção de vermelhos era maior, os alunos conseguiam verificar que havia mais balas desta cor, mas não apresentavam a proporção. |
| C3 – discute proporção de cores como causa de variação | Utilizam a proporção, mas com alguns equívocos. Um dos alunos respondeu que deveria ter metade ou mais de vermelhos. |
| C4 – discute probabilidade baseada em proporções | Os alunos utilizaram o raciocínio proporcional combinado com a probabilidade de obter a proporção da população. Um aluno (do 3º ano do ensino médio) apresentou uma ordem dos resultados mais prováveis. Para a proporção populacional de 70% de vermelhos, ele disse: “7 é o mais provável e o 6 e o 8 são os próximos prováveis”. |

E a análise hierárquica das causas de variação descreve um desenvolvimento da percepção da fonte de variação que foi descrita. Segundo Reading e Shaughnessy (2004, p. 221), o tipo de questão apresentada aos alunos pode “requerer uma ou ambas respostas de variação” (descrição ou causa). Como pode ser observada na Tabela 22, a primeira questão estimulou respostas do tipo causa de variação enquanto a terceira questão estimulou respostas descritivas de variação.

Reading (2004) realizou uma pesquisa com alunos australianos de uma sala de 6ª série, de uma sala de 8ª série e uma sala de 2ª. série do ensino médio, com o objetivo de dar continuidade a seu estudo sobre a análise hierárquica da descrição de variação apresentada por alunos. A atividade utilizada para esse estudo está apresentada na Figura 26.

| |
|---|
| **Atividade denominada Tempo**: Foi fornecida aos alunos uma tabela contendo os milímetros diários de chuva e as temperaturas máximas e mínimas de cada dia, ambas de um período de dois meses dos últimos três anos. Os alunos precisariam escolher o mês para que fosse instituída uma nova festividade na cidade, que aconteceria ao ar livre, realizando as seguintes tarefas: 1: descrever o tempo de sua amostra; 2: comparar sua descrição com a descrição feita pelos colegas e 3: Decidir o mês. |

Figura 26 - Atividade desenvolvida por Reading (2004)

Entre a descrição da temperatura e da chuva, os alunos tiveram um episódio de ensino que trabalhava o diagrama de ramo-e-folhas, valores máximo e mínimo, medidas de tendência central e amplitude para a 6ª série e, para a 8ª série foi introduzido o gráfico de *boxplot* a partir do diagrama de ramo-e-folhas que já lhes era familiar. Para o segundo ano do ensino médio foi desenvolvida uma unidade completa de trabalho estatístico.

Reading (2004) utilizou a análise hierárquica da descrição de variação desenvolvida por Reading e Shaughnessy (2004) para analisar o raciocínio dos alunos e a maioria das respostas ficou nos níveis D1 e D2.

Estes dois níveis de descrição de variação foram re-divididos em respostas qualitativas e quantitativas. Segundo Reading (2004), as respostas qualitativas são menos sofisticadas estatisticamente, mas refletem a observação e o reconhecimento da variação, sem que os alunos saibam calcular uma medida.

Cada resposta qualitativa e quantitativa foi dividida em Uniestrutural, quando o participante da pesquisa apresenta a resposta focando em apenas um elemento; Multiestrutural, quando o participante da pesquisa apresenta a resposta focando em vários elementos não relacionados e Relacional, quando foca em vários elementos cuja relação é identificada.

O Quadro 9 apresenta o modelo desenvolvido por Reading (2004) para analisar as respostas de variação apresentada pelos participantes da pesquisa.

Quadro 9 - Modelo de desenvolvimento cognitivo proposto por Reading (2004, p. 97)

| Resposta Qualitativa | Resposta Quantitativa | Relação com a análise hierárquica |
|---|---|---|
| **Uniestrutural:** frases relacionadas à magnitude (exemplos: tempo razoavelmente constante, muito imprevisível) e relacionadas com à posição dos elementos em relação a outros elementos do conjunto (uniformemente distribuído, espalhado) | **Uniestrutural**: utilizaram valores máximos e mínimos ou usaram a amplitude, ou ainda os valores internos, em que agrupam os dados em pequenos conjuntos de dias. | Como as respostas de D1, valores extremos ou valores internos. |
| **Multiestrutural:** surgiram frases relacionadas ao limite (muito frio, não muita chuva) e relacionadas seqüencialmente (há muitos dias secos seguidos de dois dias de chuva e então muitos dias secos) | **Multiestrutural**: Utilizaram duas medidas ao mesmo tempo. Os exemplos apresentados pela autora trabalharam com as temperaturas máxima e mínima. | Como as respostas de D2 |
| **Relacional**: combinaram as categorias já discutidas anteriormente (mês de Janeiro tem chuva regular, com períodos de 3 a 5 dias secos) | **Relacional**: Trabalhava com a diferença entre os valores (um dia com -0,9 e o dia seguinte com 7 graus Celsius) | Para as respostas qualitativas, combinaram as categorias D1 e D2. Para as respostas quantitativas, que sugeriam noção de desvio, a autora relacionou-a com D3. |

Reading (2004, p. 100) concluiu que o uso de contexto real, apesar de considerado mais significativo, impede os alunos de reconhecer uma oportunidade de utilizar as habilidades adquiridas em sala de aula. “Poucos alunos da 6ª série fizeram um gráfico para descrever os dados e somente 3 alunos usaram seus gráficos para explicar. Nem o episódio de ensino resultou no aumento do uso de gráficos.”

Isso indica que a aprendizagem não foi significativa e que o conceito não foi construído. Para Brousseau (1997), o primeiro contato com um conceito deve ser contextualizado. Na fase de institucionalização o professor descontextualiza e nas re-aplicações, o aluno deve recontextualizar para resolver o problema proposto, exatamente como todos os estudos mostrados até aqui. Possivelmente a elaboração da atividade de Reading (2004) precisaria ser repensada, como salienta Pfannkuch (2005), que questionou se as características da tabela de dados podem ter induzido os estudantes a não mudar para uma outra representação, considerando-se que só três alunos fizeram gráficos.

É possível notar que todos os estudos apresentados lidavam com variabilidade, ou seja, a capacidade de variar, e não especificamente com a medida de variação. Exploraram gráficos, mas sem relacioná-los com as medidas de variação, que foram apenas duas: a amplitude (no estudo de Ben-Zvi, 2004 e de Reading, 2004) e o desvio absoluto mediano (no estudo de Lehrer e Schauble, 2002).

### 4.2 O conceito de Variabilidade apresentado por alunos de graduação.

Com alunos de graduação, estão apresentados apenas os resultados da tese de Meletiou (2000), e seus trabalhos posteriores, provenientes dela (Meletiou e Lee, 2002 e Meletiou-Mavrotheris e Lee, 2002).

A tese da Meletiou (2000) tinha como objetivo estabelecer uma ligação entre o entendimento intuitivo e o entendimento formal de variação a partir de uma estrutura de ensino que ela estava testando, em que a variação era a essência da disciplina. A disciplina teve duração de cinco semanas, com quatro encontros semanais de duas horas cada um. Todo o experimento foi conduzido por um estatístico e a autora foi observadora não participante.

Ela trabalhou com <u>trinta e três alunos (dezenove homens e quatorze mulheres) de graduação</u> (EUA), com pouco conhecimento matemático. Embora tenha trabalhado com trinta e três alunos, a parte qualitativa foi desenvolvida com apenas oito alunos, que representavam as características da classe toda.

No primeiro dia foi aplicado um questionário para investigar o entendimento intuitivo dos alunos e direcionar a etapa de ensino. Logo após o questionário, a autora fez entrevistas individuais para entender as respostas do questionário. Os exercícios utilizados por Meletiou (2000) no pré-teste (primeiro dia) e os resultados obtidos estão sintetizados a seguir.

1) Baseado em sua experiência, o que significa variabilidade para você? Dê uma explicação verbal ou um exemplo.

Figura 27 - Questão 1 do pré-teste aplicado por Meletiou (2000).

Muitos alunos relacionaram variabilidade com variedade, com múltiplos valores, como medidas de coisas diferentes, variação de um mínimo para um máximo ou alguma coisa que não era constante.

2) Em cada caso abaixo, é melhor que a variabilidade seja alta ou baixa?

a) idade das árvores numa floresta nacional;

b) Diâmetro de novos pneus saídos de uma linha de produção;

c) Chuva diária;

d) Peso de um pacote de cereal.

Figura 28 - Questão 2 do pré-teste aplicado por Meletiou (2000).

Segundo a autora, os alunos reconheciam que a variabilidade baixa poderia ser boa ou ruim, dependendo do contexto. Por exemplo, eles imaginaram que seria ruim para o desempenho do carro se a variabilidade do diâmetro dos pneus fosse grande.

3) Dois alunos que estavam cursando uma disciplina de Estatística tiveram as seguintes notas (numa escala até 100)

Estudante A – 60, 90, 80, 60, 80

Estudante B - 40, 100, 100, 40, 90

Se você tivesse que fazer um teste estatístico, quem você escolheria para trabalhar em dupla?

Figura 29 - Questão 3 do pré-teste aplicado por Meletiou (2000).

Nesta questão, 50% dos alunos escolheram o estudante A e justificaram dizendo que apesar de ambos terem a mesma média, o estudante A é mais consistente. 33% escolheram o estudante B, pois como ele tirou algumas notas máximas, eles acreditavam no seu potencial e 15% dos alunos disseram que tanto faz, pois ambos tinham a mesma média.

7) Qual das seguintes distribuições tem mais variabilidade?

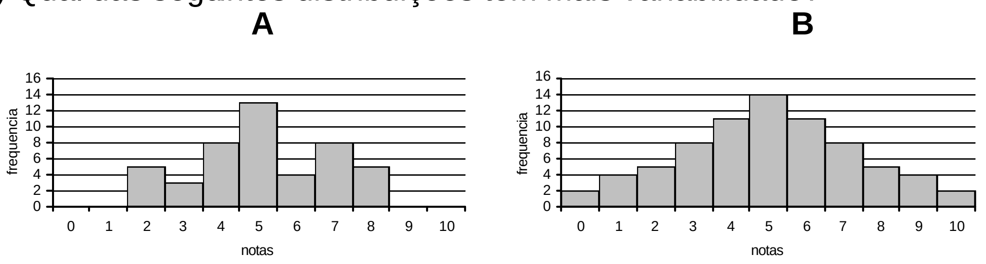



Figura 30 - Questão 7 do pré-teste aplicado por Meletiou (2000).

Nesta questão, 22 alunos consideraram que a Distribuição B tinha mais variação e os alunos que escolheram a Distribuição A explicaram que era devido à irregularidade das colunas.

A questão 10 do pré-teste também observava a variação por meio de histogramas. O objetivo dessa questão era verificar como os alunos iriam relacionar as características de uma distribuição com seu formato.

10) Considere a seguinte lista de variáveis:

a) idade de morte de uma amostra de 34 pessoas;

b) o último dígito na securidade social de 40 pessoas;

c) notas num teste de estatística razoavelmente fácil;

d) peso de um grupo de adutos;

e) número de medalhas dos países vencedores da Olimpíada de Inverno de 1992.

Use o seu conhecimento da variável (pergunte a você mesmo se a distribuição é aproximadamente simétrica ou não) para relacionar cada histograma com cada variável.

Figura 31 - Questão 10 do pré-teste aplicado por Meletiou (2000).

A autora apresentou cinco histogramas e somente três alunos escolheram todos os gráficos corretamente. Muitos alunos pensaram que cada barra do histograma representava uma observação, o que não permitiu avaliarem a densidade de freqüência. Nesse sentido, Loosen, Lioen e Lacante (1985) fizeram um estudo apresentando um gráfico em que cada barra era uma observação da variável e concluíram que não é a melhor estratégia para iniciar o ensino de variabilidade.

Há que se destacar que o histograma é um tipo de gráfico que gera muita dificuldade. No estudo de Meletiou (2000), um dos alunos que relacionou corretamente cada variável com seu gráfico, interpretou o histograma como o gráfico de barras. Meletiou e Lee (2002) argumentam que o histograma é gráfico que representa muita dificuldade para os alunos e que isto prejudica o desenvolvimento da intuição de variação.

Relacionada à representatividade amostral, Meletiou (2000) fez a pergunta que está apresentada na Figura 32.

Todos alunos consideraram que as Pesquisas 1 e 3 apresentavam vieses e poucos alunos disseram que não era aleatória. Sobre a Pesquisa 4, a maioria dos alunos disse que não representava a escola toda e quatro alunos consideraram que era aleatória, pois todos os alunos da escola tinham a mesma oportunidade. Já na Pesquisa 5, os alunos detectaram mais problemas do que na Pesquisa 4, pois só iriam entrevistar quem estava interessado em ganhar um vídeo game.

A maioria dos alunos preferiu a Pesquisa 2, que usava uma amostra aleatória e a pesquisa 6 que trabalhava com amostra aleatória estratificada. Alguns alunos justificaram esta última escolha, pois além de ser aleatória tem boa representatividade da população, e dois alunos usaram a palavra variabilidade com conotação de amostra representativa “boa variabilidade dos alunos – ambos os sexos e todas as séries”

6) Estudantes de ensino fundamental estavam querendo angariar dinheiro para uma viagem e resolveram rifar um vídeo game. Seis alunos fizeram uma pesquisa para verificar o número de alunos que comprariam a rifa. Cada um entrevistou 60 alunos usando metodologias diferentes, conforme descrito abaixo.

Pesquisa 1: Tom perguntou para 60 amigos (75% sim e 25% não)

Pesquisa 2: Shannon pegou o nome dos 600 alunos da escola, colocou-os num chapéu e pegou 60 deles (35% sim e 65% não)

Pesquisa 3: John perguntou a 60 alunos num encontro no Games Club, onde tinham encontros semanais para jogar diferentes games computadorizados. (90% sim e 10% não).

Pesquisa 4: Ann aplicou um questionário para todos os alunos da escola e usou os primeiros 60 que responderam (50% sim, 50% não)

Pesquisa 5: Claire colocou uma cabine de pesquisa na lanchonete com a seguinte mensagem: Ganhe um vídeo game. Quem quisesse respondia o questionário. Assim que ela obteve 60 respostas, ela parou. (100% sim)

Pesquisa 6: Kyle queria o mesmo número de meninos e meninas e alguns alunos de cada série. Então, ele perguntou para 5 meninos e 5 meninas de cada série, num total de 60 entrevistados (30% sim, 70% não).

Questões:

a) Para cada uma das entrevistas realizadas, responda: o que você acha sobre a maneira que foi conduzida? Você acha que foi feita de maneira adequada? Seus resultados proporcionam uma boa idéia de quantos alunos na escola comprariam um número da rifa do vídeo game? Explique o porquê.

b) Se você tivesse que escolher uma das seis maneiras de fazer a pesquisa, qual você escolheria? Explique sua resposta.

c) Qual você acha que é a melhor estimativa para a porcentagem de alunos que comprará a rifa?

Figura 32 - Questão 6 do pré-teste aplicado por Meletiou (2000).

Meletiou (2000) explica que observou, no inicio da disciplina, um entendimento informal de questões relacionadas à variação amostral. Na questão 8, apresentada na Figura 33, somente 20% dos alunos pensaram em vermelho ou preto como igualmente prováveis. 67% dos alunos esperavam que saísse preto para balancear a distribuição.

8) Uma roleta tem 18 números pretos (B) e 18 vermelhos (R). A probabilidade de uma bola cair num número vermelho é a mesma que cair num número preto. Um jogador observa a bola cair seis vezes no número vermelho, na seqüência RRRRRR. Em que cor você acha que a bola cairá na próxima rodada? Por quê?

Figura 33 - Questão 8 do pré-teste aplicado por Meletiou (2000).

Como já discutido em muitos trabalhos sobre média aritmética, os alunos esperam que os próximos valores compensem os valores já observados e desprezam o conhecimento estatístico já adquirido em favor de uma noção do senso comum (equilíbrio)[26].

A questão 4, apresentada na Figura 34, tinha como objetivo verificar como os alunos lidavam com as idéias de variabilidade e representatividade amostral.

4) Suponha que você levou seu sobrinho ao desfile de Páscoa. O coelhinho da Páscoa distribuiu pacotes de confetes coloridos para todos os alunos. Cada pacote tinha seis confetes. Para fazer os pacotes, o coelhinho da Páscoa pegou dois milhões de confetes verdes e um milhão de vermelhos, colocou tudo num grande pote e misturou tudo e fez muitos pacotes de seis confetes, sempre pegando um punhado de confetes e colocando nos pacotes até que todos os pacotes tivessem sido preenchidos.

a) quando você chegou em casa, você abriu o seu pacote. Quantos confetes verdes você acha que pode ter em seu pacote? Você pode explicar como pensou?

b) Você acha que todos os alunos pegaram *n* verdes, onde *n* é o número de confetes verdes que você pegou? Você pode explicar por quê?

c) Se você pudesse olhar os pacotes de 100 alunos, quantos alunos você acha que pegou *n* confetes verdes?

d) Lembre-se que o desfile de Páscoa começou com dois milhões de confetes verdes e um milhão de vermelhos. Ele usou até o fim uma cor antes de outra quando estava preenchendo os pacotes ou ambas as cores ficaram até próximo do fim? Por quê?

Figura 34 - Questão 4 do pré-teste aplicado por Meletiou (2000)

[26] Considere que sair vermelho ( R ) é 1 e sair preto ( B ) é zero. A proporção de vermelho em n jogadas é obtida pela soma de todos os vermelhos que saíram em n jogadas, dividida pelo número de jogadas, cujo cálculo é o mesmo da média aritmética.

Todos os alunos responderam quatro verdes e duas vermelhas como estimativa e todos imaginaram que nem todos alunos tinham obtido quatro verdes porque havia variabilidade ("eles entenderam que seleção aleatória lida com variação" – MELETIOU, 2000, p. 140).

A autora comenta que os alunos entenderam intuitivamente que probabilidade é o limite da freqüência relativa, que se sustenta de maneira aproximada para os dados reais. Este é o enfoque freqüentista do conceito de probabilidade. No entanto, Coutinho (2001) explica que, intuitivamente, os alunos fazem o amálgama entre freqüência e probabilidade. Tal amálgama pode ser um obstáculo reforçado pelas opções didáticas do professor, caso não se mostre efetivamente que, na realidade, esses dois conceitos são distintos, referindo-se inclusive a campos distintos (Teoria das Probabilidades e Estatística Descritiva).

Outra questão sobre variabilidade e representatividade amostral está apresentada na Figura 35. Nessa questão, somente 35% dos alunos pensaram, corretamente, que alguém deveria esperar aproximadamente um número igual de homens e de mulheres, pois quem foi selecionado até aqui não afeta quem será selecionado na seqüência do experimento. Como haviam sido selecionadas mais mulheres do que homens, 33% dos alunos argumentaram que esperaram o oposto acontecer e 16% tentaram achar causas por trás da diferença que (dado o pequeno número de pessoas entrevistadas até aqui) poderia ser facilmente explicada pela variação ao acaso enquanto o resto, empregando a lei dos grandes números, pensaram que a tendência de selecionar mais mulheres do que homens deveria continuar.

9) Circule a melhor resposta para o seguinte problema:

Numa escola de ensino médio da redondeza, metade dos alunos são mulheres e metade são homens. Um trabalhador de uma organização estudantil quer entrevistar alunos sobre as recentes mudanças no fundo governamental de ajuda financeira. O trabalhador quer obter uma boa representação dos estudantes e vai para muitas diferentes áreas do campus. Três ou quatro estudantes foram entrevistados em cada lugar visitado. Os últimos 20 estudantes entrevistados, treze foram mulheres e sete homens. Agora, você não sabe que hora do dia é, que parte do campus o trabalhador já foi ou onde o trabalhador está indo. Os próximos 20 estudantes que o trabalhador vai entrevistar, você acha que serão mais homens ou mulheres?

a) o trabalhador parece entrevistar mais mulheres do que homens. Poderia haver muitas razões para isto. Talvez mulheres são mais disponíveis para falar sobre suas opiniões. Ou, talvez o trabalhador vai para áreas do campus onde há mais mulheres que homens. Desta forma, é provável entrevistar mais mulheres do que homens nos próximos 20 estudantes.
b) Como metade dos alunos no campus são homens e metade são mulheres, você esperaria uma divisão 50/50 entre o número de homens e mulheres que o trabalhador entrevistou. Como houve mais mulheres do que homens, eu espero o oposto. Nos próximos 20 entrevistados, haverá mais homens do que mulheres para que as coisas comecem a se equilibrar.
c) Metade dos alunos são homens e metade são mulheres. Isto quer dizer que tem uma chance de 50/50 de entrevistar um homem ou uma mulher. Não deveria interessar quantos homens ou mulheres o trabalhador entrevistou. Nos próximos 20 alunos, aproximadamente metade seriam homens e metade mulheres.
d) Até aqui, a tendência parece ser mais mulheres a serem entrevistadas do que homens. Nos próximos 20 estudantes eu esperaria a mesma coisa acontecer. O trabalhador provavelmente entrevistará mais mulheres do que homens.

Segundo Meletiou (2000), a tendência de subestimar o papel da variação devido ao acaso é mais evidente em contextos do mundo real. Apesar de os estudantes parecerem atentos aos perigos envolvidos quando tomavam decisões baseadas em pequenas amostras, no momento em que eram solicitados a fazer seus próprios julgamentos acerca dos dados, eles freqüentemente ignoravam esses perigos e, exagerando na confiabilidade das informações providenciadas, não hesitavam em usar pequenas amostras como base para inferência. As respostas dos oito alunos entrevistados na Questão 5 (que está apresentada na Figura 36) são indicativas disso.

5) Em média, há 600 mortes por ano na cidade devido a acidentes de trânsito. Uma pessoa observou o seguinte: Em Fevereiro, o número de mortes na semana 1 foi 3, na semana 2 foi 12, na semana 3 foi 21, na semana 4 foi 14 e em Março, na semana 5 foi 2. Considerando que em nenhuma dessas semanas tinham um feriado, suponha que a manchete de um jornal dizia que a semana 3 tinha sido desastrosa e que a razão era a velocidade. A semana 4 foi descrita como evidência que dirigir na cidade está piorando. Ao final da semana 5 a polícia se vangloriou pela baixa razão das mortes – suas patrulhas tinham tido sucesso. O que você diria a essa pessoa?

Figura 36 - Questão 5 do pré-teste aplicado por Meletiou (2000)

Quase metade dos alunos apresentou resposta determinística, tentando encontrar as causas para a queda na razão das mortes. Muitos alunos disseram que o número de semanas observadas era muito pequeno e que deveria esperar para ver os resultados das próximas semanas.

Meletiou (2000) explica que os alunos foram mais dispostos a reconhecer o papel da variação devido ao acaso na questão sobre lançar 50 vezes uma moeda e obter 27 caras e lançar novamente depois de dois dias e obter 30 caras (questão 3 da entrevista). A mesma aluna que tinha dito que deveria haver uma razão para diminuir o número de acidentes disse que os dois resultados não eram suspeitos, pois há variação nos resultados. De acordo com Coutinho (2001), os alunos associam o acaso aos jogos de azar de forma espontânea, o que pode ser explicado pela própria história da probabilidade, mas que, em atividades de ensino, outras situações devem ser apresentadas aos alunos, visando à construção do significado desse conceito.

Para investigar o efeito do tamanho da amostra na variação, foram feitas três outras questões na entrevista, apresentadas nas duas figuras seguintes.

| |
|---|
| 1) Todo ano na Nova Zelândia aproximadamente 7 crianças nascem com um defeito num membro. No último ano, as crianças nascidas com essa anormalidade foram localizadas no mapa da Nova Zelândia, conforme desenho (o mapa mostra cinco regiões cujas freqüências foram 0, 2, 2, 3, 0, numa ordem de região inferior para superior respectivamente)<br>O que você acha? (Na Nova Zelândia é conhecimento comum que um terço da população mora na região superior e um sexto em cada uma das outras regiões) |

Figura 37 - Questão 1 da entrevista realizada por Meletiou (2000).

De acordo com Meletiou (2000), os autores da questão (Pfannkuch e Brown) julgaram pobre o entendimento de variação dos alunos em pequenas amostras neste contexto. Enquanto uma análise, combinando pensamento probabilístico e determinístico teria sido mais apropriada, todos estudantes que eles entrevistaram deram explicações determinísticas, e foi somente depois de repetidas sondagens que alguns sugeriram a necessidade de mais dados. Os resultados de Meletiou (2000) foram semelhantes aos de Pfannkuch e Brown, em que um aluno relatou que não queria viver no meio da Nova Zelândia e outra aluna estava convencida que deveria haver um fator externo causando a diferença e explicou "há sempre uma chance de alguma coisa acontecer, mas 3 e 0 no outro...deve haver uma razão para isto" (MELETIOU, 2000, p. 145).

Quando foi perguntado aos alunos o que eles achavam da possibilidade de obter o resultado {3,3,3,4,4,5,5} quando jogado um dado honesto 7 vezes (que era a segunda questão da entrevista), nenhum estudante achou o resultado surpreso. Os alunos responderam a este problema muito diferentemente do problema da Nova Zelândia, “apesar de serem análogos - obter 1 ou 2 no dado corresponde à região do topo no mapa onde um terço da população mora e obter um 3, 4, 5 ou 6 corresponde a cada uma das outras regiões do mapa” (MELETIOU, 2000, p. 145).

Meletiou (2000, p.145-146) explica que “os alunos não estão completamente errados, pois muitos outros fatores podem influenciar a ocorrência de defeitos nas crianças nascidas, mas eles deveriam imaginar que 7 crianças é uma amostra muito pequena para tomar decisões”. Segundo a autora, os alunos deveriam ter mostrado a mesma sensibilidade para o efeito do tamanho da amostra como foi mostrado no problema 4 da entrevista, apresentado a seguir, na Figura 38.

| |
|---|
| 4) Um psicólogo infantil está envolvido num estudo sobre dois brinquedos infantis. Das primeiras cinco crianças estudadas, 4 mostraram uma preferência por um mesmo brinquedo. O psicólogo concluiu que a maioria das crianças mostrará uma preferência por esse mesmo brinquedo. Você acha que o psicólogo apresentou uma conclusão válida? |

Figura 38 - Questão 4 da entrevista realizada por Meletiou (2000).

Todos os estudantes entrevistados individualmente desafiaram a conclusão do psicólogo.

> Tim disse: 4 de 5, eu sei que 4 de 5 dentistas preferem esta pasta de dente, mas eu diria você precisa de pelo menos 100 crianças... eu poderia pegar 5 crianças e persuadir 4 delas. Esta resposta contrasta com o que ele respondeu para a questão do defeito de nascimento.
> Entrevistador: Somente olhando o mapa, você vê alguma conexão entre onde moram e quantas crianças nasceram com defeito no membro?
> Tim: Sim. Eles correlacionam porque 1/3 que mora tem zero porque provavelmente há mais doutores e mais hospitais e somente 1/6 mora lá, então deve haver coisa errada por lá. Sim, tem que haver uma razão.
> Entrevistador: Você vê que os números são pequenos? Você acha que isto deveria ser levado em consideração?
> Tim: Por que? (MELETIOU, 2000, p. 146)

Após o diagnóstico, o professor deu início à fase de ensino, cuja estratégia utilizada foi a solução de problemas, em que os conteúdos necessários seriam introduzidos de maneira aplicada e articulada com outros conteúdos. Ou seja, para a solução de um determinado problema, o professor discutia as ferramentas estatísticas possíveis para descrever e interpretar aquela situação.

Quando foram discutidas as medidas de tendência central e dispersão, Meletiou (2000) explica que a ênfase era dada no significado e que a medida de variação mais difícil foi o desvio padrão.

Na etapa de ensino, a questão 10 do pré-teste (que relaciona o conceito de histograma e de assimetria da distribuição) foi explorada com a inserção dos *boxplots* de cada distribuição, para que os alunos fizessem a relação entre histogramas e *boxplots*. Para decidirem se a distribuição tinha um desvio padrão alto, eles mobilizaram dois conceitos: grande número de observações e barras altas no final da escala. O professor disse que era um quebra-cabeça, pois eles iriam analisar tudo junto: média, mediana, desvio padrão, histograma, *boxplots* e verificar como tudo isto estava relacionado com variabilidade. A primeira intervenção do professor foi concordar com os alunos que o formato da distribuição (simétrico ou assimétrico) é uma estratégia para avaliar a variabilidade e fizeram a relação entre média e mediana nos histogramas simétricos e assimétricos.

Trabalharam a comparação entre médias de dois grupos, desde que levando em conta a variação dentro de cada grupo (assim como a mediana e o intervalo interquartílico). Após a etapa de ensino, os autores pediram que os alunos analisassem os dois *boxplots* e chegassem a uma conclusão. A maioria dos alunos percebeu que o intervalo interquartílico era muito semelhante nos dois grupos, o que não permitiria comparar os escores. Os autores, por sua vez, perceberam que alguns alunos ainda pensavam que a linha do meio no *boxplot* era a média.

Trabalharam probabilidade dando ênfase aos conceitos de independência e acaso. Trabalharam distribuição binomial, normal, distribuição amostral, intervalo de confiança e teste de hipótese.

Ao término da etapa de ensino, Meletiou (2000) aplicou um pós-teste, cujas questões foram divididas em cinco grandes áreas temáticas: a variação envolvida em análise exploratória de dados, a produção de dados, o conceito de independência, a variação amostral e a representatividade e inferência estatística.

Para verificar o conceito de variabilidade em análise exploratória, ela usou quatro atividades e os resultados obtidos estão discutidos a seguir.

10) Suponha que duas distribuições tenham exatamente a mesma média e o mesmo desvio padrão. Então as duas distribuições são exatamente iguais.
a) Verdadeiro
b) Falso
Explique sua resposta

Figura 39 - Questão 10 do pós-teste aplicado por Meletiou (2000).

Quase todos os alunos reconheceram que podem ocorrer muitas formas diferentes de amostras que tenham os mesmos atributos.

A questão da comparação dos dois histogramas foi aplicada novamente e 15% dos alunos ainda escolheram o histograma da *Distribuição A* como tendo maior variabilidade (resposta errada), pois não olharam o eixo horizontal e apenas a altura das barras.

Ainda trabalhando com distribuição, foi apresentada a seguinte questão:

6) Na construção do histograma para descrever a distribuição dos salários de pessoas com 40 anos ou mais, que ainda estão no mercado de trabalho, explique:
a) O que está no eixo y:
b) o que está no eixo x:
Qual seria o formato da distribuição dos salários? Explique.

Figura 40 - Questão 6 do pós-teste aplicado por Meletiou (2000).

Muitos alunos confundiram com o gráfico de dispersão (*scatterplot*) talvez relacionando salário com idade. Mas, 42% dos alunos conseguiram imaginar o histograma com a variável salário representada no eixo das abscissas e a freqüência (relativa) das pessoas no eixo das ordenadas e que a distribuição deveria ser assimétrica para a direita, pois a maioria ganharia em torno do mesmo valor e poucos ganhariam muito mais (MELETIOU, 2000).

E a última questão do pós-teste que envolvia análise exploratória era:

11) Em uma pequena sala de aula foi aplicado um teste de aritmética e foram observadas as notas obtidas. O mesmo teste foi aplicado algumas semanas depois. Os *boxplots* para os dois conjuntos de notas são mostrados (ambas as caixas tinham amplitude, mediana e valores máximo e mínimos semelhantes)

As notas tiveram mudanças significativas? Sim ou Não. Explique sua resposta.

Figura 41 - Questão 11 do pós-teste aplicado por Meletiou (2000).

Era uma questão já utilizada na etapa de ensino e quase todos os alunos (exceto 3) reconheceram que os dois intervalos interquartílicos eram quase idênticos, significando que a mudança no teste, provavelmente, não era significativa .

A autora concluiu que, ao final da disciplina, os alunos entenderam que “freqüência é como um monte de gente em cada categoria” e que a “distribuição é como o gráfico se parece... quantas pessoas estão aqui e quantas pessoas estão lá”, que uma medida de centro sempre requer uma visão da dispersão e perceberam que existem outras medidas de dispersão além do desvio padrão como a 'variação da caixa' (MELETIOU, 2000 p. 227 e 228).

### 4.3 O conceito de Variabilidade apresentado por Professores em Formação ou em Atuação

Como salientado por Canada (2006), as pesquisas com alunos do ensino fundamental têm contribuído para a compreensão do raciocínio de variação de estudantes nessa etapa escolar, mas pouco tem sido publicado acerca da concepção de professores (em formação) sobre este assunto.

Foram encontrados três estudos sobre o raciocínio de variação com professores, sendo dois com professores em formação e um com professores em atuação, mas apenas um estudo foi realizado com futuros professores de Matemática.

Makar e Confrey (2005) fizeram uma pesquisa com dezessete futuros professores de Matemática e Ciências, sendo três homens e quatorze mulheres (EUA). O objetivo do estudo era identificar a linguagem padrão e não padrão utilizadas pelos sujeitos na discussão sobre variação[27].

Os autores focaram na linguagem não padrão, pois o fato de um professor usar a linguagem padrão não permite inferir que ele assimilou o conceito, ou seja, que esteja 'vendo' a variação que está medindo. Eles fizeram duas entrevistas com os professores, no início e no final de uma disciplina de um semestre. O conteúdo da disciplina foi: gráficos (histograma, *boxplots* e gráfico de pontos), estatística descritiva (média, mediana, desvio padrão, intervalo interquartílico, formato da distribuição), regressão linear (associação, correlação, mínimos quadrados e resíduos) e uma breve introdução à distribuição amostral e inferência, com o uso do *software Fathom*.

A tarefa, idêntica nas duas entrevistas, foi solicitar que comparassem duas distribuições de dados apresentadas em um gráfico de pontos. A tarefa solicitada foi determinar a efetividade de um programa de recuperação matemática denominado *Enrichment* para alunos de 8ª séries. Foram comparadas

[27] Os autores se referem à linguagem padrão como sendo os termos estatísticos conhecidos e a linguagem não padrão aos termos que os próprios sujeitos utilizariam para explicar a variação que estavam observando.

suas notas da 7ª. Série e as notas obtidas no final da 8ª. Série, representadas pelo gráfico de pontos apresentado na Figura 42.

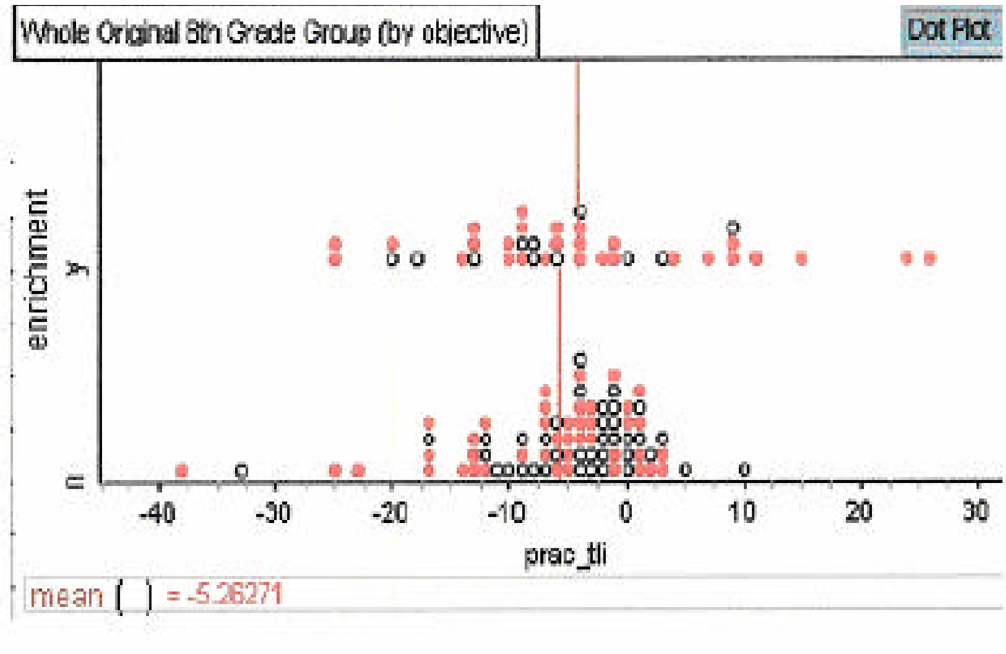



Figura 42 - Gráfico apresentado aos professores do estudo de Makar e Confrey (2005)

O gráfico de pontos na Figura 42 apresenta as notas reais de uma oitava série regular (pontos abaixo) e uma oitava série *Enrichment* (pontos acima) e cada ponto se refere à diferença entre a nota da 8ª série e a nota da 7ª série, ou seja, se o valor do eixo das abscissas for positivo, indica que a nota da 8ª série foi maior que a nota da 7ª série. Os pontos vermelhos indicam os alunos economicamente em desvantagem (mas esta característica não foi explorada neste estudo).

A média de cada grupo (marcada na Figura 42 pela linha vertical vermelha) e a média global (-5,26) foram apresentadas no gráfico o que permitiu verificar se os futuros professores interpretariam uma pequena diferença nas médias de maneira determinística ou se eles esperariam alguma variabilidade entre as médias (MAKAR E CONFREY, 2005).

Os autores explicam que os dados são confusos, justamente por serem reais, o que tornou a análise mais difícil, porém permitiu identificar como os futuros professores interpretariam a situação, que é semelhante à que iriam enfrentar quando concluíssem seu curso.

Os termos padrões utilizados para fazer a análise do gráfico foram, numa ordem decrescente de utilização: proporção ou número de alunos que melhoraram, média, máximo e mínimo, tamanho da amostra, *outliers*, amplitude, formato da distribuição e desvio padrão. Na primeira entrevista, nenhum professor usou o termo desvio padrão, enquanto, na segunda entrevista, apenas dois utilizaram. O comentário de um dos dois professores é bastante ilustrativo de seu raciocínio quanto ao conceito de desvio padrão: "Provavelmente o desvio padrão vai ser, como, realmente grande neste (*Enrichment*) quando comparado com aquele (não-*Enrichment*), pois este está bem espalhado, bem longe" (MAKAR e CONFREY, 2005, p. 38).

É interessante notar que os futuros professores usaram o termo desvio padrão, mas quando dizem bem espalhado, bem longe, não fazem referência à média e usam o termo variação como predicado e não como substantivo (como já descrito no trabalho de Bakker, 2004).

Makar e Confrey (2005) utilizaram o termo *variation-talk* para se referir às palavras não padrão que expressavam o conceito de variação. Esses termos foram organizados em duas categorias. A primeira categoria foi denominada "*spread"*, cujos termos semelhantes foram "*clustered, clumped, grouped, bunched, gathered, spread out, evenly distributed, scatterd, dispersed*"[28], todos no particípio passado, que segundo os autores, referem-se à variação como uma característica do formato em vez de uma medida (MAKAR e CONFREY, 2005, p.48).

A segunda categoria de termos refere-se aos substantivos *"triads, modal clump e distribution chunks"* (tríade, grupo modal e parte da distribuição), em que os sujeitos dividem o conjunto de dados para analisar. Os autores salientam que eles não dividiam em quatro partes, o que poderia dar uma noção do *boxplot*. Os autores discutem que, apesar dos futuros professores estarem usando termos não padrão, o conceito que eles estavam discutindo estava longe de ser simplista e que precisava ser reconhecido como conceito estatístico.

---

[28] Como o estudo referia-se à linguagem empregada, foram mantidos os termos em inglês, haja vista que a tradução pode não representar adequadamente as palavras.

Outro estudo realizado com docentes foi de Hammerman e Rubin (2004), que tinham como objetivo verificar as estratégias do raciocínio estatístico empregadas por professores para lidar com questões de variabilidade quando analisavam dados e para verificar novas oportunidades que o *software TinkerPlots*™ proporcionava. Eles trabalharam com <u>onze professores norte-americanos, sendo seis de 6ª a 8ª séries do ensino fundamental e cinco do ensino médio</u> (os autores não especificaram se eram professores de Matemática, mas explicaram que eram professores participando de um projeto de desenvolvimento profissional denominado *VISOR,* durante dois anos, com três horas-aula em dois encontros semanais). Como acompanharam os professores em sala de aula, fizeram também um experimento com doze alunos de sexta série do ensino fundamental, durante treze semanas. Foram escolhidos dois conjuntos de dados reais já utilizados em outros estudos.

Os resultados encontrados salientam que, em geral, os professores raramente utilizaram medidas de centro como primeira alternativa para comparar dois grupos apresentados graficamente. A maioria dos professores utilizou a estratégia de dividir o conjunto de dados em duas partes (o que os autores denominaram de *cut point*) e verificar a freqüência em cada grupo. Esse ponto de corte não dizia respeito à mediana e nem a outra medida de tendência central, mas sim a uma medida de contexto (os professores escolheram o valor 500 que representava células por mililitro e que os médicos consideravam o ponto de corte para diagnosticar indivíduo sadio ou não). Os professores também dividiram o conjunto de dados em mais de dois grupos (o que os autores chamaram de *slices*).

Quando representaram horas de estudo, a estratégia foi semelhante à elaboração da distribuição de freqüência com dados agrupados, pois escolheram amplitude 4 e contaram o número de casos (e também a porcentagem) em cada subgrupo de tamanho 4. Porém, na elaboração da análise, os professores consideraram apenas os intervalos de 4 a 8 e de 8 a 12 que, somados, representavam 51% (27% de 4 a 8 horas e 24% de 8 a 12 horas) dos casos em uma escola e 48% (16% de 4 a 8 horas e 32% de 8 a 12 horas) dos casos na outra escola e chegaram à conclusão que os alunos da segunda escola tinham mais tempo para as tarefas de casa, ignorando o restante da distribuição.

Uma outra estratégia utilizada por professores para analisar os dados foi criar categorias para duas variáveis quantitativas ( o que os autores denominaram de matriz de covariância) em que analisaram um conjunto de dados dos estados e fizeram a seguinte categorização: idade mediana dos eleitores nos estados americanos ( 25 a 30, 30 a 35, 35 a 40 e 40 a 45) e a porcentagem de pessoas que votaram no Bush em 2000 (25 a 35%, 35 a 45%... 65 a 75%). E tiraram conclusões como: “poucos estados em que a idade mediana está entre 40 e 45 anos tiveram uma porcentagem de votos entre 45% e 55%” e os autores explicam que com este tipo de análise os professores fizeram descrições com foco determinista dos dados em vez de levar em consideração o ruído subjacente a todos os dados. (HAMMERMAN e RUBIN, 2004, p.35).

Os autores concluíram que, à medida que está disponível uma representação gráfica, é mais difícil aceitar uma medida de centro para representar toda a distribuição que tem sua particular variação e formato. E que a função *binning* do *software,* que permite a quebra da distribuição em pequenos subgrupos em que a variabilidade seja menor, permitiu aos professores realizar uma análise. E que o fato de o *software* apresentar o número e a porcentagem de casos em cada subgrupo permite ao participante o desenvolvimento do raciocínio proporcional. Os autores justificam que o *software* permite diferentes representações e medidas e que isto pode proporcionar ao professor a discussão de diferentes maneiras de se fazer uma análise do mesmo conjunto de dados.

Ao analisar o artigo de Hammerman e Rubin (2004), Pfannkuch (2005) questiona se a freqüência disponibilizada pelo *software* para cada subgrupo fosse substituída pela elaboração de um gráfico de barras, o pensamento dos professores mudaria. Em outras palavras, Pfannkuch questiona o uso do *software* para realizar uma tarefa que já é muito discutida em livros e, geralmente, já é de domínio dos professores, que é o gráfico de barras.

Canada (2006) realizou um estudo com <u>trinta professores em formação para as primeiras séries do ensino fundamental,</u> nos Estados Unidos. O objetivo da pesquisa era verificar o raciocínio de variação, em situação probabilística, antes e depois de uma disciplina denominada Matemática II, em que foram dedicadas quatro semanas para a temática Estatística e Probabilidade. Segundo o autor, como este assunto não foi tratado na disciplina Matemática I, o conhecimento de Probabilidade e Estatística dos participantes era referente à etapa escolar anterior à graduação. Para a realização das simulações gráficas, foram utilizados os *softwares Fathom* e *ProbSim*.

Na primeira semana, o professor (que não era o pesquisador) fez um pré-teste sobre os assunto-alvo da disciplina. Nas três semanas seguintes, desenvolveu a etapa de ensino e o pós-teste foi representado pelas atividades que os participantes desenvolviam em casa.

Além das situações observadas em sala de aula, Canada (2006) contou com onze voluntários que se dispuseram a participar de entrevistas após cada aula, o que proporcionou a oportunidade de compreender mais profundamente o raciocínio de variação utilizado para resolver os exercícios.

A atividade de ensino focou três temáticas: dados e gráficos, amostragem e probabilidade. Na primeira temática, que durou uma semana, o autor pediu para que os participantes fizessem uma pesquisa na própria sala de aula e a atividade está sintetizada na Figura 43.

1) Quantos animais de estimação você tem?

2) Quantos anos você mora na cidade?

3) Quantas pessoas moram em sua casa?

4) Quantas moedas você tem hoje?

5) Meça a envergadura de seu braço. Meça também de um colega.

6) Meça sua altura.

7) Meça sua circunferência da cabeça.

8) Meça sua palma da mão

9) Conte sua pulsação em um minuto.

Figura 43 - Atividade desenvolvida na etapa de ensino de Dados e Gráficos do estudo de Canada (2006, p. 39 e 40).

As quatro primeiras questões foram utilizadas para explorar gráficos e medidas e as outras questões foram utilizadas para discutir as fontes de variação.

A segunda etapa de ensino foi sobre amostragem e a atividade utilizada está descrita na Figura 44.

A banda de uma escola tem 100 pessoas, 70 mulheres e 30 homens. Para organizar uma viagem, a banda vai montar uma comissão com 10 membros, que serão selecionados a partir de um sorteio, em que o nome de cada um foi colocado em um chapéu.

Figura 44 - Atividade desenvolvida na etapa de ensino de Amostragem do estudo de Canada (2006, p. 39 e 40).

Os alunos discutiram o que aconteceria se fossem tiradas 30 amostras de 10 pessoas e simularam esta atividade utilizando pedras. Em seguida, o professor pegou 550 pedras amarelas e 450 pedras verdes e colocou em um grande pote. Sem saber a proporção de pedras de cada cor que estavam no pote, os alunos deveriam retirar amostras de qualquer tamanho e quantas amostras quisessem para estimar o número de pedras amarelas e o número de pedras verdes que tinham no pote.

E a intervenção de probabilidade está apresentada na Figura 45.

**Caixas de Cereais**: Há cinco tipos diferentes de cereais dentro de cada caixa e cada cereal tem a mesma probabilidade de ser retirado. Quantas caixas precisam ser abertas para obter todos os cinco tipos?

**Jogo do Rio**: Dois jogadores colocam 12 pedras de cada lado do “rio”, de maneira que as pedras estejam dispostas em posições de 1 a 12. Jogam-se dois dados e a soma dos pontos dos dois dados permite retirar a pedra que ocupa a posição do somatório. Vence o jogador que acabar primeiro com as pedras.

Figura 45 - Atividade desenvolvida na etapa de ensino de Probabilidade do estudo de Canada (2006, p. 41).

Para simular a situação da caixa de cereais, o professor utilizou uma roleta dividida igualmente em 5 partes. E para o Jogo do Rio, o professor omitiu a informação que o resultado 1 era impossível, deixando que os alunos descobrissem sozinhos.

A questão utilizada para investigar o raciocínio de variação é semelhante à de Torok (2000) e Meletiou (2000) e está apresentada na Figura 46.

**Um set – O quê** : Se você jogar uma moeda 50 vezes, quantas vezes vai sair cara?

**Um set – Por quê** : Por quê você pensa isto?

**Comparação dos sets** - Se Mark resolver lançar novamente 50 vezes a moeda, como você acha que será o resultado em comparação com primeiro set de 50 lançamentos?

**Seis sets – o quê** : Mark está com muito tempo disponível e resolve jogar mais 6 sets de 50 lançamentos da moeda. Escreva uma lista dos possíveis números de caras nos 6 sets.

**Seis sets – Por quê:** Por quê você escolheu estes números?

Figura 46 - Questão do pré-teste utilizada por Canada (2006, p. 38)

Para esta situação apresentada na Figura 46, o pós-teste foi a atividade da roleta dividida ao meio (tal como a atividade de Torok, 2000), cujas perguntas eram as mesmas utilizadas no exercício em sala de aula.

Apoiado nos estudos de Wild e Pfannkuch (1999), Reading e Shaughnessy (2004) e Garfield e Ben-Zvi (2005), Canada (2006) elaborou uma estrutura conceitual para analisar seus resultados, que está sintetizada na Figura 47.

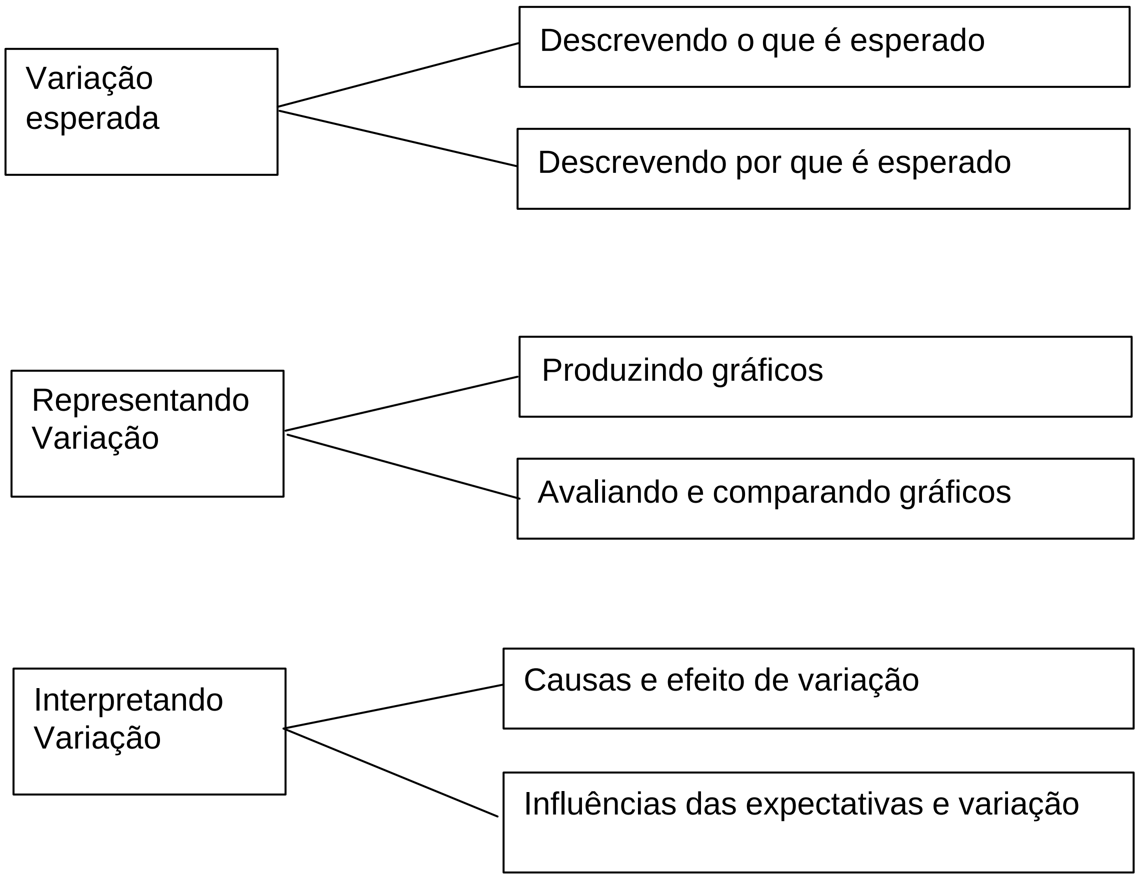



Figura 47 - Síntese da estrutura conceitual utilizada por Canada (2006, p. 42) para analisar os resultados de sua pesquisa

Para a questão ***Um set - O que**,* apenas um aluno respondeu um número em torno de 25 no pré-teste enquanto, no pós-teste, dez alunos apresentaram este tipo de resposta. A maioria dos alunos respondeu exatamente 25, em ambos os testes.

Para a questão ***Um set – Por quê*** , a maioria dos alunos explicaram utilizando os conceitos de razão ( 1:2 ) e porcentagem (50%), tanto no pré como no pós-teste. Quando os alunos ***compararam os sets***, no pré-teste, a maioria respondeu que os resultados seriam semelhantes, justificando com a utilização de termos como razão e porcentagens, enquanto, no pós-teste, a maioria respondeu que os resultados seriam semelhantes e apresentaram uma amplitude para a variação das respostas.

Para a questão de ***seis sets***, no pré-teste houve dois tipos de respostas predominantes: escolhas inapropriadas e escolhas apropriadas com algum tipo de raciocínio proporcional. A última predominou no pós-teste.

A relação entre estes resultados e a estrutura conceitual do autor restringiu-se apenas aos dois aspectos: variação esperada e interpretação da variação.

Quanto à variação esperada, no pré-teste e nas situações de ensino, mesmo os alunos que consideravam alguma variação em torno do valor 25, imaginavam que a média dos valores deveria ser 25. No pós-teste, os alunos utilizaram o raciocínio proporcional combinado com um entendimento do que era provável em face de variação e a resposta de um aluno (George) ilustra isto: “Bem, 25 seria metade, e 20 é possível. É possível obter um número alto. Você sabe, é possível obter 36? Poderia acontecer? Claro! Claro que poderia acontecer. É muito improvável” (CANADA, 2006, p. 53). Os resultados referentes à interpretação da variação estão sintetizados no Quadro 10.

Quadro 10 - Resultados do pré e pós-teste de Canada (2006) referente à interpretação de variação

| Dimensões da interpretação da variação | Pré teste | Pós teste |
| --- | --- | --- |
| Causa da Variação: busca pela explicação | Explicação física: clima, temperatura, tipo de roleta, etc. | Explicação natural: aleatoriedade |
| Efeito da variação na percepção do aluno e na sua tomada de decisão | “Alguma coisa pode acontecer” e “não sei o que”. Este tipo de pensamento gerou o seguinte problema:<br>“Alguns alunos consideraram que todos os resultados eram equiprováveis” (CANADA, 2006, p. 57) | “A maioria dos participantes pensou que, apesar de não saber o resultado, ao certo, é possível fazer afirmações razoáveis da expectativa.” (CANADA, 2006, p. 58) |
| Influencia da expectativa e variação | Os alunos compreendiam que os resultados poderiam ser 25, mas achavam que a média dos resultados deveria ser 25. | Entenderam que os resultados deveriam ser mais próximos de 25 devido a uma explicação teórica. |

Estes três estudos realizados com professores (em formação ou em atuação) permitem fazer algumas interpretações gerais. A primeira é o uso comum da medida de centro para observar uma variável. Isto não ocorreu no estudo de Hammerman e Rubin (2004) e pode ser em decorrência da ênfase na utilização da ferramenta *binning* do *software* utilizado, que permitia quebrar a distribuição em partes.

Uma segunda observação nestes estudos é que houve a percepção da variação e a busca pela sua explicação. No estudo de Hammerman e Rubin (2004), os professores buscaram explicá-la ao dividir a distribuição em partes, no estudo de Makar e Confrey (2005), os professores perceberam que havia mais variação na 8ª série *Enrichment,* pois os pontos estavam mais espalhados e no estudo de Canada (2006), embora tratasse de situação de acaso, os professores perceberam que haveria variação da proporção em diferentes amostras e que poderia haver variação na proporção em relação ao valor esperado (teórico).

Mas, em nenhum estudo foi relatada a preocupação em medir esta variação. Então, pode-se supor que os estudos restringiram-se ao estudo de variabilidade e não de variação e isso não permitiu encontrar resultados muito diferentes dos estudos realizados com alunos do ensino fundamental e médio.

### 4.4 Estudos Específicos sobre as Medidas de Variação

Segundo Watson e Kelly (2002), o desvio padrão é a medida de variabilidade mais comum, mas devido a sua natureza complexa, ela é freqüentemente evitada no currículo de matemática do ensino fundamental australiano. E Delmas e Liu (2005, p. 56) declaram que a maioria do "ensino sobre o desvio padrão tende a enfatizar a fórmula, praticar os cálculos e atrelar o desvio padrão à regra empírica da distribuição normal".

Para a compreensão do conceito de desvio padrão, Delmas e Liu (2005) consideram que o aluno precisa mobilizar três conceitos estatísticos: distribuição, média e desvios da média.

O conceito de distribuição, por sua vez, requer a mobilização de outros conceitos como os valores assumidos pela variável e a freqüência acumulada. A visualização da variação em uma distribuição requer a observação de, pelo menos, os valores da variável, a densidade de freqüência destes valores e a média aritmética como o ponto de equilíbrio de uma balança.

De posse desses conceitos, Delmas e Liu (2005, p. 56) argumentam que um aluno pode imaginar o terceiro conceito fundamental: desvios da média.

> É através da coordenação da distribuição (como representada pela coordenação de valor e freqüência) e desvio (como distância da média) que um conceito dinâmico de desvio padrão é derivado como a densidade relativa de valores em torno da média.

Hart (1984, p.24) sugere ensinar o desvio padrão a partir de um intervalo. Segundo a autora, "se a medida de variabilidade for chamada de *d*, então pode ser possível afirmar que *x %* das observações ficam entre *1d* da média", e afirma que é perfeitamente possível para o ensino do desvio médio e também pode ser possível para o ensino do desvio padrão.

Quando Hart (1984) pergunta a seus alunos o que é o desvio padrão, as respostas, na maioria das vezes, foram: o desvio padrão é uma medida de variação ou então os alunos reproduzem a fórmula do desvio. Loosen, Lioen e Lacante (1985) questionam estas respostas fornecidas a Hart. Para estes autores, a maioria dos alunos não imagina que o desvio padrão é uma medida especial de variação que "mede o quão fortemente os dados afastam-se da tendência central"

Loosen, Lioen e Lacante (1985) afirmam que os livros didáticos enfatizam a heterogeneidade entre as observações e não dos desvios em relação à tendência central. Eles realizaram um experimento com cento e cinqüenta e quatro alunos do primeiro ano de graduação em psicologia, na Bélgica.

Eles apresentaram os gráficos da Figura 48, em que o valor acima de cada barra representa o valor da observação e não sua freqüência. Logo, o conjunto A tem duas observações, cujos valores são 45 e 50.

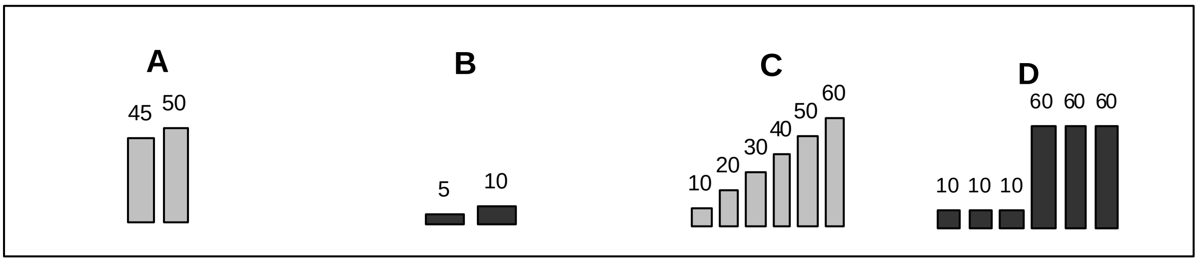



Figura 48 - Primeira seqüência de gráficos usados por Loosen, Lioen e Lacante (1985)

Com relação aos conjuntos A e B, os autores perguntaram: a) em qual dos dois conjuntos você acredita que os blocos são mais diferentes? b) Em qual dos conjuntos os blocos tem a maior variação entre eles mesmos? e c) Qual conjunto apresenta a maior flutuação com relação a altura: A ou B?

Foi verificado que 69% dos alunos disseram que B tinha maior variação, 11% disseram que A tinha maior variação e 20% disseram que tanto A quanto B apresentavam a mesma variação.

Quando foram apresentados os gráficos C e D, foram feitas as mesmas perguntas e 50% dos alunos consideraram que a diversidade em C era maior, 36% pensaram que era em D e 14% disseram que a variação era a mesma tanto em C quanto em D.

Segundo Loosen, Lioen e Lacante (1985), os resultados sugerem que, quando os alunos compararam A e B, a variabilidade entre os dados foram avaliadas relativamente em vez de absolutamente e quando compararam C e D, os alunos apresentaram uma tendência a interpretar, intuitivamente, o conceito de variabilidade em termos de *unalikability*, definida pelos autores como a escassez de observações do mesmo tipo ou a escassez de grupos com observações idênticas.

A sugestão dos autores para introduzir o conceito de variabilidade é iniciar a etapa de ensino com os gráficos A, B, C e D e, em seguida apresentar os seguintes gráficos:

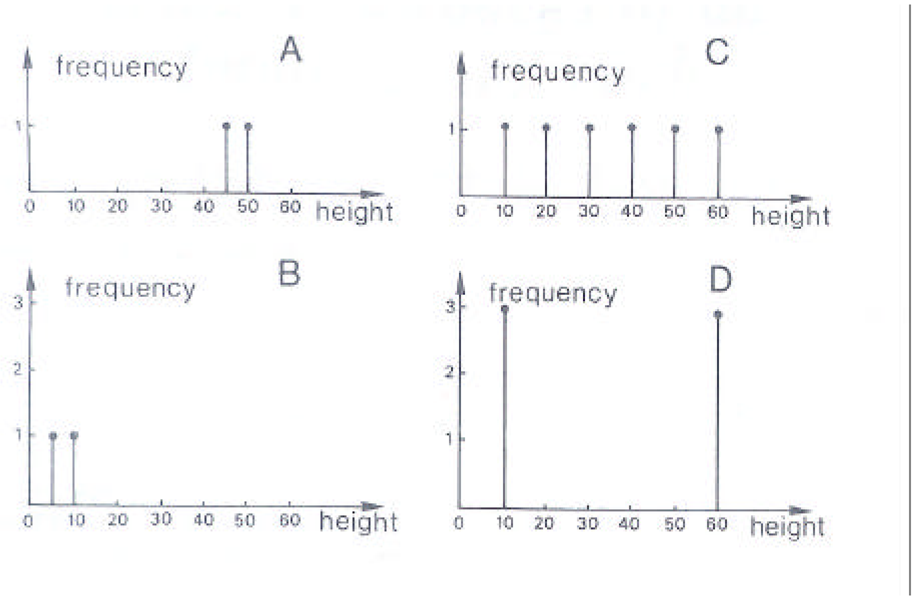



Figura 49 - Segunda seqüência de gráficos usados por Loosen, Lioen e Lacante (1985)

Loosen, Lioen e Lacante (1985, p.5) sugerem que o resultado numérico do desvio padrão proporciona pouca informação e pode ser melhor traduzido dentro da teoria de probabilidade: "uma proporção *p* das observações no conjunto de dados fica entre ± *k* desvios padrão da média". Segundo eles, quando a distribuição é conhecida, esta proporção pode ser calculada, mas, em outros casos, será necessário contentar-se com o cálculo do limite mais baixo da proporção, por meio do Teorema de Tchebichev (apresentado no Capítulo 2 deste trabalho).

Embse e Engebretsen (1996) sugerem iniciar o trabalho com alunos do ensino médio sobre o desvio padrão fazendo uso de calculadora gráfica. Para exemplificar, os autores trabalharam com as calorias de 20 doces e representaram da seguinte maneira: cada doce foi representado por uma barra no gráfico, o eixo das ordenadas apresentava a escala de calorias, a média aritmética foi representada por uma reta paralela ao eixo das abscissas e outras duas retas paralelas ao eixo das abscissas foram traçadas para representar um desvio padrão acima da média e um desvio padrão abaixo da média.

Os autores salientam a diferença existente entre gráfico de barras e histograma, porém denominaram de histograma o gráfico das calorias dos 20 doces. Como já salientado por Meletiou e Lee (2002) existe grande dificuldade dos alunos para interpretar um histograma e é possível que esta representação colabore ainda mais para esta dificuldade. Além disso, Loosen, Lioen e Lacante, já em 1985 salientavam complementar esta informação com o gráfico de bastão (conforme Figura 49).

Delmas e Liu (2005) realizaram um estudo para verificar o raciocínio dos alunos sobre a magnitude do desvio padrão, quando as observações estão representadas em uma distribuição de freqüência. Os autores utilizaram um aplicativo desenvolvido por eles mesmos, que permitia ao aluno movimentar as barras do gráfico e perceber a mudança nas medidas de tendência central e dispersão. Além disso, cada barra do gráfico apresentava o desvio absoluto em relação à média aritmética. A Figura 50 apresenta um exemplo do tipo de gráfico fornecido pelo aplicativo desenvolvido por eles.

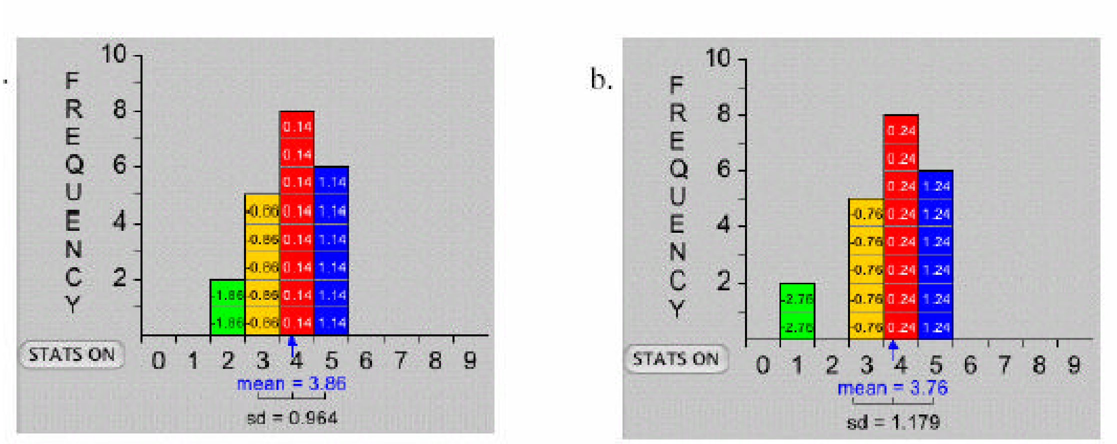



Figura 50 - Exemplo de gráficos fornecidos pelos aplicativo desenvolvido por Delmas e Liu (2005)

Como pode ser observado na Figura 50, o deslocamento da barra verde (que continha duas observações) do valor 2 para o valor 1 no eixo das abscissas proporcionou um valor absoluto maior do desvio padrão (de 0,964 para 1,179).

Para os autores, a interação com o *software* poderia permitir ao aluno entender que o desvio da média e a freqüência, ambas combinadas, determinam o valor do desvio padrão, e que para isto é necessário fazer a distinção entre valor e freqüência, reconhecer a distância de cada valor em relação à média, entender que uma distribuição e sua imagem espelhada tem o mesmo desvio padrão e entender que o valor do desvio padrão é independente de onde a distribuição é centrada. Além disso, o aluno poderia compreender a relação entre o formato da distribuição e a magnitude do desvio padrão.

Foram participantes da pesquisa doze estudantes voluntários (cinco homens e sete mulheres) que tinham participado de uma disciplina de Introdução à Estatística em uma universidade nos EUA. Foram feitas três entrevistas, em uma fase introdutória, exploratória e de teste. Na fase introdutória os sujeitos deveriam movimentar duas barras de um gráfico e observar a mudança dos valores das medidas (média, desvio padrão, desvios absolutos, desvios quadrados e a fórmula do desvio padrão). Após alguns minutos na fase introdutória, os alunos decidiam quando passar para a fase exploratória, em que iriam buscar os fatores que afetavam o desvio padrão. Com cinco pares de distribuições diferentes, os alunos deveriam encontrar um formato que produzisse o maior e o menor valor de desvio e justificar a escolha.

Os autores encontraram onze categorias de justificativas para esta fase exploratória, em que cinco categorias referiam-se a justificativas para produzir um alto valor do desvio padrão: barras distantes umas das outras (sem referência à média), barras igualmente espalhadas, barras distantes da média, distribuição equilibrada de valores acima e abaixo da média, nos valores extremos da escala numérica, alto valor numérico da média. As categorias (quatro) referentes ao menor valor do desvio padrão foram: barras contíguas em ordem crescente ou decrescente de freqüências, mais barras (ou valores) no meio, distribuição simétrica, equilíbrio (balança).

Para justificar o mesmo desvio padrão foram encontradas duas categorias de respostas: imagem espelhada da distribuição e locação (em que o aluno poderia perceber que a mesma combinação das barras, independente da localização, produz o mesmo valor do desvio padrão –mas não da média).

Na terceira etapa, foram elaborados 10 pares de histogramas onde o primeiro histograma de cada par continha a média e o desvio padrão da distribuição e o segundo histograma só continha a média e os alunos foram solicitados a responder se o desvio padrão seria maior, menor ou igual ao desvio padrão da primeira distribuição.

Os alunos não apresentaram dificuldade para observar que o desvio padrão é o mesmo em uma distribuição espelhada e independente da localização, mas poucos estudantes perceberam que, em uma distribuição simétrica, o desvio padrão tende a ser menor. Para trabalhar esse tipo de concepção, os autores criaram um par de histogramas em que a distribuição simétrica tinha maior desvio padrão, pois a amplitude da escala numérica era maior (tal como a questão 7 do pré-teste de Meletiou 2000). Amplitude da variação dos dados não foi um conceito levado em consideração por muitos sujeitos da pesquisa no momento de justificar o maior ou menor desvio padrão, que ratifica os resultados encontrados por Meletiou e Lee (2002). Um dos pares foi criado com o objetivo de levar os alunos a repensar que distâncias iguais entre uma barra e a seguinte produzem menor desvio.

O estudo desses autores foi o primeiro encontrado em que são levantados os conceitos que devem ser mobilizados para compreender o desvio padrão de uma distribuição: valores da distribuição, freqüência (densidade de freqüência), média e as distâncias da média, todos combinados. Porém, deve-se tomar cuidado ao reproduzir estudo semelhante. Os autores referiam-se ao valor absoluto do desvio padrão e que nem sempre corresponde a uma variação relativa maior, pois isto depende do valor da média (como já discutido no Capítulo 2 sobre o coeficiente de variação).

Os autores concluem que as idéias do desvio padrão formadas pelos alunos tais como barras contíguas, amplitude, média no meio e valores distantes resgatam aspectos importantes desse conceito, mas podem representar níveis de entendimento superficial e fragmentado enquanto outras idéias como grande média e distâncias iguais das barras são inconsistentes. Outras idéias como distâncias da média, equilíbrio, mais valores no meio e distribuição simétrica podem representar um entendimento mais completo do desvio padrão.

## 4.5 Os Aspectos do Raciocínio de Variabilidade e Variação que apareceram nos Estudos Analisados

No SRTL-3 (Terceiro Fórum de Pesquisa em Letramento, Pensamento e Raciocínio Estatístico), os autores discutiram diferentes aspectos do raciocínio sobre variabilidade, tais como variação inerente aos dados, variabilidade representada em distribuições univariada e bivariada, o papel da variabilidade na comparação de grupos, o entendimento dos alunos sobre medidas particulares de variabilidade e variabilidade em diferentes contextos amostrais (Ben-Zvi e Garfield, 2004).

Esses aspectos do raciocínio sobre variabilidade e variação foram utilizados como categorias para agrupar os estudos analisados, que estão apresentados nos quadros seguintes. Além desses aspectos, outras categorias foram criadas, como a Explicação de Variação e a Variabilidade na descrição do formato da distribuição.

Quadro 11 - Explicação de Variação apresentada pelos sujeitos

| Autores e Nível dos sujeitos de pesquisa | Respostas e as estratégias de raciocínio utilizadas |
|---|---|
| Torok (2000)<br>Alunos de: 7ª e 8ª séries do ensino fundamental | Assemelha-se a uma pequena mudança.<br>Variedade de cores |
| Meletiou (2000)<br>Alunos de graduação | Variedade; Múltiplos valores; Inconstância; Variação de um máximo para um mínimo. |

Como pode ser observada no Quadro 11, a única medida de variação sugerida intuitivamente é a amplitude. Os alunos associam variação com a diferença entre os dados, ou seja, variação inerente, cujos estudos estão sintetizados no Quadro 12.

Quadro 12 – Percepção da variação inerente aos dados

| Autores e Nível dos sujeitos de pesquisa | Respostas e as estratégias de raciocínio utilizadas |
|---|---|
| Watson e Kelly (2002)<br>Alunos de 3ª série do ensino fundamental | Os alunos não tiveram dificuldade para perceber que existia variação na quantidade de balas de cada cor |
| Lehrer e Schauble (2002)<br>Alunos de 4ª e 5ª séries do ensino fundamental | Percebem variação entre os dados, mas não conseguem perceber variação em torno de uma medida (nem mesmo aceitar uma medida para representar o todo) |

Os alunos percebem naturalmente a diferença entre os valores, mas não sentem necessidade de representar esta variação com uma medida. Ou seja, a utilização das medidas de tendência central e dispersão para representar um conjunto de observações não acontece naturalmente.

As estratégias utilizadas pelos participantes da pesquisa para representar um conjunto de observações estão sintetizadas no Quadro 13.

Quadro 13 – Raciocínio sobre variabilidade na análise de uma distribuição

| Autores e Nível dos sujeitos de pesquisa | Respostas e as estratégias de raciocínio utilizadas |
|---|---|
| Reading (2004)<br>Alunos de: 6ª e 8ª séries do ensino fundamental e 2º ano do ensino médio | -Variação entre os dados;<br>-dividem o conjunto em pequenos grupos<br>-usam valores máximos e mínimos ou a amplitude<br>-Não usaram gráficos |
| Ben-Zvi (2002)<br>Alunos de 7ª série do ensino fundamental | Na etapa de elaboração de hipóteses, alguns alunos focam em padrões gerais, mas alguns elaboram com foco em pontos específicos ou de contexto.<br>Na etapa de análise de uma série temporal, poucos alunos conseguem “ver” o todo. Centram a atenção em pontos específicos |

Na análise de uma distribuição, as estratégias utilizadas foram a divisão do conjunto de dados em pequenos grupos, o foco em alguns pontos da distribuição e a utilização da amplitude. Ou seja, a única estratégia utilizada para trabalhar com toda a variabilidade foi o uso da amplitude, pois as outras estratégias têm como objetivo diminuir a variabilidade ou ignorá-la.

Os estudos que associaram o formato da distribuição e a variabilidade estão apresentados no Quadro 14.

Quadro 14 – Raciocínio sobre Variabilidade na descrição do formato da distribuição

| Autores e Nível dos sujeitos de pesquisa | Respostas e as estratégias de raciocínio utilizadas |
|---|---|
| Watson e Kelly (2002)<br>Alunos de 3ª série do ensino fundamental | Gráfico de barras: Descreveram como montanha |
| Bakker (2004)<br>Alunos de 8ª série do ensino fundamental | Representação contínua (semi-círculo, assimetria, etc):<br>Tendência a escolher distribuição simétrica |
| Meletiou (2000)<br>Alunos de graduação | Histograma: muitos alunos acharam que uma barra representava uma única observação, o que dificultou observar a variação. Escolheram o histograma cuja distribuição era simétrica. |

Tanto no estudo de Bakker (2004) quanto no estudo de Meletiou (2000) foi observada uma preferência a distribuições simétricas como representativas de menor variação. Isso também foi observado no estudo de Delmas e Liu (2005), quando os alunos tentavam identificar a distribuição com menor desvio padrão.

Distribuição é um conceito que está diretamente relacionado à variabilidade. Segundo Petrosino et. al. (apud Bakker, 2004, p. 65) distribuição poderia proporcionar uma organização da estrutura conceitual para pensar sobre variabilidade localizada dentro de um contexto mais geral de modelar os dados.

O maior número de estudos encontrados foi referente ao raciocínio sobre variação utilizado na comparação de duas distribuições, como pode ser observado no Quadro 15.

Utilizando diferentes representações gráficas, os estudos apresentam as estratégias que os alunos utilizaram para fazer as comparações. Independente da faixa etária dos sujeitos, houve uma tendência de utilizar partes da distribuição, os valores extremos e a amplitude total para fazer a análise. E, principalmente, existiu uma tendência para a não utilização de uma medida de centro para representar cada distribuição.

Segundo Hammerman e Rubin (2004), os professores usavam o pensamento agregativo (que focava em características emergentes de um conjunto de dados), mas sem envolver medidas de centro, ou seja, analisavam um conjunto de dados, encontravam um intervalo em que havia concentração dos dados mas sem pensar na média.

Quadro 15 – Raciocínio sobre Variabilidade na comparação de duas distribuições

| Autores e Nível dos sujeitos de pesquisa | Respostas e as estratégias de raciocínio utilizadas |
|---|---|
| Watson e Kelly (2002)<br>Alunos de 3ª série do ensino fundamental | Com uso de gráfico de pontos, os alunos fizeram análises em termos de grupos, buracos, valores mais comuns e menos comuns e amplitude. Discutiram que uma observação é diferente das demais. |
| Ben-Zvi (2004)<br>Alunos de 7ª série do ensino fundamental | Gráfico de barras múltiplas e tabela de distribuição de freqüências. Análise começa com valores extremos, passando paulatinamente para os valores centrais. Alunos identificam a moda, mas não retém que as medidas representam um conjunto e nem que existe variação em torno da tendência central. |
| Meletiou (2000)<br>Alunos de graduação | Uso de histogramas.<br>Analisam o formato das colunas (irregularidade) e as freqüências. Não analisam a amplitude da amostra. |
| Meletiou (2000)<br>Alunos de graduação | Boxplots<br>Analisaram o intervalo interquartílico |
| Makar e Confrey (2005)<br>Futuros professores de Matemática e Ciências | Gráfico de pontos<br>Ordem decrescente de termos usados: proporção, média, máximo e mínimo, tamanho da amostra, *outliers*, amplitude, formato da distribuição e desvio padrão.<br>Na primeira entrevista, nenhum professor usou o termo desvio padrão e na segunda entrevista, somente dois professores.<br>Referem-se à variação como o formato da distribuição ou como pedaços da distribuição. |
| Hammermam e Rubin (2004)<br>Professores | Semelhante ao Gráfico de Pontos<br>Não usaram medidas de centro para comparar os dois grupos.<br>Dividiram o conjunto de dados em duas partes (a partir de uma medida de contexto) ou em mais partes. |

A questão que emerge destes resultados é: se os participantes das pesquisas não sentiram necessidade de uma medida de tendência central, como poderiam sentir necessidade de uma medida de variação? Ou seja, perceberam a variação, tentaram diminuí-la, mas não buscaram sintetizá-la numa medida. Isto reforça o tipo de estratégia utilizada, pois o resultado é semelhante quando os participantes analisavam uma distribuição (Quadro 13).

Isto pode estar indicando que o conceito de variação ainda não está estável para os sujeitos de pesquisa, independente da faixa etária.

Os únicos estudos encontrados especificamente sobre o raciocínio dos participantes com o desvio padrão estão apresentados no Quadro 16.

Os estudos também avaliavam o raciocínio dos participantes quando comparavam duas distribuições, mas tinham o objetivo específico de trabalhar com as medidas de variação.

Quadro 16 - Comparação de duas distribuições utilizando média e desvio padrão

| Autores e Nível dos sujeitos de pesquisa | Respostas e as estratégias de raciocínio utilizadas |
|---|---|
| Loosen, Lioen e Lacante (1985)<br><br>Alunos de graduação em psicologia | Utilização de gráficos de barras, em que cada barra representa uma observação.<br><br>Alunos “vêem” a variação entre as observações e não em relação à tendência central: *Unalikability* . |
| Delmas e Liu (2005)<br><br>Alunos de graduação | Movimentar o Gráfico de barras e observar as alterações no desvio padrão.<br><br>Conceitos relacionados: valor e freqüência das observações; distância de cada valor em relação à média; distribuição espelhada tem mesmo desvio padrão; o valor do desvio padrão é independente de onde a distribuição é centrada.<br><br>Raciocínio dos alunos: barras distantes umas das outras (sem referência à média), barras igualmente espalhadas, barras distantes da média, distribuição equilibrada de valores acima e abaixo da média nos valores extremos da escala numérica, alto valor numérico da média, barras contíguas em ordem crescente ou decrescente de freqüências, mais barras (ou valores) no meio, distribuição simétrica, equilíbrio (balança), imagem espelhada da distribuição e locação |

Conforme Loosen, Lioen e Lacante (1985) salientam, os alunos continuaram a perceber a variação que existia entre as observações e não a variação em torno da medida de tendência central.

No estudo de Delmas e Liu, os autores trabalharam exaustivamente no entendimento do desvio padrão. Utilizaram o histograma para representar uma variável discreta. Como já identificado por Meletiou (2000), utilizaram uma estratégia que o aluno, geralmente, apresenta dificuldade, pois precisa perceber o valor da variável (eixo horizontal), a freqüência em cada valor (eixo vertical), a densidade de freqüência em torno da medida de tendência central para poder escolher a menor variação em torno dessa medida. Para raciocinar desta maneira, o aluno precisa mobilizar os conceitos de distribuição de freqüência e sua representação gráfica e o conceito de média.

Da maneira como foi conduzido o estudo, os alunos foram induzidos a olhar toda a distribuição, evitando que caíssem no mesmo resultado que os outros estudos: perceber a variação inerente.

Porém deve ser notado que, quando os autores comentam o mesmo valor do desvio padrão independente de onde se situa a média, ou seja, movimentando toda a distribuição ao longo do eixo horizontal, deve-se tomar o cuidado de verificar que é o valor absoluto da variação (desvio padrão) e não o valor relativo da variação (coeficiente de variação).

Alguns estudos verificaram outros aspectos de variabilidade tal como sintetizado no Quadro 17 e 18.

Quadro 17 - Variabilidade e representatividade amostral

| Autores e Nível dos sujeitos de pesquisa | Respostas e as estratégias de raciocínio utilizadas |
|---|---|
| Bakker (2004)<br>Alunos de 8ª série do ensino fundamental | Alunos perceberam que as observações ficavam muito separadas e que não permitia tirar conclusões. |
| Meletiou (2000)<br>Alunos de graduação | 1) Amostra grande; preferência por amostras aleatórias; Detectavam vieses.<br>2) Amostra pequena; em situações reais os alunos podem ser levados a analisar o contexto e superestimar o tamanho da amostra. |

É possível notar no Quadro 17 que os alunos conseguiam perceber que amostras pequenas não permitiam emitir conclusões, mas não é possível inferir que os alunos perceberam que amostras pequenas apresentavam maior variabilidade, como já questionado por Pfannkuch (2005) sobre o trabalho de Bakker (2004).

Quando trabalhadas situações que envolviam o acaso, foi possível perceber mais claramente o raciocínio de variação dos alunos, tal como apresentado no Quadro 18.

Quadro 18 – Raciocínio sobre Variabilidade em situações de acaso.

| Autores e Nível dos sujeitos de pesquisa | Respostas e as estratégias de raciocínio utilizadas |
|---|---|
| Reading e Shaughnessy (2004)<br>Alunos de 4ª série do ensino fundamental à 3º ano do ensino médio | A atividade sobre a estimativa de balas vermelhas em uma única amostra estimula o participante a apresentar raciocínio sobre as causas de variação enquanto que a atividade que solicita escolher uma lista de possíveis resultados para seis amostras leva o participante a apresentar raciocínio sobre a descrição de variação. |
| Torok (2000)<br>Alunos de 7ª e 8ª séries do ensino fundamental | Inicialmente os alunos pensam de maneira determinística. Com o aumento das repetições, alguns alunos percebem que os resultados giram em torno de um valor. |
| Meletiou (2000)<br>Alunos de graduação | Em jogos de azar - Tendência de balancear os resultados.<br>Em contexto real – Percebem que existe variação, mas não levam em consideração o tamanho da amostra. Focam em causas físicas. |
| Canada (2006)<br>Professores (em formação) das séries iniciais do Ensino Fundamental | No pré-teste, a maioria dos alunos apresentaram a proporção de casos como o resultado, sem levar em consideração a variação.<br>No pós-teste, os alunos perceberam que devido à variação inerente em situações de acaso, tanto a proporção de um set de 50 rodadas como os resultados de vários sets tendem a ficar próximos da proporção (devido à teoria), mas podem variar e que resultados muito distantes da proporção são improváveis. |

Torok (2000, p. 25) explica que variação em situações de acaso é aquela que nasce em um processo aleatório. Este autor explica, também, que variação amostral (*sampling variation)* é a idéia que repetir um experimento aleatório várias vezes (nas mesmas condições) ou tomar várias amostras aleatórias de uma população e produzir uma variedade de resultados.

Meletiou-Mavrotheris e Lee (2002) alertam para as diferentes competências requeridas para entender variação em um plano aleatório e variação em contexto real. Nesse último, o professor precisa levar em consideração a grande variedade de crenças, concepções e interpretações que os estudantes trazem para cada situação.

Baseado nos estudos de Wild e Pfannkuch (1999) e em Reading e Shaughnessy (2004), Ben-Zvi (2004) resume 6 aspectos de variação que devem ser considerados: observação e reconhecimento, medida e modelagem (para as propostas de predição, explicação ou controle), explicação e manuseio da variação, desenvolvimento de estratégias de investigação em relação variação, descrição e representação.

De acordo com esses aspectos identificados por Ben-Zvi (2004), os estudos apresentados conseguiram trabalhar com a observação e o reconhecimento da variação (quase todos), com as diferentes maneiras de representar graficamente esta variação, mas sem utilizar uma medida (exceto para o estudo de Delmas e Liu (2005) que tratava especificamente do desvio padrão).

Ressalta-se que Ben-Zvi (2004) está denominando de aspectos de variação o que Snee (1990) denominou de elementos do pensamento estatístico e que Wild e Pfannkuch (1999) denominaram de pensamento fundamental do pensamento estatístico.

Desta maneira, o próximo subcapítulo discute os diferentes aspectos de variação e sua relação com os níveis de raciocínio sobre este objeto.

## 4.6 A Relação entre os Aspectos de Variação/Variabilidade com os Níveis de Raciocínio sobre este objeto.

Como apresentado no Capítulo 1 deste trabalho, o modelo de raciocínio estatístico desenvolvido por Garfield (2002) apresenta cinco níveis crescentes (idiossincrático, verbal, de transição, de procedimento e raciocínio completo), em que um aluno só tem adquirido completamente o conceito quando apresenta o último nível de raciocínio estatístico.

Este modelo não foi utilizado pelos autores dos estudos analisados neste trabalho. Portanto, o objetivo deste subcapítulo é verificar se é possível estabelecer uma relação entre os modelos desenvolvidos pelos estudos analisados neste trabalho com o modelo de raciocínio estatístico apresentado por Garfield (2002).

A maioria dos autores tem utilizado o termo *aspecto do raciocínio sobre variação*, como, por exemplo, identificar a variação, medir a variação, variação em situação aleatória, etc., sem identificar o nível do raciocínio em cada aspecto.

Canada (2006) utilizou três situações diferentes (amostragem, dados e gráficos e situações probabilísticas) e elaborou uma estrutura conceitual para analisar o raciocínio de variação apresentado pelos alunos. Esta estrutura era composta de três aspectos: variação esperada, representação de variação e a interpretação de variação.

Segundo o próprio autor relatou, sua estrutura conceitual relacionava-se com outros modelos. Canada (2006, p. 44) explica que tem relação com o modelo de Wild e Pfannkuch (1999) no que diz respeito ao "reconhecimento, medida, modelagem e explicação de variação". Estes aspectos fazem parte do que Wild e Pfannkuch (1999) denominaram de pensamento fundamental do pensamento estatístico.

Embora o que o autor tenha designado como estrutura conceitual não se refere ao raciocínio utilizado, os resultados apresentados permitem uma compreensão do raciocínio. Por exemplo, no aspecto sobre a interpretação da variação, o aluno iniciou apresentando causas físicas e no pós-teste, atribuiu as causas devido à aleatoriedade, que significa um nível superior do raciocínio sobre as causas da variação; na percepção da variação, no pré-teste, o aluno entendia que alguma coisa poderia acontecer, mas sem saber exatamente o que. Já no pós-teste, este raciocínio apresentou um nível superior, pois os alunos já pensavam que seria possível fazer afirmações razoáveis. E assim por diante. Ou seja, embora não fosse objetivo do autor estabelecer um nível de variação em cada aspecto observado, isto poderia ter sido feito.

Diferentemente do estudo de Canada (2006), Reading e Shaughnessy (2004) elaboraram uma estrutura hierárquica do raciocínio de variação em situação de acaso que está apresentada no Quadro 19.

Quadro 19: Estrutura hierárquica do raciocínio de variação (Reading e Shaughnessy, 2004).

| Análise hierárquica da descrição de variação | Análise hierárquica da causa de variação |
|---|---|
| D1: concentra-se em valores no meio do conjunto ou valores extremos;<br>D2: concentra-se em ambos valores internos e extremos;<br>D3: Discute desvios de um valor e<br>D4: discute desvios de um valor central. | C1: identificação de causas irrelevantes de variação;<br>C2: discute freqüência de cores;<br>C3: discute proporção de cores e<br>C4: discute probabilidades baseada em proporções |

Observando o estudo de Reading e Shaughnessy (2004), é possível verificar que eles exploraram a variação em situação de acaso (aspecto de variação) e, a partir das respostas dos alunos, identificaram um nivelamento do raciocínio utilizado para descrever a variação e um nivelamento do raciocínio utilizado para explicar as causas de variação.

Reading e Shaughnessy (2004) não utilizaram o modelo de raciocínio estatístico proposto por Garfield (2002), mas criaram seu próprio nivelamento a partir da técnica denominada análise hierárquica.

No estudo de Ben-Zvi (2004), o autor observou o raciocínio de variação utilizado para analisar uma distribuição e duas distribuições, identificando se alunos estariam observando, reconhecendo, lidando intuitivamente e descrevendo a variabilidade.

O autor detectou em seu estudo sete fases de desenvolvimento do raciocínio sobre variabilidade. Diante da detecção destas fases, pode-se supor os aspectos priorizados em cada fase e relacioná-los com os níveis de raciocínio estatístico de Garfield (2002). Esta inferência sobre o trabalho de Ben-Zvi (2004) está apresentada no Quadro 20.

Quadro 20: Relação entre as fases do desenvolvimento do raciocínio sobre variabilidade (Ben-Zvi, 2004) e os níveis de raciocínio estatístico (Garfield, 2002).

| Fases do raciocínio sobre variação | Aspecto do raciocínio de variação | Nível do raciocínio estatístico |
|---|---|---|
| 1) foco em informações irrelevantes ou locais | observação | Idiossincrático |
| 2) descrição informal de variabilidade no rol de dados (entender a comparação da variável em questão) | reconhecimento | Verbal |
| 3) formular uma hipótese estatística que leva em conta a variabilidade ( a maioria de) | Lidando com a variação | De transição |
| 4) explicar a variabilidade em um tabela de distribuição de freqüência | descrição | |
| 5)uso de medidas de centro e dispersão para comparar grupos | descrição | Procedimento |
| 6) modelar variabilidade lidando com os outliers | descrição | |
| 7) observação e distinção da variabilidade dentro e entre as distribuições, a partir do gráfico. | descrição | Completo |

Garfield e Ben-Zvi (2005) fizeram uma discussão sobre os artigos publicados no *SERJ* de Novembro de 2004 e Maio de 2005 e, a partir dos resultados encontrados nos artigos, sugeriram um modelo epistemológico para quando os estudantes raciocinam sobre variabilidade na solução de problemas estatísticos. Resumidamente, o modelo é apresentado no Quadro 21.

O que os autores estão denominando de componentes do modelo epistemológico, pode ser entendido como os aspectos do raciocínio de variação. Por exemplo, o componente *Uso de variabilidade para fazer comparações*, que é um aspecto de variação investigado, pode ter diferentes respostas, com diferentes conceitos mobilizados, que permite estabelecer um nivelamento destas respostas.

Quadro 21: Síntese do modelo epistemológico desenvolvido por Garfied e Ben-Zvi (2005)

| Componentes do modelo epistemológico | Idéias chave | Avaliação e/ou verificação |
|---|---|---|
| Desenvolvimento de idéias intuitivas de variabilidade | Reconhecer que a variabilidade está em todo lugar (onipresença da variabilidade).<br>Existe variação em medidas repetidas da mesma variável e existe variação nas observações de variáveis coletadas diferentes indivíduos.<br>Variabilidade é uma entidade ao invés de pontos individuais ou uma combinação de centro e valores extremos. | Descrever variáveis como idade e altura das crianças, descrevendo a variabilidade ou o formato da distribuição.<br>Pedir aos alunos para prever a distribuição de uma variável.<br>Comparar dois ou mais gráficos e buscar razões para que um tenha medidas de variabilidade maior ou menor que o outro. |
| Descrição e representação de variabilidade | Diferentes gráficos podem revelar aspectos diferentes da variabilidade num conjunto de dados e é importante estudar mais do que um simples gráfico.<br>Uso de um número para representar a variação tais como o desvio padrão, amplitude ou o intervalo interquartílico, aliado a sua medida de centro. | Interpretar a variabilidade de uma variável que esteja representada graficamente ou numericamente.<br>Escolher entre medidas apropriadas para distribuições assimétricas (mediana e o intervalo interquartílico) e simétricas (média e desvio padrão)<br>O efeito dos outliers nas medidas de variabilidade. |
| Uso de variabilidade para fazer comparações | A partir de gráficos com a mesma escala e medidas de centro e variação ao invés de comparar dados individuais ou partes do gráfico. | Tomada de decisão diante de dois gráficos ou explicar em qual dos dois gráficos apresenta menos ou mais variabilidade. |
| Reconhecimento de variabilidade em tipos especiais de distribuições | Numa distribuição normal, a partir do conhecimento do desvio padrão e da média, é possível determinar a porcentagem de observações dentro de um, dois e três desvios padrões da média.<br>Numa distribuição bivariada, conhecer a variabilidade de y para os valores individuais de cada x.<br>A variabilidade da um conjunto bivariado de dados (covariação) pode revelar uma relação entre as variáveis. | Apresentação de média e desvio padrão de uma distribuição normal e pedir ao aluno que utilize estes dados para elaborar gráficos mostrando esta variação.<br>Apresentação de gráficos para dados bivariados e solicitar aos alunos verificar se a variabilidade de y pode ser explicada pela variabilidade de x. |
| Identificar padrões de variabilidade no ajustamento de modelos | Ajustar uma curva normal para uma distribuição de dados ou uma reta para um gráfico de dispersão de dados bivariados. | Verificar se um conjunto de dados assemelha-se a uma distribuição normal ou se uma gráfico de dados bivariados sugerem uma relação linear. |

Continuação do Quadro 21

| Componentes do modelo epistemológico | Idéias chave | Avaliação e/ou verificação |
|---|---|---|
| Uso de variabilidade para predizer amostras ou resultados aleatórios | "Amostras maiores têm mais variabilidade do que amostras menores, quando retiradas aleatoriamente de uma mesma população. Entretanto, as estatísticas de amostras maiores variam menos do que as estatísticas de amostras menores." (GARFIELD e BEN-ZVI, 2005, p. 95) | Escolher estatísticas amostrais (por exemplo, proporções) de uma população específica (por exemplo, doces coloridos) para um dado tamanho amostral e perguntar qual seqüência de estatísticas é mais provável.<br>Perguntar aos alunos qual resultado é mais provável em um experimento aleatório em que todos os resultados são igualmente prováveis. |
| Consideração da variabilidade como parte do pensamento estatístico | Na exploração de dados e na solução de problemas estatísticos: inicia-se discutindo a variabilidade nos dados; pensando na variabilidade na produção dos dados; tentando explicar a variação, procurando efeitos sistemáticos escondidos na variabilidade aleatória. Tudo isto faz parte do pensamento estatístico | Dar um problema para que os alunos investiguem com um conjunto de dados, que requeira gráficos, descrição e explicação da variabilidade.<br>Permitir ao aluno realizar os passos de uma investigação estatística, revelando se e como os alunos consideram a variabilidade dos dados. |

Segundo Garfield e Ben-Zvi (2005, p. 95), a lista de idéias cada vez mais sofisticadas oferecem: 1) as maneiras nas quais este conjunto de conhecimento pode ser estruturado para que possa ser compreendido pelo aluno; 2)uma seqüência efetiva para apresentar o material relacionado com variabilidade; 3) um planejamento para re-visitar variabilidade como um progresso do aluno no currículo estatístico e 4) um suporte para construir novos níveis de entendimento profundo de variabilidade. E os autores trabalham com uma perspectiva construtivista de aprendizagem, em que a construção de significado não é linear, mas complexa e melhor entendida numa imagem de progressão espiral.

Como os próprios autores salientam, o Quadro 21 apresenta uma lista de situações (aspectos de variação) em que pode ser explorado o raciocínio sobre este objeto. E, certamente, quanto mais situações um aluno vivenciar, mais elevado é o nível de seu raciocínio sobre variação.

> Um entendimento completo de variabilidade significa desenvolver um modelo cognitivo que inclui vários componentes e suas conexões, de maneira a usar este modelo para raciocinar sobre variabilidade em diferentes contextos. (GARFIELD e BEN-ZVI, 2005, p. 93 )

Segundo Pfannkuch (2005 p.90) raciocínio sobre variação envolve o questionamento dos dados e requer uma relação recíproca entre o contexto e a variação nos dados; envolve detectar padrões através da conexão recíproca entre a medida de tendência central e a variação. Esta explicação de Pfannkuch é essencial no sentido que o contexto da problemática estatística deve fazer sentido para o aprendiz, deve ter significado para ele e principalmente deve promover uma vontade de interpretar os resultados.

Essa retomada dos aspectos de raciocínio sobre variação desenvolvidos pelos autores foi elaborada com o objetivo de facilitar a comparação com o modelo de raciocínio de variação apresentada neste trabalho e justificar a problemática apresentada no capítulo seguinte.

## 5 Descrição da problemática da pesquisa

Como já salientado anteriormente, o nível de raciocínio estatístico torna-se mais avançado à medida que o indivíduo tem condições de vivenciar diferentes situações, ou seja, trabalhar diferentes aspectos do raciocínio sobre o conteúdo estatístico alvo.

Os estudos de Reading (2004), Canada (2006) e Reading e Shaughnessy (2004) elaboraram uma estrutura conceitual para nivelar o raciocínio sobre variação, definido por esses dois últimos autores como o processo cognitivo envolvido na descrição do fenômeno observado em situações em que haja variabilidade. Utilizando o modelo de nivelamento do raciocínio estatístico proposto por Garfield (2002), o problema principal desta pesquisa é: **Qual o nível de raciocínio de variação utilizado pelo professor de Matemática em diferentes etapas do ciclo investigativo do pensamento estatístico?**

Este trabalho pretende investigar o raciocínio sobre variação utilizado para resolver problemas estatísticos que emergiram durante uma investigação idealizada e realizada pelos próprios professores participantes da pesquisa, ou seja, em diferentes aspectos de variação.

Utilizando o modelo de pensamento estatístico elaborado por Wild e Pfannkuch (1999) e o modelo de raciocínio estatístico apresentado por Garfield (2002), o problema principal é dividido em subproblemas de pesquisa.

Na fase de problema e planejamento de um ciclo investigativo, as questões propostas são: **os professores de Matemática processam o pensamento fundamental do pensamento estatístico (Wild e Pfannkuch, 1999)? Se o fazem, qual o nível de raciocínio de variação apresentado pelo professor nesta etapa do ciclo investigativo?**

Na fase de análise de dados do ciclo investigativo existem vários subproblemas de pesquisa, em que cada um relaciona-se com algumas etapas do modelo epistemológico desenvolvido por Garfield e Ben-Zvi (2005), cujo objetivo foi identificar o nível de cada aspecto de variação.

Uma primeira situação apresentada aos professores de Matemática foi a elaboração de representações gráficas de diferentes variáveis, apresentadas num banco de dados, em que eles poderiam analisá-las separadamente ou elaborar grupos e comparar as distribuições. Se os professores elaboraram a representação de cada variável, independentemente, indaga-se: **qual o nível de raciocínio de variação dos professores de Matemática quando elaboram uma representação gráfica de uma distribuição? Eles percebem a necessidade de uma medida de tendência central e uma medida de variação?**

Se os professores elaboraram mais de uma representação a partir da criação de grupos, as perguntas são: **qual o nível do raciocínio sobre variação apresentado pelos professores quando comparam distribuições? Estes professores de Matemática analisam a variação entre as observações (*unalikeability*) ou em torno de uma medida de tendência central?**

Quando se estimula a utilização das medidas de tendência central, pergunta-se: **há o surgimento natural da necessidade de uma medida de variação? Qual o nível de raciocínio de variação apresentado pelos professores quando eles foram estimulados a utilizar as medidas de tendência central para representar uma distribuição?**

Quando estes professores interpretaram a medida de variação obtida, a indagação é: **existe uma tendência para olhar a homogeneidade, aspecto reforçado pelos livros didáticos, ou conseguem perceber a variação em torno da média? Qual o nível de raciocínio de variação quando interpretam o desvio padrão?**

Não foi aplicada nenhuma atividade que trabalhasse o contexto de variação em situação de acaso, e por este motivo, não é apresentado nenhum problema acerca deste aspecto de variação.

## 6 Metodologia

Esta pesquisa seguiu os pressupostos de uma pesquisa-ação que Tripp (2005, p. 447) definiu como “uma forma de investigação-ação que utiliza técnicas de pesquisa consagradas para informar a ação que se decide tomar para melhorar a prática.”

Para este autor, investigação-ação apresenta o ciclo demonstrado na Figura 51.

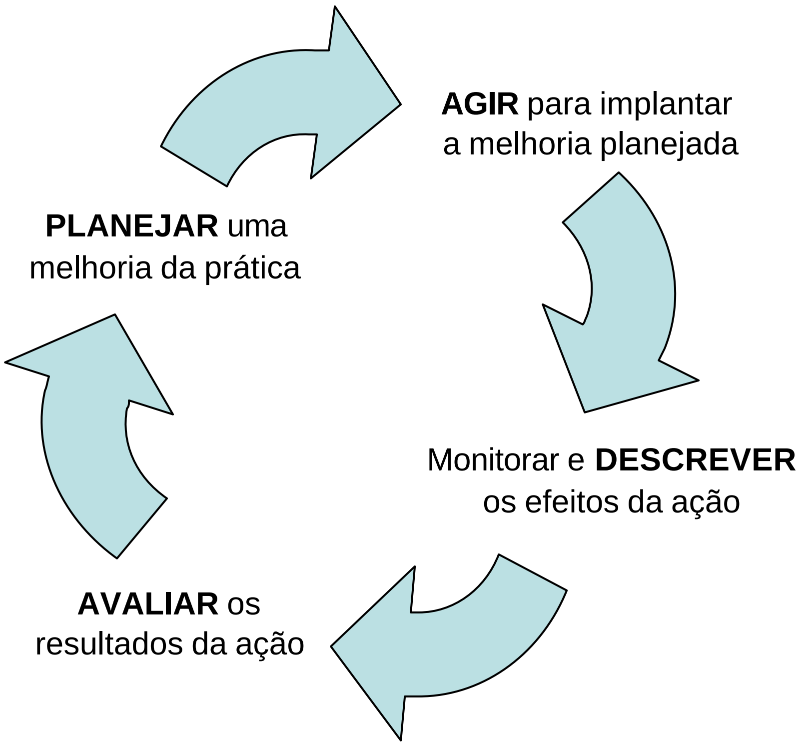



Figura 51 - Ciclo básico da investigação-ação, de acordo com Tripp (2005)

Como as quatro fases do ciclo da investigação-ação dizem respeito à uma ação, Tripp (2005 p. 453) explica que o ciclo da pesquisa-ação pode ser resumido da seguinte maneira: planejamento (o que vai ser feito num primeiro encontro e um planejamento dos outros encontros a partir dos resultados da etapa anterior), implementação (produção de dados) e avaliação ( da mudança da prática e do processo de investigação-ação).

Seguindo estes princípios, a partir da revisão bibliográfica realizada, do conhecimento do perfil dos participantes da pesquisa e do conteúdo estatístico de interesse destes participantes, foi feito um planejamento inicial da ação. Segundo

Tripp (2005), a pesquisa-ação começa com um reconhecimento do contexto, das práticas atuais, dos participantes e envolvidos.

Este planejamento inicial foi implementado em 04 de Março de 2005 (primeiro encontro com os participantes), em que foram observados os resultados e foi feito um planejamento para o novo encontro com os participantes. Os encontros foram semanais (toda sexta-feira, das 8h00 as 11h00) e, após cada encontro, repetia-se a análise dos resultados obtidos e planejava-se a implementação para o próximo encontro. Este procedimento é explicado como iteratividade da pesquisa-ação (TRIPP, 2005).

Na fase de planejamento, as ações foram pensadas e discutidas ora com a orientadora da pesquisa ora com o grupo de pesquisa. Esta reflexão conjunta permite caracterizar esta pesquisa como participativa no grupo de pesquisa.

A fase de implementação foi conduzida de 04 de Março de 2005 à 26 de Agosto do mesmo ano, totalizando 16 encontros, 48 horas de formação continuada.

A fase de avaliação aconteceu com a orientadora e/ou o grupo de pesquisa, em que foram observados os resultados e foram planejadas as novas ações. Como explica Franco (2005 p. 491), o método da pesquisa-ação deve "contemplar o exercício contínuo de espirais cíclicas: planejamento; ação; reflexão; pesquisa; ressignificação; replanejamento, ações cada vez mais ajustadas às necessidades coletivas, reflexões, e assim por diante".

Como a pesquisa-ação envolve a pesquisa e a ação, este trabalho teve como ação uma formação continuada sob a temática Estatística, por solicitação dos próprios participantes, e teve como pesquisa a verificação do nível de raciocínio sobre variação utilizado para resolver os problemas estatísticos emergentes em um ciclo investigativo.

Considerando que o maior interesse da pesquisadora era verificar o raciocínio dos participantes, esta pesquisa-ação pode ser considerada como ação-pesquisa, que segundo Barbier (2004, p. 41-42) é um dos tipos de pesquisa-ação e refere-se "às pesquisas utilizadas e concebidas como meio de favorecer mudanças intencionais decididas pelo pesquisador". Segundo Tripp (2005, p. 452) uma ação-pesquisa é quando se prioriza mais o conhecimento obtido do que o aprimoramento da prática.

Entendendo que este trabalho seguiu os pressupostos de uma pesquisa-ação, a seguir está descrito o perfil dos participantes, o papel do observador, da formadora e dos próprios participantes, o contexto e cronograma da intervenção e a forma de análise dos resultados.

## 6.1 Participantes da Pesquisa

Foram participantes da pesquisa nove professores do ensino fundamental e médio e dois alunos do curso de Matemática da Universidade de São Paulo.

Quatro professores e um aluno participaram das três pesquisas antecedentes a esta, cujas temáticas foram fração, função e equação, respectivamente das autoras Silva (2005), Rossini (2006) e Lima [2007].

Em 2001, estes quatro professores haviam solicitado ao Professor Dr. Saddo Ag Almouloud, coordenador do grupo de pesquisa, uma formação em Probabilidade e Estatística, haja vista a necessidade de lecionarem este conteúdo em suas aulas de Matemática.

No primeiro encontro desta pesquisa, todos os participantes receberam uma carta de esclarecimento sobre o desenvolvimento da pesquisa e duas vias do Termo de Consentimento Livre e Esclarecido - TCLE (Apêndice 3)[29], que foram lidos coletivamente. Após a leitura e explicação sobre a necessidade deste procedimento (seguindo as recomendações do Comitê de Ética em Pesquisa da Universidade São Judas Tadeu, que avaliou eticamente este trabalho) foi solicitado aos professores que concordassem em participar, que assinassem uma via do TCLE e entregasse à pesquisadora.

Também no primeiro encontro foi solicitado que respondessem a um questionário (Apêndice 2) para que fosse possível conhecer o perfil destes professores e sua experiência com a Estatística.

O grupo era composto por dois homens e sete mulheres. O professor mais novo tinha 26 anos e o mais velho tinha 59 anos. Todos os professores eram graduados em Matemática exceto uma professora, cuja formação era em Química. Todos ministravam aula no Ensino Médio e tinham uma carga horária semanal alta, como pode ser observado no Quadro 22.

---

[29] Protocolo aprovado com recomendação pelo Comitê de Ética da Universidade São Judas Tadeu, sob registro 0141.0.000.219-05

Quadro 22 - Perfil dos professores participantes da pesquisa

| Participante * | Formação | Disciplina que leciona | Séries em que leciona** | Número de horas-aula semanais |
|---|---|---|---|---|
| LH<br>59 anos | Matemática | Matemática e Física | 1º, 2º e 3º anos do E.M. | 18 |
| AM<br>39 anos | Matemática | Matemática | 5ª, 6ª e 7ª E.F.<br>1º, 2º e 3º do E.M. | 47 |
| OB<br>49 anos | Ciências e Matemática | Matemática | 8ª do E.F.<br>1º e 2º do E.M. | 32 |
| RN<br>50 anos | a) Arquitetura e Urbanismo e<br>b) Matemática | Matemática | 7ª do E.F.<br>1º e 2º do E.M.<br>3º E.J.A. | 43 |
| MA<br>39 anos | Química | Matemática para E.F.<br>Física e Química para EM | 6ª do E.F.<br>1º , 2º e 3º do E.M. | 17 |
| SB<br>29 anos | Matemática | Matemática | 5ª e 8ª do E.F.<br>7ª E.J.A.<br>3º do E.M. | 45 |
| LF<br>37 anos | Ciências e Matemática | Matemática | 1º e 2º do E.M. | 28 |
| IS<br>26 anos | Matemática | Matemática | 6ª e 7ª do E.F.<br>1º do E.M. | 30 |
| RS<br>44 anos | Ciências e Matemática | Matemática | 5ª a 8ª E.F.<br>5ª a 8ª E.J.A.<br>2º e 3º E.M. | 48 |

* Foram utilizadas as letras iniciais dos nomes dos professores

** E.F. - Ensino Fundamental; E.M. – Ensino Médio – E.J.A.– Ensino de Jovens e Adultos

A carga horária semanal da maioria dos professores é muito extensa e muitos verbalizaram que só tinham um período livre na semana (a manhã da sexta-feira), e que dedicavam à formação continuada. Talvez esse fosse um dos motivos para o grande número de ausências, como pode ser observado na Tabela 23.

Tabela 23 - Freqüência dos professores à formação continuada (em 16 encontros)

| Professor | LH | AM | OB | RN | MA | SB | LF | IS | RS |
|---|---|---|---|---|---|---|---|---|---|
| Freqüência ao projeto | 16 | 14 | 14 | 14 | 13 | 10 | 09 | 08 | 08 |

Considerando que a formação continuada teve 16 encontros, é possível observar na Tabela 23 que três professores tiveram uma freqüência muito baixa nos encontros (cerca de 50% de freqüência). A professora LF parou de participar do grupo de pesquisa, pois ingressou no curso de Mestrado em Educação Matemática da própria PUC-SP. A professora IS era a mais jovem do grupo e justificava suas faltas pela necessidade de utilizar este horário para elaborar suas aulas. A professora RS foi convidada pela professora RN a integrar o grupo e começou a participar da pesquisa após um mês de seu início.

Para identificar a experiência prévia desses professores com conteúdos estatísticos, foi perguntado se estavam trabalhando ou já tinham trabalhado algum conteúdo de estatística em suas aulas e o Quadro 23 apresenta esses resultados.

Quadro 23 - Conteúdo estatístico já trabalhado em sala de aula

| Participante | Conteúdo estatístico trabalhado |
|---|---|
| LH | --- |
| AM | Freqüência, gráfico, média e desvio padrão |
| OB | --- |
| RN | Levantamento de dados, organização em gráficos e tabelas |
| MA | Construção de gráficos de barras e setores |
| SB | Leitura de gráficos e probabilidade |
| LF | Levantamento de dados, porcentagem, tabelas e gráficos (leitura e interpretação) |
| IS | --- |
| RS | Pesquisa, tabela, gráficos, análise de dados e porcentagem |

É importante observar no Quadro 23 que três professores nunca ensinaram nenhum conteúdo estatístico e que os que já tinham experiência, em sua maioria, trabalharam com gráficos. Somente o professor AM já tinha trabalhado com medidas de tendência central e dispersão, restringindo-se à média e ao desvio padrão.

Além de conhecer a experiência prévia desses professores com o conteúdo estatístico, buscou-se saber se eles conheciam as orientações dos Parâmetros Curriculares Nacionais e os resultados estão apresentados na Tabela 24.

Tabela 24 - Número de participantes que conheciam (ou não) os PCNs

| Parâmetros Curriculares | Conhece | Não Conhece | Conhece em parte |
|---|---|---|---|
| Do Ensino Fundamental | 6 | 0 | 3 |
| Do Ensino Médio | 6 | 0 | 3 |
| PCN+ | 2 | 6 | 1 |

Somente dois professores disseram conhecer o PCN+. Ou seja, somente dois professores conheciam os conteúdos e metodologias sugeridas pelos parâmetros curriculares para o trabalho com conteúdos estatísticos no ensino médio. É possível que utilizassem o livro didático como referência, mas isto não foi perguntado a eles.

Quando perguntado se utilizavam os PCNs para elaborar suas aulas, quatro professores disseram que usavam e dois professores responderam que usavam de vez em quando.

Além desses nove professores, participaram da pesquisa dois alunos de graduação em Matemática da Universidade de São Paulo. A aluna de Matemática (AG – 18 anos) já estava participando do grupo de pesquisa desde 2004, convidada pelo professor AM, que tinha sido seu professor de Matemática durante o ensino médio. Após quatro encontros dessa pesquisa, a aluna pediu permissão para que seu namorado (LJ – 19 anos) pudesse participar. A freqüência deles ao projeto foi de 11 encontros (aluna) e 8 encontros (o aluno). Ambos não tinham experiência como professores de Matemática.

Além dos nove professores e dos dois alunos, dois outros professores participaram de apenas um encontro e depois não compareceram mais por incompatibilidade de horário. Sua participação foi considerada na análise dos resultados, mas não foi possível identificar o nível de raciocínio de variação destes professores.

## 6.2 Ambiente da implementação da ação

Franco (2005) explica que a realização da pesquisa-ação deve ser feita em ambientes onde acontecem as próprias práticas, o ambiente natural, e isto não foi possível acontecer, impossibilitando seguir rigorosamente os princípios da pesquisa-ação.

Por decisão dos pesquisadores do grupo de pesquisa e dos participantes da formação continuada, em 2004 as atividades passaram a acontecer em uma escola pública do município de Arujá, onde se desenvolveu a pesquisa de Silva (2005). Devido às dificuldades encontradas pelos participantes, eles solicitaram que os encontros voltassem a acontecer na universidade sede do grupo de pesquisa. Logo, por decisão dos próprios participantes, esta pesquisa foi realizada no Centro de Ciências Exatas e Tecnologias da PUC-SP, campus Marquês de Paranaguá, e não na escola em que estes professores lecionavam.

Toda a implementação da ação aconteceu no laboratório de Matemática, em que tinha disponível lousa, giz, gravadores e computadores com software *OpenOffice*. Alguns encontros foram feitos no laboratório de informática deste centro devido ao fato de os computadores terem o Microsoft Excel.

Além dos encontros presenciais, foi disponibilizado um fórum virtual. Todos os participantes foram cadastrados e estimulados a escrever suas dúvidas ou ansiedades que pudessem acontecer durante a semana. Embora tenha sido estimulado este meio de comunicação, dois ou três professores fizeram uso no início do trabalho e logo abandonaram. Por este motivo, os resultados desta comunicação não serviram para analise dos dados.

## 6.3 Papel do Observador – observação participante periférica

Em todos os encontros havia, pelo menos, três observadores. Estes observadores eram pesquisadores do grupo de pesquisa e que decidiram pela participação como observadores ou pela necessidade de se integrar aos participantes ou para compreender o tipo de pesquisa desenvolvido no grupo ou mesmo para contribuir com a pesquisadora.

As observadoras VG e RL eram alunas de doutorado e professoras da PUC-SP. Elas participavam do grupo de pesquisa e acompanhavam os participantes desde o início dos trabalhos (em 2000). Por este motivo, a participação delas era fundamental para que o pesquisador coletivo existisse. Segundo Barbier (2004), o pesquisador coletivo é quando os membros envolvidos na pesquisa têm vontade de resolver o problema, quando os participantes tornam-se ativos e aliados ao pesquisador.

A observadora MM era aluna de mestrado da PUC-SP e estava interessada em desenvolver sua pesquisa com a temática Estatística e por este motivo solicitou sua participação como observadora. O mesmo aconteceu com LL, porém com menor envolvimento.

A observadora AA era aluna de graduação de matemática da PUC-SP e orientanda de iniciação científica de RL.

Estes observadores assumiram o papel de observação participante periférica (OPP) que, segundo Barbier (2004), é quando o pesquisador aceita uma implicação parcial para poder ser considerado como membro sem, entretanto, ser admitido no centro das atividades do grupo.

As observadoras VG e RL anotavam os acontecimentos e faziam intervenções esporádicas, ora questionando o tipo de desenvolvimento da atividade ora propondo reflexões acerca do assunto. Estas intervenções foram muito importantes para o entendimento do raciocínio utilizado pelos participantes.

As observadoras MM, LL e AA assumiram um papel passivo. Apenas anotavam todo o desenvolvimento das atividades.

Antes de cada encontro, a formadora explicava aos observadores o objetivo das atividades. Quando os participantes dividiam-se em grupos para a realização das atividades solicitadas, cada observador acompanhava um grupo, para que fosse possível ter registro de todos os debates.

## 6.4 Papel do Formador – observação participante completa

A formação foi feita pela própria autora. Para que pudesse ser aceita pelo grupo de participantes, a autora assumiu o papel de observação participante periférica na pesquisa de Silva (2005), cujas atividades aconteceram em 2003 e 2004. Mesmo utilizando esta estratégia de aproximação com o grupo, no início das atividades a formadora era uma pessoa estranha ao grupo. Tal como lembra

Franco (2005), o pesquisador deve perceber que está lidando com um grupo, de alguma forma estruturado, que possui uma dinâmica própria e que ele, pesquisador, de início, não faz parte desse grupo.

Como formadora, a autora assumiu o papel de observação participante completa (OPC) em que, segundo Barbier (2004 p. 126), “o pesquisador ou está implicado desde o início, porque já era membro do grupo antes de começar a pesquisa; ou ele se torna membro do grupo por conversão, porque provém de fora”. Neste caso, houve um pouco dos dois motivos. Não é possível considerar que a autora fazia parte do grupo, mas não se pode dizer que era estranha ao grupo.

O papel da formadora era o de propiciar situações que motivassem os participantes a se engajar no projeto e que permitisse a eles tirar suas próprias conclusões e verificá-las no momento do *feedback*. Segundo Franco (2005) o papel do pesquisador em uma pesquisa-ação é de facilitador, só devendo intervir quando houver necessidade.

Segundo Barbier (2004, p. 18), em uma pesquisa-ação o pesquisador desempenha “seu papel profissional numa dialética que articula constantemente a implicação e o distanciamento, a afetividade e a racionalidade, o simbólico e o imaginário, a mediação e o desafio, a autoformação e a heteroformação, a ciência e a arte”.

Assim, a autora navegou entre o papel de formadora e pesquisadora, intervindo ou mantendo distância para observar o tipo de desenvolvimento, ora questionando-se sobre os benefícios da formação ora questionando-se sobre os resultados da pesquisa. Este papel é intrigante e Franco (2005, p. 492) coloca que a grande questão é a da necessária interpenetração de papéis: “como passar de pesquisador a participante, continuando a ser prioritariamente pesquisador, pois o pesquisador estará, por certo, prioritariamente envolvido na pesquisa e nos resultados desta”.

Uma das preocupações e intenções da formadora era propiciar uma ação comunicativa que permitisse aos participantes agir efetivamente, questionando sobre o conteúdo, refletindo sobre sua prática e permitindo à pesquisadora entender o tipo de raciocínio que utilizavam.

> Uma ação comunicativa é uma ação eminentemente interativa, nasce do coletivo, da equipe. Essa ação não pretende garantir a eficiência a qualquer custo, não é individualista, não persegue o êxito, mas, ao contrário, é uma ação dialógica, vitalista, que emerge do mundo vivido. Essa ação nasce da situação e lhe oferece saídas. É comunitária, busca o entendimento, persegue a negociação, o acordo; busca o consenso; é axiológica porque acredita na validade das normas discutidas. Mansa na escuta e forte na tomada de decisão. (ROJO APUD FRANCO, 2005 p. 492)

Esta ação comunicativa só passa a acontecer quando o clima de confiança está estabelecido, pois os questionamentos feitos pelos professores podem ser entendidos como os não-entendimentos dos professores acerca de um conteúdo que ele deveria conhecer e isto pode bloquear o bom desenvolvimento da pesquisa-ação.

É preciso que o pesquisador saiba construir esse sentimento de parceria, construindo um clima de grupo que permita a livre participação dos sujeitos. Segundo Franco (2005), torna-se mais fácil estabelecer este clima cooperativo quando o grupo solicita o trabalho aos pesquisadores, mas que nem sempre isto é vontade de todos. Nesta pesquisa, não é possível dizer que isto era vontade de todos, mesmo porque seis participantes eram recém chegados ao grupo de formação continuada.

O objetivo principal da formadora era “produzir nos sujeitos envolvimento, participação, comprometimento e produção de saberes” e principalmente “produzir conhecimentos novos a serem incorporados no campo científico”, que Franco (2005, p. 497) explica que são os objetivos de uma pesquisa-ação.

## 6.5 O papel dos participantes – atores da pesquisa-ação

Diferentemente do trabalho de Silva (2005), em que os professores participantes assumiam o papel de conhecedores do assunto “fração”, neste trabalho eles verbalizavam e afirmavam não conhecer Estatística e, por este motivo, haviam solicitado uma formação sobre este conteúdo. Eles assumiram, desde o início desta formação, o papel de alunos.

Esta postura assumida pelos participantes facilitou o estabelecimento da ação comunicativa. Como já salientado anteriormente, nem todos os participantes haviam solicitado uma formação em Estatística e alguns deles eram novos no grupo. Por este motivo, o momento de integração foi essencial nesta pesquisa-ação.

Os participantes LH, AM e OB, solicitantes da formação, participaram ativamente de todos os encontros, perguntando, expondo suas dificuldades e buscando soluções para as situações vividas em sala de aula. As participantes LF, CI e RN, embora novas no grupo, eram muito extrovertidas e envolvidas com o trabalho, participando ativamente das discussões e buscando alternativas.

As participantes IS e RS falavam muito baixo, participavam pouco, dando a impressão que preferiam não se expor. Esta postura foi se alterando paulatinamente.

Mas uma atenção especial deve ser dada à participação de SB. Ela já participava do grupo desde 2003 e lecionava na mesma escola que AM. Ela prestava atenção em todo o debate das sessões e podia ser observado que estava o tempo todo refletindo sobre sua prática. Muitas vezes foi solicitado que ela verbalizasse o que estava pensando, pois ela falava baixo apenas para AM. E sempre que era possível ouvir suas reflexões, as discussões eram muito ricas e suas colocações muito pertinentes. Vale registrar que ela também estava grávida e a criança nasceu no final de Junho de 2005, o que a impossibilitou de participar de alguns encontros.

Segundo Franco (2005), uma das ações esperadas pelos sujeitos da pesquisa é que participe ativamente da elaboração da problemática da ação, da pesquisa, da busca de soluções, enfim de todas as etapas, mas, como pode ser observado na descrição acima, isto aconteceu com intensidade diferente com cada participante.

Franco (2005) ainda coloca uma questão acerca da participação do professor como participante da pesquisa-ação: como passar de professor, sujeito da pesquisa, a pesquisador de seu fazer, mantendo-se prioritariamente no papel de professor? Embora este papel seja difícil, era o esperado pela formadora.

Vale ressaltar que os professores participaram ativamente das decisões da ação, ficando a cargo da autora e do grupo de pesquisa as decisões do curso da própria pesquisa. Assim, entende-se que este trabalho seguiu os pressupostos de uma pesquisa-ação e não se realizou uma propriamente dita, pois segundo Franco (2005), os participantes devem colaborar nas tomadas de decisão, tanto nas questões de pesquisa quanto nas questões da ação.

## 6.6 Coleta de dados e análise dos resultados

Todas as sessões foram áudio-gravadas e registradas pelos observadores. As gravações foram transcritas e as observações foram digitadas, o que permitiu o acesso aos dados coletados.

Também foram obtidos dados para análise através de materiais desenvolvidos pelos professores e entregues pessoalmente ou por e-mail.

A partir destes dados obtidos, foi feita a interpretação do nível do raciocínio sobre variação apresentado em todas as etapas do ciclo investigativo. A análise interpretativa dos resultados, que segundo Franco (2005), é o que configura uma pesquisa-ação, foi elaborada a partir das falas dos professores e, por este motivo, tornou a transcrição das fitas o material mais importante.

Foram analisadas as discussões quando os professores realizavam as atividades solicitadas em pequenos grupos e, também, no momento do *feedback*. Segundo Barbier (2004, p. 55), a análise e interpretação dos dados são produtos de discussões de grupo. Isso exige uma linguagem acessível a todos. "O traço principal da pesquisa-ação – *o feedback* – impõe a comunicação dos resultados da investigação aos membros nela envolvidos, objetivando a análise de suas reações".

## 6.7 Organização das Atividades desenvolvidas na fase de implementação da pesquisa-ação

Os professores haviam solicitado uma formação continuada em Estatística devido à necessidade de ensinar seu conteúdo, como sugerido pelos Parâmetros Curriculares Nacionais (BRASIL, 1998; BRASIL, 2000 e BRASIL, 2002).

Nos Parâmetros Curriculares Nacionais para o terceiro e quarto ciclos do ensino fundamental (BRASIL, 1998), a Estatística faz parte do bloco de conteúdo denominado Tratamento de Informação e tem como objetivo:

> a finalidade é fazer com que o aluno venha a construir procedimentos para coletar, organizar, comunicar dados, utilizando tabelas, gráficos e representações que aparecem freqüentemente em seu dia-a-dia. Além disso, calcular algumas medidas estatísticas como a média, mediana e moda com o objetivo de fornecer novos elementos para interpretar dados estatísticos. (BRASIL, 1998, p. 52)

Com relação ao ensino médio, o PCN+ (Brasil, 2002) caracteriza a Estatística no bloco de conteúdos denominado Análise de Dados e as sugestões estão apresentadas no Quadro 24.

Quadro 24 - Sugestões do PCN+ para o trabalho de Estatística no ensino médio brasileiro

| Série | Conteúdos sugeridos de estatística | Explicação do PCN+ (2002 p. 127)* |
|---|---|---|
| Primeira | Descrição de dados | "identificar formas adequadas para descrever e representar dados numéricos e informações de natureza social, econômica, política, científico-tecnológica ou abstrata" |
| | Representações gráficas | "Ler e interpretar dados e informações de caráter estatístico apresentados em diferentes linguagens e representações, na mídia ou em outros textos e meios de comunicação" |
| Segunda | Análise de dados: médias, moda e mediana, variância e desvio padrão | "Obter médias e avaliar desvios de conjuntos de dados ou informações de diferentes naturezas"<br>"Compreender e emitir juízos sobre informações estatísticas de natureza social, econômica, política ou científica apresentadas em textos, notícias, propagandas, censos, pesquisas e outros meios" |

* A explicação por sub unidade não foi especificada pelo PCN+ , mas foi uma interpretação da autora.

Estes conteúdos foram discutidos na formação continuada, que foi implementada de maneira semelhante ao trabalho de Meletiou (2000), em que a variação era o tema central de toda a formação.

Como uma pesquisa-ação, embora houvesse um planejamento inicial da implementação, cada etapa foi repensada e re-elaborada de acordo com os resultados obtidos na etapa anterior.

Para verificar como os professores atribuíam significado à Estatística e como os conceitos relacionados à variabilidade e/ou variação faziam parte desta significação, foi feito um diagnóstico no primeiro encontro (04 de Março de 2005).

Diante dos resultados obtidos no diagnóstico, foi desenvolvida a fase de sensibilização da pesquisa-ação, em que foram discutidos textos para responder a algumas indagações dos professores. Esta fase aconteceu nos dias 11 e 18 de Março.

As discussões dos textos permitiram identificar que alguns professores já trabalhavam conteúdos estatísticos por meio da realização de pesquisa, mas as análises restringiam-se à representação gráfica.

Com o objetivo de ampliar esta análise, ou seja, utilizar outros conteúdos estatísticos para analisar dados, foi decidido pela realização de uma pesquisa, cujo cronograma do desenvolvimento das atividades está apresentado no Quadro 25.

Quadro 25 – Cronograma das atividades do ciclo investigativo

| Data | atividade | Objetivo da ação |
|---|---|---|
| 01 e 08 de Abril | Planejamento e desenvolvimento de pesquisa | Desenvolver o Pensamento Geral do Pensamento Estatístico (Wild e Pfannkuch (1999) |
| 15 e 29 de Abril | Distribuição de freqüência simples e suas representações gráficas | a) Discutir as diferentes maneiras de representar graficamente uma mesma distribuição de freqüências;<br>b) Apresentar a distribuição de freqüências conjuntas;<br>c) Analisar as distribuições. |
| 06 e 13 de Maio | Organização do banco de dados da pesquisa | Discutir a elaboração de um banco de dados, com o auxílio de uma planilha eletrônica. |
| 20 de Maio e 03 de Junho | Distribuição de freqüência com dados agrupados | a) Discutir e interpretar a representação gráfica deste tipo de distribuição;<br>b) Apresentar outras maneiras de representação gráfica de variáveis contínuas. |
| 10, 17 e 24 de Junho | Medidas de Tendência central e dispersão | a) discutir diferentes medidas de tendência central;<br>b) proporcionar condições do surgimento natural de uma medida de variação em torno da média aritmética;<br>c) Compreender o conceito de desvio padrão. |

Embora o objetivo da ação tenha sido proporcionar condições de desenvolvimento do pensamento estatístico, o objetivo da pesquisa era verificar o nível de raciocínio sobre variação destes professores enquanto elaboravam, realizavam e analisavam dados de uma pesquisa.

Em algumas situações, a relação dialética entre a pesquisa e a ação não teve um equilíbrio adequado e, por este motivo, no dia 19 e 26 de Agosto de 2005 foi realizada uma atividade com estes professores para verificar o nível do raciocínio sobre variação em diferentes situações, de maneira que pudesse complementar os resultados já obtidos.

## 7 Análise do Diagnóstico

A primeira atividade desenvolvida na fase de implementação desta pesquisa-ação foi um diagnóstico sobre os primeiros significados de Estatística que estes professores apresentavam. Para realizá-lo, foi solicitada a tarefa que está apresentada na Figura 52.

| |
|---|
| PSQ: “Vamos pensar em palavras que estejam relacionadas com a Estatística. Eu vou escrever na lousa as palavras que vocês falarem. Em grupo, vocês vão organizar estas palavras em categorias e nomeá-las. A partir das categorias, vocês vão elaborar uma frase para representar *Estatística* utilizando as categorias que vocês encontraram e vão pensar numa representação para esta frase”. |

Figura 52 - Tarefa 1 do diagnóstico

A indicação das palavras, bem como a associação que fariam com elas, permitiria atingir o objetivo do diagnóstico, que era verificar se o conceito de variação e/ou variabilidade fazia parte do arcabouço de Estatística destes professores.

### 7.1 As palavras indicadas

Os professores começaram a indicar palavras que eles consideravam relacionadas com a Estatística, conforme apresentado no Quadro 26.

Quadro 26 - palavras relacionadas à Estatística, indicadas pelos professores*.

| | | | | |
|---|---|---|---|---|
| pesquisa | futuro | força interna | regra de três | previsão |
| álgebra | organizar | média | interpretar | amostra |
| representar | desconto | número | tabela | discreto |
| contínuo | consumo | fatos | sociedade | opinião pública |
| política | IBGE | censo | modelo | taxa de mortalidade |
| taxa de natalidade | variável | incógnita | cálculo | dados |
| analisar | informática | determinar | solução | estimativa |
| experimentar | experiência | freqüência | relativo | absoluto |
| Desvio padrão | chato | calculadora | aprendizado | |

* As palavras estão em ordem (por linha) de aparecimento.

Quando a palavra estimativa (35ª) foi pronunciada, eles disseram que já era o suficiente. Pode-se supor que, neste primeiro momento, o conjunto de palavras constantes no quadro já relatavam o que eles entendiam como Estatística, mesmo sem o surgimento de nenhuma palavra referente à variação. Ainda assim, o debate com a pesquisadora os incentivou a continuar na busca de novas palavras do seu repertório. Vale ressaltar que o desvio padrão apareceu depois deste episódio (42ª palavra pronunciada).

As palavras que expressavam afetividade em relação à Estatística só apareceram depois da indicação do desvio padrão, devido à intervenção do professor AM, como pode ser observado no trecho da discussão a seguir.

> AM: É, eu tô imaginando, não tá dando raiva, não tá dando medo, será o quê? Ninguém tá habituado?
>
> OB: Chato.
>
> AM: Não, mais é..., eu não acho.
>
> OB: Então, vamos começar pelos nossos sentimentos.... calculadora.
>
> RN: Sentimento como pessoa ou como professor? A gente não vai ensinar isso daí?

A primeira interpretação deste diálogo está ligada às atitudes vinculadas à prática docente que, no que tange a este tema, mostrou-se inexistente. A segunda interpretação é que o desvio padrão tenha despertado sentimentos, possivelmente negativos, em relação à Estatística, pois as palavras verbalizadas foram chato, medo e raiva. Ambas interpretações só poderiam ser confirmadas ou refutadas ao longo da formação continuada.

Mesmo no início do diagnóstico já era possível elaborar uma primeira hipótese: os professores participantes provavelmente se colocariam na condição de aluno, dispostos a aprender um conteúdo para elaborar suas aulas.

Dá-se continuidade ao processo de indicação das palavras.

Quadro 27 – Continuação da indicação das palavras

| | | | | |
|---|---|---|---|---|
| expectativa | Jornal | revista | tv | ilusão |
| época eleitoral | representante | dúvidas | ansiedade | gráfico |
| grupo | trabalho | empresa | profissão | entrevista |
| saúde | educação | lazer | probabilidade | Matemática |
| possibilidade | altura | medida | medir | situação |

Quando a professora OB pronunciou a palavra ilusão, o professor AM verbalizou:

> AM: É, na época eleitoral!
>
> OB: Analfabetismo no Brasil, a estatística diz que não existe analfabeto no Brasil! Quase, né?

Este episódio permite uma segunda hipótese: os professores não confiavam nas estatísticas veiculadas em mídias, o que precisaria ser discutido durante a formação continuada.

Para realizar a tarefa de organizar as palavras em categorias e elaborar uma frase que pudesse expressar como eles davam significado à Estatística, os professores dividiram-se em três grupos de três pessoas cada e os resultados da discussão interna do grupo, da elaboração da frase e da apresentação estão a seguir.

## 7.2 Resultados do Grupo 1

O grupo, formado pelos professores AM e LH e pela aluna AG, apresentou a seguinte frase: "Um planejamento consciente, com o auxílio de <u>ferramentas</u> e elementos matemáticos, levando também em consideração os sentimentos envolvidos, propõe um processo estatístico eficiente" e a representação desta frase, que fora elaborada por eles está apresentada na Figura 53 (grifo nosso).

A frase apresentava termos gerais como planejamento consciente, ferramentas e processo estatístico eficiente, o que não permitiu verificar como os professores percebiam e apreendiam a variabilidade e/ou variação. Isto foi explorado no momento do *feedback*, que está apresentado no subcapítulo 7.5.

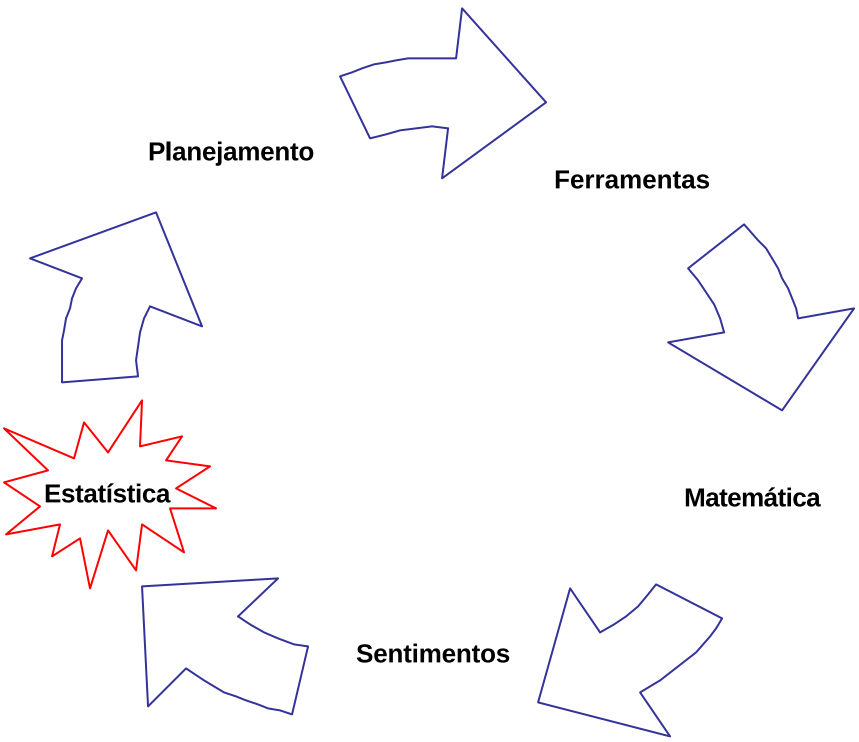



Figura 53 – Representação da significação de Estatística do Grupo 1.

A maneira como os professores distribuíram as palavras nas categorias pôde dar uma idéia do que eles entendiam como Estatística.

A categoria Matemática englobava gráficos, média, porcentagem, tabela e conteúdos matemáticos enquanto que a categoria Planejamento continha outros conteúdos estatísticos como freqüência e desvio padrão.

Esta divisão do conteúdo estatístico na categoria Matemática e na categoria Planejamento permite inferir que os professores não tinham clareza sobre os conteúdos pertencentes à Estatística. No momento do *feedback* foi questionado o que eles entendiam como Planejamento, inclusive para compreender como eles estavam interpretando o desvio padrão.

### 7.3 Resultados do Grupo 2

O segundo grupo era composto pelas professoras RN, IS e LF. Elas elaboraram a seguinte frase: “Com alguns procedimentos e cálculos podemos analisar fatos, e com sentimento interferir no campo de atuação da sociedade”. A frase foi construída a partir das seguintes categorias: Campo de atuação, Informações, Sentimentos, Procedimento e Cálculo e a representação dela está na Figura 54.

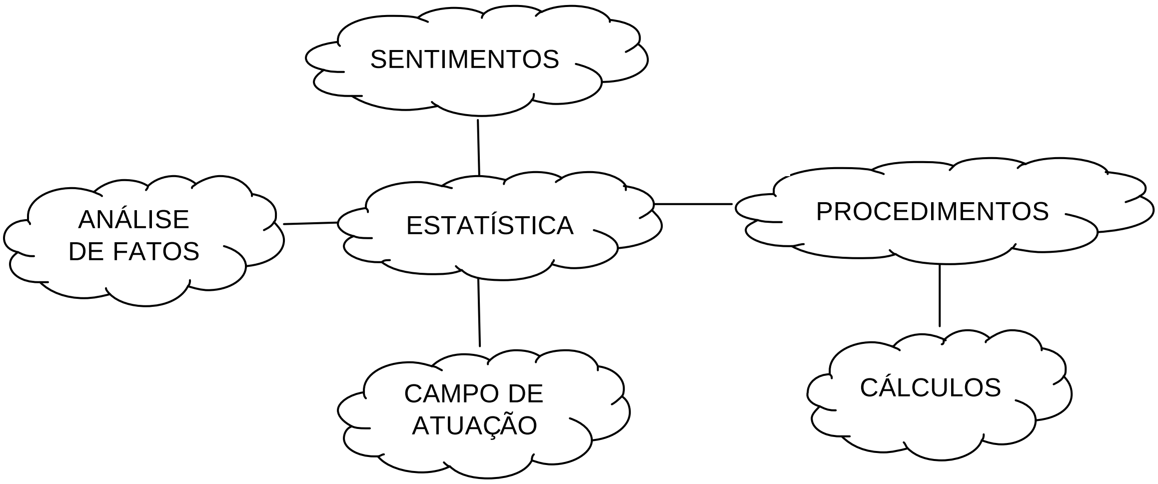



Figura 54 - Representação da significação de Estatística do Grupo 2.

A categoria Cálculo envolvia conteúdos matemáticos e estatísticos e a categoria Procedimento envolvia as ações que seriam realizadas com os cálculos tais como organizar, analisar, agrupar, representar, etc.

Diante da apresentação oral do grupo foi possível verificar que a Estatística era compreendida como uma ferramenta para analisar fatos. Provavelmente isto se devia à participação da professora RN que já realizava pesquisa com seus alunos para ensinar gráficos estatísticos.

### 7.4 Resultados do Grupo 3

O terceiro grupo, composto pelas professoras OB, SB e CI apresentou a seguinte frase: “Para o aprendizado de Estatística precisamos buscar fontes, coletar dados de diversos meios e com a utilização de <u>ferramentas</u> chegar ao nosso objetivo. A interação entre alunos e professores é importante para sanar as dúvidas e expectativas que aparecem como incógnitas, que são as dificuldades nos cálculos”, cuja representação foi reproduzida na Figura 55 (grifo nosso).

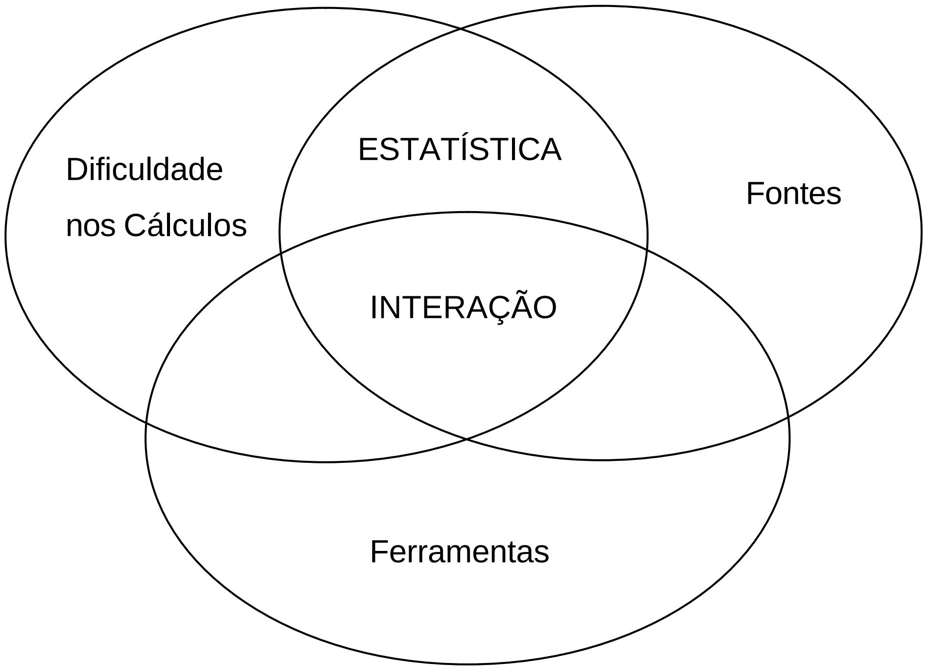



Figura 55 - Representação da significação de Estatística do Grupo 3.

As professoras verbalizaram que "chegar ao objetivo" era o aprendizado de Estatística e que, para isto, elas buscariam fontes, coletariam dados e utilizariam ferramentas (que era a própria Estatística utilizada para analisar os dados).

Isto mostra certa confusão na significação da Estatística, vista como sinônimo de Ferramenta e compreendida na intersecção entre Fontes e Dificuldade nos Cálculos. Pode-se inferir que as professoras deste grupo estavam entendendo a Estatística como um conteúdo que se usa para analisar dados de pesquisa, mas que elas tinham dificuldades nos cálculos. E isto pode ser devido à forte influência da professora OB que já havia verbalizado isto.

Deve-se ressaltar que as Ferramentas estatísticas (categoria) que seriam usadas para analisar os dados só tinham gráficos e tabelas.

Uma outra categoria era Dificuldade nos cálculos, que englobava todos os conteúdos matemáticos e outros conteúdos estatísticos. As Fontes, uma outra categoria, foi colocada ora como local onde os dados seriam obtidos ora como a aplicação da estatística. E fora mencionado, diversas vezes, que o professor teria o papel de sanar as dificuldades nos cálculos, que pertencia à categoria Interação.

Este terceiro grupo foi o único que relacionou Estatística com o processo ensino-aprendizagem, porém o foco não estava na discussão dos conceitos estatísticos com seus alunos, mas provavelmente sanar as próprias dificuldades do professor com os cálculos.

## 7.5 *Feedback*

Após a apresentação das frases, a pesquisadora, que está identificada por PSQ, fez intervenções para compreender o significado de Estatística atribuído pelos professores e verificar se havia alguma identificação de variação e/ou variabilidade neste significado. Este momento é considerado como a fase de *feedback* do primeiro encontro da pesquisa-ação.

Enquanto os grupos estavam elaborando as categorias e as frases, a PSQ observou alguns pontos que mereceriam atenção e discussão: partição do conteúdo estatístico em diferentes categorias, diversidade do significado de ferramentas, a Estatística como meio de manipulação de massa e a Estatística como objeto de ensino e interação.

A discussão sobre a distribuição dos conteúdos estatísticos em diferentes categorias está parcialmente reproduzida a seguir.

> PSQ: Todos os grupos colocaram direta ou indiretamente a dificuldade nos cálculos. O que vocês entendem com isto?
>
> RN: É bem complicado. Você consegue agitar bastante a sala, porque você pode trabalhar com jornal, com revista, alguma coisa assim, com material diferenciado, vai trabalhar em grupo, deixa de ser um trabalho individual e você estimula eles a fazerem pesquisa, então essa parte estimula bastante. Agora, quando você trabalha com... é... principalmente os livros didáticos, aí você cai nisso daí. Tem sempre números bonitinhos, né? Só na pesquisa, você não vai conseguir os números bonitinhos. Esse é o grande nó na hora de ensinar estatística.

A professora RN estava se referindo à não possibilidade de obter 100% quando efetuava a soma da porcentagem de cada categoria de resposta de uma pesquisa.

> OB: eles não sabem nem usar a regra de três. Eles não conseguem achar a porcentagem. Média, a gente tá tirando a média da nota deles e quando a gente fala em média eles não fazem nenhuma associação. O desvio padrão, eles não entendem o que é desvio padrão e é isto que derruba a gente nos concursos. A curva, a bendita curva! Puxa, não alcancei! Não passei!
>
> AM: Ou então, a curva está baixa, o nível tá baixo.

> OB: Ou então, olha, por um pouquinho eu não fico na curva. A calculadora, os alunos só sabem fazer mais ou menos. Tem função na calculadora que eu não conheço. Probabilidade, meu Deus! É um caos. Então, são estas as dificuldades.

Embora nem todos os professores tenham se pronunciado, ficou claro que haviam pedido a formação continuada em Estatística devido à necessidade de sanar suas próprias dificuldades com o conteúdo. A discussão permitiu identificar algumas dessas dificuldades: cálculo de porcentagem em uma dstribuição de freqüências, à não compreensão de média aritmética, a interpretação do desvio padrão e o cálculo de probabilidades.

A PSQ explicou que os conteúdos probabilísticos seriam desenvolvidos por outro pesquisador e que, na discussão sobre Estatística seria possível observar que há mais de uma maneira de apresentar um resultado. Esta expressão da PSQ levou o professor LH a se pronunciar: “como não é exato, não é Matemática”.

A afirmação do professor LH despertou um longo debate no grupo. Parte dos professores disse que Estatística era Matemática, pois estava nos livros didáticos enquanto que outra parte dos professores disse que usava Matemática e fizeram a analogia com a confecção de um bolo (usa matemática, mas não é matemática). O *feedback* a este respeito foi feito no encontro seguinte, por meio do texto de Silva e Coutinho (2005)[30].

Quanto à categoria ferramentas, o objetivo era verificar se estavam entendendo a Estatística como uma ferramenta de pesquisa ou se estavam levantando as ferramentas necessárias para ensinar Estatística.

> PSQ: Duas frases utilizam a categoria ferramentas. O que vocês entendem como ferramentas?
>
> AM (G1): Excel, ....., jornais, tudo o que se usa para conseguir fazer um estudo estatístico.
>
> PSQ: são as ferramentas que vocês precisam para ensinar estatística?
>
> AM: Isto mesmo.
>
> PSQ: No Grupo 2, procedimentos são as ferramentas?
>
> RN: Procedimentos que você precisa para levantar dados, organizar, calcular, desenhar gráficos, mais nesse sentido.

---

[30] Este texto não está disponível nos Apêndices deste trabalho devido ao fato de estar disponível no endereço ftp://ftp.usjt.br/pub/revint/191_41. ZIP

> PSQ: Procedimentos necessários para fazer estatística?
>
> RN: Isto.

Vale destacar que, neste diálogo, os professores responderam às questões da PSQ, o que não significa que era o que eles efetivamente tinham como representação, já que o pronunciamento não foi espontâneo. Portanto, é uma inferência da autora que o primeiro grupo entendia ferramenta como o material necessário para ensinar Estatística enquanto que o segundo grupo estava compreendendo a Estatística como uma ferramenta de análise de dados, o mesmo já verbalizado pelo terceiro grupo.

Apenas o Grupo 3 mencionou a relação professor-aluno, embora a preocupação estivesse centrada na solução de suas próprias dificuldades. A discussão sobre esta Interação (categoria) fez surgir o papel da Estatística como meio de integrar as disciplinas na escola quando do desenvolvimento de projetos.

Outro papel da Estatística explorado foi sua utilização para manipulação. Foram levantadas duas hipóteses acerca da compreensão dos professores sobre essa temática: Estatística é falsa ou a Estatística pode ser usada para obter os resultados desejados.

> PSQ: Vocês verbalizaram, nas discussões, sobre a ilusão da Estatística, manipulação de dados.
>
> AM: O sentimento que me causa é muita raiva, muito ódio quando vai falar de política. Tem sempre dados manipulados. Nunca é a verdade. O candidato sempre puxa para seu lado.
>
> PSQ: Seus alunos percebem a manipulação na Estatística?
>
> AM: Meus alunos não enxergam esta manipulação.
>
> PSQ: Isto não vai incomodar você na hora de dar aula?
>
> AM: Isto sempre aparece! Não tem jeito!

Mas, em um outro momento da discussão das frases, o professor AM relatou que “A estatística não é mentirosa. Depende de quem a usa, de quem a faz. A estatística não é mentirosa. É quem a faz, quem a usa que é mentiroso”, que se relaciona com o que ele havia denominado de processo estatístico eficiente: “sempre buscar a verdade dos fatos que estão sendo analisados”.

A fala deste professor permite confirmar, parcialmente, a segunda hipótese da autora sobre a interpretação da Estatística como objeto de manipulação. Como os outros professores não se pronunciaram, não foi possível verificar sua compreensão a este respeito.

## 7.6 Análise do Raciocínio sobre variação/variabilidade identificado no diagnóstico

O diagnóstico permitiu identificar pontos frágeis na significação de Estatística para estes professores. As frases refletiam o uso de pesquisa para ensinar estatística, mas o conteúdo que possivelmente era explorado restringia-se aos gráficos e tabelas.

Existia, junto a estes professores, uma concepção de Estatística como a ferramenta para enganar as pessoas e surgiu a preocupação de quanto esta concepção poderia afastá-los de ensinar este conteúdo.

A professora SB citou o desvio padrão na indicação das palavras e a professora OB disse que ele é o motivo do fracasso nos concursos. Porém, este conceito estatístico não foi apropriado pelo grupo, não surgindo nem como ferramenta nem como conteúdo na significação da Estatística. Por tudo o que foi apresentado, pode-se classificar o nível de raciocínio sobre variação como idiossincrático, pois apenas verbalizaram o desvio padrão sem levar em consideração sua aplicação como ferramenta de pesquisa ou como conteúdo a ser ensinado.

## 8 Fase de sensibilização da pesquisa ação.

O diagnóstico apontou dois aspectos que precisariam ser discutidos: os conteúdos e estratégias metodológicas sugeridas para o trabalho com a Estatística desde a 5ª série do ensino fundamental até o 3º ano do ensino médio e a própria definição de Estatística e sua relação com a Matemática.

Para trabalhar o primeiro aspecto, no dia 11 de Março de 2005 foram discutidas as sugestões dos Parâmetros Curriculares Nacionais do ensino fundamental e médio.

> PSQ: Como vocês usam os PCNs na escola?
>
> AM: Ninguém usa. Vou ser sincero! Eu uso para explicar alguma coisa que eu quero.
>
> LF: Para fazer projetos, pois lá tem explicação.
>
> OB: Para fazer planejamento no início do ano.

Como uns professores afirmaram não lembrar o que estava escrito nos Parâmetros Curriculares Nacionais (tanto para o Ensino Fundamental – PCN-EF – como para o Ensino Médio – PCN-EM e PCN+) e outros professores admitiram que nunca leram esses documentos, se dividiram em três grupos para fazer uma leitura rápida das sugestões para o ensino de Estatística.

- Grupo 1 – Professora IS e o professor EVE (que só participou deste encontro e que está identificado como eventual) leram os PCNs para 5ª e 6ª séries;
- Grupo 2 – professores LF e OB leram os PCNs para 7ª e 8ª séries;
- Grupo 3 – professores AM, CI, LH , RN e a aluna AG leram o PCN-EM e o PCN+.

A tarefa solicitada está apresentada na Figura 56.

| |
|---|
| PSQ: Que conteúdos os PCNs sugerem (para a Estatística) e que propostas de trabalho eles fazem para estes conteúdos? |

Figura 56 - Tarefa 1 da sensibilização

Cada grupo fez uma leitura de, aproximadamente, uma hora, após o que fizeram um relato dos pontos que consideraram importantes. A discussão e a apresentação de cada grupo estão sintetizadas nos subcapítulos 8.1 a 8.3, as questões que foram formalizadas estão apresentadas no subcapítulo 8.4 e a síntese da análise do raciocínio de variação/variabilidade estão apresentadas no subcapítulo 8.5.

Na apresentação dos resultados, algumas frases foram sublinhadas por serem consideradas como uma evidência do raciocínio do professor.

## 8.1 Resultados do Grupo 1

O grupo relatou que os conteúdos sugeridos para as 5ª e 6ª séries do ensino fundamental são média aritmética e espaço amostral, enfatizando o significado para o aluno e evitando o trabalho com termos e fórmulas.

Os professores consideraram importante o fato de trabalhar com pesquisa, principalmente sem a imposição do tema, mas apresentaram dúvidas quanto à possibilidade de sua efetivação, seja pela possibilidade de algazarra dos alunos, seja pela não compreensão e autorização de outros membros da escola para que isto fosse feito fora da sala de aula.

O professor EVE relatou que estilo musical é uma temática de interesse dos alunos, embora os PCNs sugerissem idade e altura. Após a realização da pesquisa, os alunos calculariam a média e colocariam em tabelas e gráficos, permitindo leituras e interpretação.

A afirmação do professor EVE sobre o ensino de média aritmética de estilos musicais requer a escolha de um estilo, por exemplo rock, e considerar 1 para quem selecionou-o e considerar 0 para quem selecionou outros estilos. Levanta-se a hipótese de que o professor não havia pensado nesta estratégia e, possivelmente, teria a intenção de calcular a média de uma variável qualitativa com mais de duas possibilidades de resposta.

## 8.2 Resultados do Grupo 2

A observadora LL explicou ao grupo que o termo Tratamento da Informação referia-se ao bloco de conteúdos da Estatística, da Probabilidade e da Análise Combinatória dos PCNs.

Após a compreensão do termo, a professora LF explicou que trabalhava Estatística de duas formas diferentes, sendo uma delas junto com juros simples, no momento de organizar em tabelas os preços das mercadorias observados em três estabelecimentos e por meio da interpretação de gráficos veiculados em jornais e revistas.

A professora LF pretendia trabalhar os conteúdos matemáticos de maneira interligada, permitindo ao aluno revisitar o que já aprendeu e utilizar este conhecimento para facilitar a nova aprendizagem. Entretanto, sua verbalização sobre a utilização de tabela para apresentar os preços de três estabelecimentos, indica que não estava trabalhando com distribuição de freqüência, que foi confirmado após a intervenção da observadora LL.

> Observadora LL: Você consegue atingir todo o conteúdo sugerido pelo PCN no bloco de Tratamento de informação?
>
> LF: Eu chego até um certo ponto. Eu sempre trabalhei com pesquisa, coleta de dados, mas eu nunca me importei como tratar isto.

No momento do relato sobre a leitura dos PCNs, as professoras verbalizaram que a estratégia de trabalho sugerida para a Estatística é a mesma já apresentada pelo Grupo 1 e que a Estatística pode ser usada para retomar outros conhecimentos matemáticos já aprendidos, o que pode tornar menos cansativa a aprendizagem e ajudá-lo a compreender fatos do cotidiano. A pesquisadora fez uma intervenção:

> PSQ: E quais os conteúdos sugeridos pelo PCN?
>
> OB: Razão, porcentagem, regra de três, <u>juros simples</u>, freqüência, organização dos dados, né? Primeiramente, a leitura e interpretação de dados expressos em gráficos, construções de recursos visuais, né? Principalmente para apresentar globalmente os dados, destacar aspectos relevantes, sintetizando as informações, permitindo a elaboração de inferências (...) freqüências, tem as medidas, que é a média, a mediana, né? e a moda, né? E a compreensão de seu significado também, né? A construção do espaço amostral, utilizando o principio multiplicativo e a iniciação da probabilidade por meio da razão. E a elaboração de experimentos e simulação para estimar probabilidades e verificar probabilidades previstas.
>
> PSQ: E a estratégia a ser utilizada?
>
> OB: Fazer pesquisa que tenha interesse para o aluno, né? Que, como ele [professor EVE] falou, o peso, a altura, a idade, que isto tem significado para ele.

A professora OB entendeu que a professora LF trabalhava juros simples como um conteúdo estatístico e não que utilizasse a Estatística quando trabalhava matemática financeira. Além disso, parte de sua verbalização foi a leitura dos conteúdos sugeridos pelos PCNs.

### 8.3 Resultados do Grupo 3

Este grupo prendeu-se mais em troca de experiências do que na leitura dos PCNs. Poucos trechos foram realmente lidos e discutidos.

A professora RN explicou que faz pesquisa com os alunos sobre algum tema que os próprios alunos sugerem, por exemplo, sabor do sorvete. Monta-se o rol das respostas. Organiza-se a tabela (7ª e 8ª séries do ensino fundamental).

O professor AM explicou que divide os alunos do ensino médio em grupos, em que cada um escolhe o tema a ser pesquisado, faz dez perguntas e escreve uma sinopse sobre como, onde e com quem fizeram a pesquisa. Quando os alunos voltam com os resultados, o professor AM conduz a maneira de analisar os resultados, com a freqüência de cada resposta. Para cada questão, os alunos fazem um gráfico diferente, com o uso do computador. E os alunos fazem "no lápis" o que o computador fez, para pensar o conteúdo.

Ambos professores que relataram sua experiência já trabalhavam a Estatística da maneira como os PCNs sugeriam, porém, o conteúdo estatístico discutido restringia-se a distribuição de freqüências e a elaboração de gráficos.

A leitura do PCN+ motivou o professor AM a questionar sobre o desvio padrão.

> AM: Você trabalha com o desvio padrão? [perguntando ao grupo] Eu me lembro de ter sido tão massacrado nesse negócio! Fazia, fazia, e...
> Observadora RL: O que é desvio padrão?
> AM: Eu digo assim. Vamos ver se eu não estou falando besteira. Quando eu falo em desvio padrão, por exemplo, eu dou exemplo de uma sala de aula, vai. Eu pego a nota de todos os alunos e saio comparando. <u>Quanto mais próximo todas as notas de cada aluno analisada for próximo ao zero, significa que eles estão, a sala está em uma hegemonia. Ou estão aprendendo todos iguais ou nenhum está aprendendo igual. Então é para isto.</u> Coisa que ninguém pensa nisto dentro da escola. Pelo menos, eu nunca vi nenhuma escola, nenhuma coordenação que se preocupe em chegar no final do ano e fazer uma avaliação desta. Que eu acho que é o papel do coordenador.

A verbalização de AM permite duas interpretações. A primeira é que, possivelmente, este professor estivesse comparando as notas com a média zero da Distribuição Normal Reduzida e o fato de que as notas estariam próximas do zero, próximas da média, haveria um desvio padrão menor. Se este fosse seu raciocínio, poderia ser considerado num nível de transição (GARFIELD, 2002), pois utilizou uma medida de tendência central e percebeu a densidade de freqüência em torno da média, duas dimensões do conceito de desvio padrão.

Uma segunda interpretação da verbalização do professor AM, mais provável, é que ele estivesse se recordando da apostila que utilizava no ensino médio, que colocava como resposta de um exercício a seguinte explicação: "quanto mais próximo de zero é o desvio padrão, mais homogênea é a distribuição dos valores da variável", que poderia ser classificado como um nível verbal de variação, pois está correta a relação entre o valor do desvio padrão e a homogeneidade das observações, mas isto não permite perceber que as observações são mais homogêneas em relação à um ponto de referência: a média.

Porém, a verbalização do professor AM foi referente às notas mais próximas do zero e não o desvio padrão mais próximo do zero, o que pode permitir uma interpretação de raciocínio idiossincrático de variação, pois ele conhece algumas palavras, mas ainda não consegue relacioná-las adequadamente, ou seja, o conceito não está formado.

Durante a leitura, uma afirmação do PCN+ (Brasil, 2002, p. 126) chama a atenção do professor AM: "A Estatística e a Probabilidade devem ser vistas, então, como um conjunto de idéias e procedimentos que permitem aplicar a Matemática em questões do mundo real". O professor AM entendeu como uma resposta ao que havia sido discutido na sessão anterior e verbalizou: "o PCN coloca a Estatística como um ramo da Matemática", mas quando o grupo apresentou a definição de Estatística e de Probabilidade, não houve destaque para a interpretação do professor AM.

Durante a explicação do grupo sobre as estratégias de ensino utilizadas pelos professores AM e RN, os conteúdos estatísticos do ensino médio não foram especificados, que levou o professor EVE a questionar e LH respondeu: "variância e desvio padrão".

## 8.4 *Feedback*

Como apresentado no início deste capítulo, a fase de sensibilização foi marcada pelas discussões sobre o conteúdo e metodologia sugeridos pelos PCNs e sobre a definição de Estatística e sua relação com a Matemática.

A leitura do PCN-EF, PCN-EM e PCN+ limitou-se àquela feita na sessão presencial e, portanto, foi feita de maneira superficial.

Na fase de planejamento da próxima ação (iteratividade da pesquisa-ação), o grupo de pesquisa sugeriu que fossem utilizados outros textos para dar o *feedback* dos assuntos levantados. Foi feita uma síntese do Capítulo 5 de Ponte, Brocardo e Oliveira (2003), que está no Apêndice 4 deste trabalho, para discutir o aspecto metodológico do ensino de Estatística e foi utilizado um texto introdutório sobre história da Estatística e da Probabilidade ( SILVA e COUTINHO, 2005) para discutir as diferentes definições de Estatística conforme o contexto em que está inserida.

Como a discussão só aconteceu no próximo encontro, foi solicitado aos professores que aprofundassem a leitura dos PCNs e fizessem uma busca na Internet sobre a história da Estatística.

O *feedback* foi no dia 18 de Março de 2005, com a retomada da definição de Estatística utilizada pela professora RN: “Estatística trata de um conjunto de métodos utilizados para obtenção de dados, sua organização em tabelas, gráficos e análise destes dados” e fez-se a leitura da definição de Estatística elaborada por Costa Neto (2002, p.1): “Estatística é a ciência que se preocupa com a organização, descrição, analise e interpretação de dados experimentais” que suscitou a discussão sobre duas visões da Estatística: é um ramo da Matemática e é uma ciência independente que usa a Matemática como ferramenta.

O texto de Silva e Coutinho (2005) foi lido parcialmente, dando ênfase no surgimento da Estatística com os recenseamentos, no surgimento da probabilidade com os jogos e o possível momento de integração destes dois campos. A afirmação do professor AM fez surgir novamente a discussão dos PCNs.

> AM: Eu sempre olhei para a Estatística e a Probabilidade como duas coisas separadas.
>
> PSQ: Elas são separadas? Como estão nos PCNs?

> AM: Para mim eram separadas, pois na apostila que eu uso, são capítulos separados e não são linkados (...) Eu tenho a impressão, quando eu pego o conteúdo, que a abordagem é muito resumida, em Estatística. Em Probabilidades, abre, né! Aquele monte de exercícios. Agora, a Estatística fica resumido a média e desvio. Eu faço meia dúzia de exercícios e pronto.

A pesquisadora explicou que esta visão da Estatística separada da Probabilidade pode ser em conseqüência de que o conteúdo da Estatística sugerido para a escola básica restrinja-se à estatística descritiva. Somente quando se dá início à inferência estatística, é que se torna possível observar a interligação dos dois campos.

Neste momento, a pesquisadora solicitou uma reflexão aos professores, apresentada na Figura 57.

| |
|---|
| PSQ: Depois da leitura dos PCNs, se vocês fossem trabalhar Estatística na série em que lecionam, como vocês fariam? |

Figura 57 – Questão 1 para reflexão durante a fase de sensibilização

> RN: Com jornais e revistas para os alunos interpretarem gráficos e isto agita bem a sala.
> AM: Você não precisa falar para ele (aluno) o que é freqüência relativa, cai a ficha e depois você só dá o nome. Você não precisa falar para ele (aluno) o que é desvio padrão. Lá no final, cai a ficha. O que é variância, cai a ficha. Depois você só dá o nome. Ele chega e você não precisa abordar nada. Fica até fácil você dar aula. Depois a resolução dos exercícios, aí é outra estória, né?
> LH: Fica próximo do cotidiano deles.

Utilizando a verbalização do professor AM, a PSQ lembrou a todos que os conteúdos estatísticos não se restringem apenas aos gráficos e tabelas, como os grupos apresentaram. Existem também as medidas de tendência central e dispersão, que quase não foram citadas. Neste momento, a professora OB faz uma intervenção:

> OB: Vocês lembram quando nós fomos fazer o trabalho da Renata [Rossini (2006), que trabalhou com função], foi pedido para eles desenharem gráfico. Que gráfico eles desenharam?
> AM: Ah é!
> OB: Gráfico estatístico. Né! Fizeram gráfico de setor e barras.

A ênfase no ensino de gráficos estatísticos parece não ser privilégio dos professores deste estudo, levando alunos de outros professores a associar gráficos com a Estatística e não com a Função. Este fato parece sugerir que a Estatística está diretamente relacionada com gráficos da mesma maneira que Estatística estava relacionada com *médiamodamediana* no estudo de Bakker (2004).

A discussão voltou-se para o ensino de Estatística e deu-se início à leitura completa da síntese do texto de Ponte, Brocardo e Oliveira (2003). O primeiro parágrafo gerou estranheza por parte dos professores, que não entendiam como as aulas de Estatística poderiam ser chatas.

> AM: Engraçado. Diferente do que a gente discutiu. [se referindo ao aborrecimento em aulas de estatística].
>
> Observadora RL: Leia a frase anterior: "as abordagens usuais deste tópico enfatizam os aspectos computacionais e procedimentais". É só cálculo. Calcula a média, calcula isto, calcula aquilo.
>
> OB: Lembra do professor da universidade? Como era chata aquela aula de Estatística dele. Nossa!
>
> PSQ: E como era a aula dele?
>
> OB: Ele chegava lá e: Hoje vamos estudar desvio padrão. Ele colocava o conceito e depois dava aqueles exercícios...a gente pegava a calculadora e ficava lá e fazia.

A professora OB associou a aula chata com desvio padrão e com a realização de cálculos, o que permite inferir que a aula "legal" é aquela em que é feita uma pesquisa e são elaborados gráficos.

> AM: A maior raiva que eu tenho de química. Fiz dp [dependência] desta matéria. Por quê? Porque chegava no laboratório, né, fazia aquele monte de experiência. Depois calculava a média de não sei o que, desvio padrão, lá, lá, lá aquele monte de coisa. No final, você entregava o relatório e eu não sabia o que eu tinha feito. Ah, tá, tá totalmente certo [o professor dizia]. É, que bom! Eu não sabia o que eu tinha feito.

O professor AM relatou que não entendia desvio padrão quando estava cursando a graduação, mas não se posicionou sobre sua compreensão na atualidade.

> OB: Eu fui para exame em Estatística e Calculo Diferencial. Eu estudei Cálculo pra caramba. Decorei o caderno de Cálculo e Estatística, é só fórmula mesmo, né? Era só levar a calculadora e ele deixava ainda, né? Ô matéria nojenta, chata! Tirei 9,75 de Cálculo e de Estatística, ele

> começou a corrigir e olhar para minha cara, né? Daí ele começou a colocar certo, pois ele viu que eu ia desmaiar, né? [Todos riem]. Era muito chato. Eu não conseguia estudar estatística do jeito que ele queria.
>
> PSQ: É isto que o Ponte está colocando aqui, né? Se fica só na fórmula, o aluno vai perguntar: para que eu preciso deste monte de cálculo? Lembram-se do trabalho da Zezé [Silva, 2005]: a palavra forte era significado, né? É isto! E aqui nós queremos discutir o significado destas medidas.

Ao término da leitura da síntese, a PSQ tinha o objetivo de estimular a percepção de variabilidade e fez a intervenção apresentada na Figura 58.

| |
|---|
| PSQ: Vamos utilizar o exemplo do Ponte. Se a gente pegar o perfil do aluno. Se o ET fosse ler o relatório dos alunos, <u>ele teria um retrato fiel da realidade</u>? |

Figura 58 - Questão 2 para reflexão durante a fase de sensibilização

> LH: Ele teria um retrato do que ele observou. Se é fiel, ai é que está!.. é o que ele observou.
>
> OB: Depende de como ele ia interpretar aqueles dados.
>
> LH: Se ele vai ver a cor dos alunos, <u>quanto mais informações ele pegar daquele assunto, daquilo que ele está analisando, mais preciso está.</u> Aí acho que está o pensamento cuidadoso. Quanto mais dados, informações ele tem de alguma coisa, mais próximo ele está da realidade.

A precisão relatada pelo professor LH não permite compreender à que ele se referia e isto não foi questionado no momento da formação continuada. Entretanto, no contexto em que está inserida, sugere que quanto mais pessoas ele estivesse observando, ele teria uma observação da maioria com maior fidelidade, o que permitiria classificar seu raciocínio como verbal, possivelmente observando a moda para representar a variação das observações. Vale ressaltar que este foi o raciocínio predominante do professor LH durante a formação continuada.

A pesquisadora percebeu que usou um termo inadequado (retrato fiel) que poderia reforçar a idéia de falsa Estatística. Então, fez uma nova intervenção, apresentada na Figura 59.

| |
|---|
| PSQ: Vamos pensar naqueles caras da ONU. Eles pegam aqueles relatórios da realidade do Brasil. Então, eles vão lá e pegam o relatório do Brasil que diz que a renda média do brasileiro, o número médio de filhos da mulher brasileira e eles tem que enxergar o Brasil diante daquele relatório. É o mesmo relatório que foi dado na mão do ET. Então, alguém de fora pega um relatório da realidade brasileira. Como esta pessoa deve ler este relatório? Como ela deve interpretar este relatório? |

Figura 59 - Questão 3 para reflexão durante a fase de sensibilização

> LH: Ela tem que saber o que foi pesquisado.
>
> OB: Espaço amostral, também né? Digamos que pesquisaram num estado para lá, Nordeste, né? E ele vem <u>para São Paulo, é uma realidade completamente diferente.</u>

Embora a professora OB tenha utilizado o termo espaço amostral, sua verbalização permite inferir que se referia à amostragem, a uma possível generalização de resultados obtidos em uma localidade para outra localidade. Mas, também pode ser entendido que estava observando a variação de um estado brasileiro para outro, o que permite categorizar como um raciocínio de variabilidade.

> PSQ: Se tivesse sido pesquisado renda de São Paulo e de uma cidade do Nordeste. Como esta pessoa teria que ler este relatório? Que é a pergunta que ela [autora do artigo] faz aqui. Será que ela tinha que fazer um perfil das moças e dos moços? Será que as meninas são muito diferentes dos meninos e eu precisaria fazer um relatório separado? Ou faria .... ela deixou para os alunos decidirem.
>
> LH: Depende do objetivo.
>
> PSQ: Traçar o perfil!
>
> OB: Provavelmente, né? Então, é quando entra a probabilidade né? <u>Provavelmente aqui tem alguém que ganhe este salário. Nem todos aqui, né? Então é um levantamento de todos, né? E não quer dizer que cada um daqueles ganhe exatamente aquilo que está na pesquisa.</u> É o caso do aluno típico, né? Aqui tem vários... perfil do aluno e isto aqui não prova que cada aluno tem que ser exatamente o que saiu na...(grifo nosso).

Quando a professora entende que existe um determinado valor e que nem todos os valores devem ser iguais a este, ela já está num nível de transição do raciocínio de variação. Ele percebe que uma amostra é representada por um valor (uma dimensão do conceito de variação) e que os outros valores não são necessariamente iguais a este (outra dimensão do conceito).

> RN: É uma estimativa.
>
> PSQ: Quando a gente fala em estimativa, o que a gente está pensando?
>
> LH: Um chute.
>
> OB: Espera-se ...
>
> PSQ: É um valor aproximado e nem todos serão iguais aquele. E é isto que Snee diz que é o pensamento estatístico. Ele vai sempre trabalhar na Estatística com um valor aproximado, uma estimativa, eu tenho que ler uma média e saber que há pessoas acima, vão ter pessoas abaixo, se é renda, tenho que saber que pessoas ganham menos, pessoas ganham mais e tem pessoas que não ganham nada.
>
> OB: Por isso que queriam o socialismo. Quando fizessem estatística, <u>todo mundo ganha exatamente igual</u>, todos teriam o mesmo padrão de vida. Por isso que queriam ...
>
> LH: Se pegar o exemplo que você deu. Se pegar a região Sul, faz o padrão de vida de cada um, faz o perfil do brasileiro. <u>A região Sul tá diferente da região Nordeste</u>. Ai você pega tudo isto daí e faz a média, vai cair mesmo o padrão, mas daí vai ficar bom o padrão, entre aspas né?

A professora OB fez surgir novamente o raciocínio sobre variabilidade, pois observou que os valores do salário poderiam ser todos iguais. No entanto, o professor LH observou que os valores da região Sul são diferentes da região Nordeste, raciocínio sobre variabilidade, mas também percebeu que a média é equilibrada devido à essas diferenças, o que permite inferir que estava pensando em distâncias da média e classificar seu raciocínio como de transição.

> PSQ: Então, é por isto que numa investigação em sala de aula, é importante que o aluno comece a pensar na questão. Por que ele vai começar a perceber...
>
> RN: Por isto que naquela brincadeira do sorvete que eu fiz com eles, tinha quatro salas, né? Na hora de fechar o trabalho, eu fechei com as quatro salas e eu não fechei só na sala que eu estava trabalhando. <u>Eu coloquei o resultados das quatro salas.</u> E lancei a pergunta: quem for abrir uma sorveteria, quais sabores você ia escolher? Foi daí que um aluno disse: <u>Nunca mais eu vou pegar de uva. Eu achei bonitinha a resposta dele. Jamais eu compraria de uva</u> .

Essa simpatia com a decisão do aluno em não mais comprar o sorvete de uva (possivelmente porque a freqüência de alunos que gostavam deste sabor era a mais baixa), representando a observação da minoria, o que Ben-Zvi (2004) relacionou com a noção intuitiva de *outliers* (discrepantes), uma dimensão do conceito, permitindo classificar seu raciocínio de variação como de transição.

Para encerrar a sessão de *feedback,* a pesquisadora fez uma síntese verbal do trabalho de Watson e Kelly (2002), que trabalharam com o número de pessoas da família. Os professores entenderam que trabalhar estatística em sala de aula requer planejamento, desde a não manipulação do entendimento do assunto até a forma de coleta e análise de dados.

A questão do planejamento incomodou os professores, pois recordaram da falácia do planejamento da escola. Como não era objetivo analisar este assunto, não foi transcrita a discussão.

**8.5 Análise do raciocínio de variação/variabilidade identificado na fase de sensibilização da pesquisa-ação.**

Foi possível observar que os professores utilizavam pesquisa como metodologia de ensino e gráficos com conteúdo disponível para analisar os resultados, que complementa o diagnóstico realizado inicialmente, em que estes professores buscaram uma formação continuada sobre o conteúdo estatístico pelo fato de terem um conhecimento restrito sobre o assunto.

O debate sobre a síntese do texto de Ponte, Brocardo e Oliveira (2003) permitiu observar a facilidade na admissão da variabilidade.

O entendimento de que as observações são diferentes é um raciocínio adequado sobre variabilidade, mas não sobre variação, pois não foi relacionado com nenhuma medida de variação. Isto já foi identificado nos estudos de Watson e Kelly (2002), Lehrer e Schauble (2002), Loosen, Lioen e Lacante (1985).

A partir das verbalizações, foi possível observar noções intuitivas de variação, que foram niveladas seguindo o modelo de raciocínio estatístico proposto por Garfield (2002).

Quadro 28 – Nível de Raciocínio sobre Variação na fase de sensibilização da pesquisa-ação

| Raciocínio sobre variação | Professor | Verbalização | Nível do raciocínio |
|---|---|---|---|
| Qualquer valor é possível | LH | Estimativa é um chute | verbal |
| Valores próximos do zero, mais homogeneidade | AM | Quanto mais próximo todas as notas de cada aluno analisada for próximo ao zero, significa que eles estão, a sala está em uma hegemonia. Ou estão aprendendo todos iguais ou nenhum está aprendendo igual. Então é para isto | idiossincrático |
| Moda | LH | Quanto mais informações ele pegar daquele assunto, daquilo que ele está analisando, mais preciso está. | verbal |
| Variação em torno da estimativa | OB | Provavelmente aqui tem alguém que ganhe este salário. Nem todos aqui, né? Então é um levantamento de todos, né? E não quer dizer que cada um daqueles ganhe exatamente aquilo que está na pesquisa. | Transição |
| Diferenças em torno da média | LH | A região Sul tá diferente da região Nordeste. Aí, você pega tudo isto daí e faz a média, vai cair mesmo o padrão, mas daí vai ficar bom o padrão, entre aspas né? | Transição |
| Percepção da minoria aliado à maioria | RN | Explicou sobre a decisão do aluno em não comprar sorvete de uva se fosse dono de uma sorveteria, devido à pequena freqüência. | Transição |

Como pode ser observada no Quadro 28, a consideração que estimativa é um chute indica um nível verbal do raciocínio de variação, pois admite sua existência.

Mesmo com a indução da pesquisadora, quando fez a seguinte afirmação: “eu tenho que ler uma média e saber que há pessoas acima, vão ter pessoas abaixo, se é renda, tenho que saber que pessoas ganham menos, pessoas ganham mais e tem pessoas que não ganha nada”, não houve a percepção de que os valores poderiam variar em torno da média aritmética e não surgiu a necessidade de uma medida de variação.

Não foi possível observar o raciocínio sobre variação/variabilidade de todos professores, mas certamente os que não se pronunciaram ouviram as discussões e possivelmente refletiram sobre o assunto.

# 9 Análise dos resultados do ciclo investigativo

Na fase de planejamento da próxima ação, o grupo de pesquisa discutiu sobre uma maneira de dar feedback aos participantes, principalmente sobre a estratégia de se trabalhar os conceitos estatísticos. A decisão foi trabalhar Estatística com os professores por meio da realização de pesquisa, para que eles pudessem vivenciar não só o processo de coleta de dados, o que eles já faziam, mas realizar o tratamento das informações obtidas utilizando os conteúdos sugeridos pelos Parâmetros Curriculares Nacionais do ensino fundamental e médio.

Em outras palavras, o grupo de pesquisa decidiu desenvolver o ciclo investigativo, de acordo com Wild e Pfannkuch (1999), junto com estes professores, para que fosse possível observar o desenvolvimento do raciocínio sobre variabilidade e variação em cada etapa deste ciclo.

Desta maneira, os resultados foram divididos em duas fases. A primeira fase referiu-se à análise do raciocínio sobre variação/variabilidade na elaboração do problema e do questionário, no planejamento da coleta de dados e na organização destas informações obtidas num banco de dados. Estes resultados estão apresentados no Capítulo 10.

A segunda fase foi composta pela identificação do raciocínio de variação/variabilidade na análise dos resultados obtidos na pesquisa, que está apresentado nos Capítulos 11, 12, 13 e 14.

É importante destacar que os resultados foram construídos de acordo com o aspecto de variabilidade e variação observado e que, por este motivo, não seguiu a ordem cronológica da implementação da ação, conforme apresentado no capítulo de Metodologia. Por exemplo, para a análise do raciocínio sobre variabilidade na representação de uma distribuição, os resultados foram construídos da seguinte maneira: tarefa proposta, resultados e discussão dos grupos, que aconteceu em 15 de Abril de 2005, o *feedback* da PSQ, que aconteceu em 29 de Abril do mesmo ano e a análise do raciocínio sobre variabilidade, que aconteceu no último encontro da formação continuada, 19 e 26 de Agosto.

## 10 Raciocínio sobre Variação/variabilidade na fase de planejamento do ciclo investigativo

A etapa de planejamento do ciclo investigativo foi composta pelo desenvolvimento das seguintes atividades: elaboração do objetivo da pesquisa, construção do questionário, definição dos procedimentos para a coleta de dados e a elaboração do banco de dados.

O objetivo da ação era propiciar condições dos professores desenvolverem o pensamento geral do pensamento estatístico[31] (WILD e PFANNKUCH, 1999) por meio da discussão sobre: o que vai ser pesquisado; como será pesquisado; o custo e o material necessário para o desenvolvimento da pesquisa; o que já se conhece sobre o assunto e quais os conceitos estatísticos do problema que terá influência direta na maneira como se analisará os dados.

O objetivo da pesquisa era verificar se haveria progresso no nível do raciocínio sobre variação e variabilidade na medida em que o pensamento geral do pensamento estatístico fosse se desenvolvendo.

Esta etapa ocupou dois encontros (01 e 08 de Abril de 2005) e a Figura 60 apresenta a primeira tarefa da fase de planejamento do ciclo investigativo.

| |
|---|
| PSQ: Diante da discussão sobre o texto do Ponte e das recomendações dos PCN para o ensino de Estatística, pudemos verificar a indicação de trabalhar com pesquisa em sala de aula. Logo, para iniciarmos nosso trabalho e nossa reflexão sobre o ensino e a aprendizagem da estatística vamos começar reconstruindo o caminho indicado pelo texto e pelos PCN. Ou seja, vamos fazer juntos uma pesquisa e a temática deve ser de nosso interesse. A partir da definição do nosso objetivo, decidimos como vamos obter as informações, com quem e assim por diante. |

Figura 60 – Tarefa 1 do planejamento do ciclo investigativo

Devido à importância que os professores atribuíram à liberdade do aluno em escolher o tema da pesquisa, deixou-se a critério deles próprios escolherem o que seria pesquisado bem como o objetivo e os sujeitos da pesquisa.

---

[31] O pensamento geral do pensamento estatístico está descrito no Capítulo 1 deste trabalho.

O coordenador do projeto, Professor Dr. Saddo Ag Almouloud, sugeriu o tema Estatística e os professores aceitaram, de imediato.

Todos os professores, exceto OB e SB, escolheram realizar a pesquisa com os professores de todas as disciplinas para verificar o que eles utilizavam de Estatística em suas aulas, uma espécie de diagnóstico para que o professor de Matemática soubesse o que deveria ser priorizado em suas aulas.

A professora OB e SB pretendiam fazer uma pesquisa com os professores de Matemática para verificar o que o professor de matemática estava ensinando de Estatística e que tipo de dificuldade encontrava. O trecho a seguir exemplifica a decisão destas duas professoras bem como seu raciocínio sobre variação.

> OB: É porque... olha! <u>O professor de Biologia, ele não vai usar a Estatística para ensinar o aluno dele</u>. O professor de Física, de Química, também não. Então, quem usa mesmo a Estatística na escola, para ensino e aprendizagem é o professor de Matemática. Ai as perguntas seriam voltadas até para..., até para uma análise nossa, né? Onde a gente tá errando, porque a gente tá errando, né? Por que os alunos não estão conseguindo assimilar, gostar da Estatística. Porque o professor de Matemática não gosta de ensinar estatística, também. Por isso que eu falei isto. (grifo nosso)

É possível notar um raciocínio determinístico na fala de OB, contrariando o que tinha sido observado nos encontros anteriores.

A pesquisadora informou sobre a dificuldade de se trabalhar com duas pesquisas com objetivos diferentes e, venceu, por maioria, a realização da pesquisa com os professores de todas as disciplinas, com o objetivo de verificar a opinião deles sobre a Estatística.

Os professores dividiram-se em grupos para elaborar as questões da pesquisa. Os grupos foram:

- Grupo 1: Professores AM, OB, IS e LF
- Grupo 2: Professores SB, LH e RN
- Grupo 3: Professores CI, RS e EVE (uma professora que só participou deste dia)

## 10.1 Construção do instrumento de pesquisa em pequenos grupos

Os Grupos 1 e 2 iniciaram a elaboração do questionário por perguntas que permitissem a identificação do respondente. Preocuparam-se, principalmente, com o tempo de magistério, pois a experiência do professor seria uma variável determinante da utilização ou não da Estatística em suas aulas.

A única questão de identificação sugerida pelo Grupo 3 foi a disciplina que ministrava, pois perceberam que professores de disciplinas distintas poderiam usar diferentes conteúdos estatísticos.

Enquanto elaboravam as questões, todos os grupos preocupavam-se com as possíveis respostas, que oscilavam entre uma visão determínistica e uma consideração da variabilidade. Segue um trecho de discussão do Grupo 3.

> RS: Se ele usa o gráfico, o que ele usa? Ele vai falar, tabelas e gráficos, jornais, tabelas e gráficos. Em que situação?
>
> PSQ: Você quer saber se ele pula a análise do gráfico ou ele lê, que pergunta você vai fazer?
>
> CI: <u>Ele vai falar: eu nem olho isto</u>. Você vai falar: o que você usa? E ele vai dizer: eu nem uso isto. Eu nem leio isto. O que a gente responde?
>
> RS: Enfoca a parte da analise do gráfico e das tabelas ou você enfoca mais a parte de leitura e interpretação? [pergunta elaborada pela professora RS pensando na leitura do texto do livro didático]
>
> PSQ: Você tem que fazer as perguntas de acordo com os “achômetros” de vocês.
>
> CI: A gente está com a resposta, agora. <u>A gente sabe que a resposta dele é esta: eu nem olho isto aqui.</u> Mas, então a pergunta é mesmo?

Devido ao fato de que os grupos estavam elaborando perguntas semelhantes, a pesquisadora solicitou que todos socializassem suas perguntas para que fosse elaborado um único questionário.

Não houve polêmica quanto as questões demográficas, o que foi rapidamente decidido (que correspondem as questões 1 a 7 do Apêndice 5). O mesmo não aconteceu com as questões que responderiam diretamente ao objetivo proposto. O professor AM leu, ininterruptamente as questões elaboradas pelo Grupo 1.

> AM: Primeira: Qual a disciplina que leciona? Segunda, para você, o que lembra a palavra estatística? Três: na sua concepção, qual a visão do seu aluno quanto a palavra estatística? Quatro: Você usa estatística em sua disciplina? Se usa, como? Cinco: com que freqüência você usa estatística em sua disciplina? Seis: Qual a sua dificuldade em usar estatística em sua disciplina? Sete: Qual a aceitação do seu aluno quando você usa a Estatística em sua disciplina? Oito: Como você avalia a compreensão ou aprendizagem dos seus alunos a respeito da

> estatística aplicada em sua disciplina? Nove: qual a sua intervenção em relação as dificuldades do seu aluno em Estatística? Dez: Qual a importância da Estatística em seu cotidiano? E de seu aluno?

A professora RN verbalizou o que estava pensando e que ainda não havia tido tempo para discutir com o grupo, haja visto que detalharam exaustivamente as questões demográficas.

> RN: Eu tinha feito um esqueminha de como seriam as questões, que é: como o professor... o que ele sabe? como usa isto? e o que a gente espera do aluno?

A professora CI leu as questões do Grupo 3.

> CI: Qual a importância da Estatística? Qual a relação que você faz com sua disciplina? Qual a sua maior dificuldade? Você usa? Quando? Porque? Como? Em que situação? O que você avalia, como você avalia o seu aluno?

A pesquisadora lembrou sobre a dificuldade de analisar as respostas de questões abertas e enfatizou que o objetivo da pesquisa-ação era discussão dos conteúdos estatísticos.

Como a sessão estava terminando, a pesquisadora solicitou duas tarefas distintas, conforme pode ser visto na Figura 61.

| |
|---|
| Tarefa 2: Pensem nas possíveis respostas para que possamos fazer alternativas nas questões e disponibilizem no fórum. No momento em que vocês estiverem pensando nessas possibilidades, eu gostaria que vocês levassem em consideração o objetivo do nosso trabalho e não apenas o objetivo da pesquisa que vocês vão realizar. Pensem que nosso objetivo é trabalhar os conteúdos estatísticos sugeridos pelos PCNs: gráficos, média, desvio padrão, etc. e que estas perguntas da pesquisa devem permitir este trabalho.<br><br>Tarefa 3: A professora CI trouxe uma material muito interessante sobre a Estatística. É da escola onde ela leciona [e mostra o material para os professores - dois volumes da coleção “Projeto Escola e Cidadania”, intitulados *As estatísticas revelam...* (Zampirolo, Scordamaglio e Cândido, 2000a) e *As estatísticas não mentem jamais?* (Zampirolo, Scordamaglio e Cândido, 2000b)], que discutem exatamente o que estamos fazendo agora. Eu gostaria que vocês lessem este material. Se a escola de vocês não tiver este material, conversem com a professora CI. |

Figura 61 – Tarefas 2 e 3 do planejamento do ciclo investigativo

## 10.2 Construção do instrumento de pesquisa pelo grupo todo

Não houve discussão no fórum, apenas a digitação das questões lidas no encontro anterior. Desta maneira, o grupo de pesquisa decidiu que o encontro seguinte (08 de Abril) seria destinado à elaboração final do questionário.

A PSQ iniciou os trabalhos com a leitura de trechos do material: *As estatísticas não mentem jamais?* Mas, não houve interesse dos professores, o que foi considerado estranho, pois o texto discutia exatamente as etapas do ciclo investigativo que estavam vivenciando: a) o que será pesquisado; b) como a pesquisa será feita; c) Organização e apresentação dos dados e d) análise dos resultados.

Isto mostra, mais uma vez, que os professores participantes da pesquisa-ação não tinham interesse pelas leituras e, possivelmente, estavam em busca de receitas prontas para ensinar Estatística.

A discussão do questionário foi retomada. Os professores não conseguiram pensar em outras questões que não fossem relacionadas ao uso de gráficos e tabelas e sequer conseguiram pensar em possíveis respostas para fechar as questões.

> RN: O principal é analisar, né? Já que o objetivo da estatística é levantar dados para você poder fazer análise do que está acontecendo e prever o que vai acontecer, precisa ver se eles [professores de outras disciplinas] estão com este enfoque, né? Ou se está sendo um simples levantamento numérico e pronto. Se eles estão usando com este enfoque. Se o professor de geografia for pegar alguma coisa, a análise de uma região, será que ele está debatendo aqueles dados, com a importância que deveria ser? Ou simplesmente apresenta a tabela e pronto. Acontece isto, isto, isto. Acho que a gente precisa parar para pensar neste enfoque, né? A análise da tabela, do, do, dos dados.
>
> PSQ: Você está falando duas coisas ai. Se ele usa o conteúdo e como ele usa o conteúdo.
>
> PSQ: Se vocês perguntarem para o professor de qualquer outra disciplina sobre o que ele usa de Estatística, o que vocês acham que vem como resposta?
>
> RN: Isto que eu acabei de falar. <u>Talvez ele possa usar como uma mera ilustração.</u> Ele não para e fica analisando os dados, observando o que acontece. Debatendo com a turma.
>
> AM: Eu acho que é até um reflexo que ele traz de fora da escola. Quem é que, realmente, pega um jornal, uma revista e dá uma analisada no

> gráfico? Pensa se aquilo é realmente verdade ou não é? Qual o objetivo daquilo? Acho que ele acaba levando isto para a escola, tá no livro dele, tá lá no livro dele, o professor de geografia, de historia.
>
> LH: Mas isto é julgamento nosso.
>
> AM: Sim. Tô fazendo analise, né? Mas se a gente não tem certeza, tem que abordar, né?

É possível observar que o professor AM transitava entre uma visão determinística, em que todos não analisavam gráficos e uma visão em que há variabilidade devido à incerteza, quando verbalizava que não tinha certeza do tipo de resposta. O mesmo aconteceu com a professora RN, como pode ser observado no diálogo seguinte.

A professora RN tinha uma visão da Estatística como ferramenta de pesquisa, mas novamente restringia-se a gráficos e tabelas. Na discussão com a pesquisadora, ela não "ouviu" sua intervenção.

> RN: Eu ainda estou em duvida na questão: Você incentiva a leitura do texto ilustrado por tabelas e gráficos?
>
> PSQ: E se a gente colocar de conteúdos estatísticos? Não só de tabelas e gráficos? Porque ele pode interpretar uma média também. Não pode? Média é um conteúdo que está sugerido pelo PCN.
>
> RN: Mas aí ele vai pegar um texto, né? A Estatística entra para ele mais como uma ilustração do texto, complemento do texto.
>
> Observadora RL: Isto é o que você está falando. Não chegou a resposta do questionário ainda para você saber.

O professor AM começou a pensar como seria feita a análise dos dados que obteriam, ou seja, começou a ampliar o pensamento geral do pensamento estatístico. O diálogo que segue ilustra este fato.

> AM: Deixa eu fazer uma pergunta aqui. Eu anotei aqui para não esquecer. É... É.... Quando você vai fazer um estudo estatístico, existe um padrão? Um número mínimo de entrevistados? E quando você faz esta entrevista, necessariamente você deve mostrar valores em porcentagens ou pode ser freqüência relativa ou absoluta? Tem um padrão para isto?
>
> PSQ: Não. E é isto que nós vamos discutir aqui. Quando você vai escolher freqüência absoluta ou porcentagem.
>
> AM: Esta pergunta veio depois que eu vi a qualificação da RL e o avaliador fez este comentário na qualificação. Alguma coisa sobre a quantidade.

> Observadora RL: Tinha treze professores e eu pus... quando eu separei, eu pus em porcentagem, sei lá eu. Sete e seis e ele perguntou porque.
> AM: E eu fiquei pensando nesta semana. Pô, tem um padrão? Quando é que eu devo fazer? Aí eu fiquei pensando nesta semana, em que situação eu devo transformar?

Foi explicado verbalmente sobre a adequação da porcentagem e a questão sobre o número de pessoas a serem entrevistadas foi deixada para uma futura discussão.

E, como nenhum professor pensou na possibilidade de elaborar questões cuja variável fosse qualitativa e quantitativa (proposta da Tarefa 2), a pesquisadora fez uma sugestão para a questão de importância da Estatística.

> PSQ: Eu queria dar uma sugestão para a questão da importância. Ao invés da gente deixar aberta: qual a importância da Estatística para sua área? Eu queria sugerir que ele desse uma nota de importância da Estatística, por exemplo, para sua área. Então, dê uma nota de zero a dez para a importância da Estatística para sua área. Então ele pode dizer que na minha área é sete mas na minha vida é dez. Então, dê uma nota de zero a dez para a importância da Estatística, sendo zero nada importante e dez muito importante.
> Observadora VG: Menos aberta, né?
> PSQ: É! E a gente depois pode trabalhar mais coisa, né? É uma medida. E a gente pode trabalhar mais coisa na Estatística, na hora da gente discutir a análise.

Os professores (principalmente SB) acharam interessante a estratégia de atribuir pontuação para a importância, devido ao fato de fechar a questão, mas não pelo fato de permitir uma discussão de outros conteúdos estatísticos.

Esta mesma professora, no momento de sua argumentação para justificar sua escolha em realizar a pesquisa somente com professores de matemática, verbalizou “Eu acho que, dificilmente, vai aparecer algum professor que já trabalhou com Estatística”, o que mostra sua prontidão para lidar com a variabilidade nas respostas.

Quando o questionário já estava quase pronto, os professores começaram a discutir os procedimentos que adotariam, mesmo sem saber. Levantaram questões como a presença necessária do professor-pesquisador no momento em que o professor da escola fosse responder ao questionário; solicitar ao diretor da escola a utilização de uma HTPC para aplicar os questionários; a importância de mostrar para o diretor da escola os resultados que obteriam.

Desta discussão, a PSQ apresentou brevemente a questão ética envolvida na pesquisa, a necessidade de solicitar autorização e dar devolutiva dos resultados aos participantes entre outros procedimentos.

Os participantes ficaram eufóricos em dar devolutiva para o diretor da escola e se comprometeram em fazê-lo numa outra HTPC, com a presença dos outros professores.

O questionário elaborado bem como a carta de apresentação para a escola participante encontram-se no Apêndice 5. Eles foram aplicados no período de 18 de Abril a 29 de Abril de 2005.

### 10.3 Construção do banco de dados

Os professores gostaram da experiência de coletar dados e comentaram sobre alguns fatos que consideraram interessantes, mas não houve grande entusiasmo em dar início à elaboração do banco de dados, que foi feito pela PSQ.

Com a utilização do Microsoft Excel, a PSQ digitou as nove primeiras questões e entregou para cada professor um disquete contendo três arquivos:

- banco de dados com 110 linhas numeradas, cada coluna com o nome da variável e as nove primeiras questões digitadas;
- um arquivo de texto contendo a codificação utilizada nas questões digitadas e
- um arquivo de texto contendo todas respostas de cada questão aberta.

A categorização das questões abertas demorou dois encontros (06 e 13 de Maio de 2005) e foi realizada no laboratório de informática para que fosse possível utilizar o Microsoft Excel.

Os professores tiveram muita dificuldade em elaborar categorias que permitissem abranger todas as diferentes respostas, o que lhes causou desânimo e certeza de que não trabalhariam questões abertas com seus alunos.

A variação na interpretação das respostas dos entrevistados causou desentendimentos entre os participantes, porém ninguém questionou sobre o fato que a variação dificultava a realização da tarefa, situação comum em diferentes áreas do conhecimento.

Até o final da formação continuada, algumas respostas dos professores entrevistados ainda não havia sido interpretada, como pode ser observado no Apêndice 7.

### 10.4 Análise do raciocínio sobre variação/variabilidade na fase de planejamento do ciclo investigativo

Foi possível observar que os professores se envolveram completamente com o planejamento da pesquisa, esquecendo-se que fariam a análise dos resultados. Não conseguiram atribuir possibilidades de resposta para as questões, embora tenham havido inúmeras intervenções da PSQ para isto.

Os professores oscilaram entre uma visão determinística, em que apenas uma resposta poderia sair, e uma visão da variabilidade devido à incerteza, quando consideravam que dificilmente poderia acontecer tal coisa ou quando atribuíam diferentes respostas para a pergunta. Diante disto, pode-se inferir que não houve progresso no nível de raciocínio sobre variação, ou seja, não foi a vivência do planejamento do ciclo investigativo e da elaboração do instrumento de pesquisa que permitiu um avanço no nível do raciocínio sobre variação.

Em contrapartida, os professores preocuparam-se com todos os itens elencados por Wild e Pfannkuch (1999) como parte do pensamento geral do pensamento estatístico, mesmo sem saber disto. O único item menos explorado pelos participantes foi o julgamento do tipo de analise que seria feito e se isto mudaria o questionário.

Além da preocupação do professor AM com a utilização das porcentagens, o aluno LJ queria saber como poderia analisar as notas de importância se cada professor lecionava uma disciplina diferente e que, possivelmente, tinham percepções diferentes da Estatística. Este raciocínio demonstra uma preocupação com a análise que seria feita, buscando as causas para possíveis diferenças, que é descrito por Wild e Pfannkuch como ciclo interrogativo do pensamento estatístico.

A utilização de pesquisa como metodologia de ensino não proporcionou avanço no raciocínio sobre variação, mas proporcionou a mobilização de três dimensões do pensamento estatístico: ciclo investigativo, pensamento geral e ciclo interrogativo, indicando um progresso no pensamento estatístico dos professores.

Parte dessas dimensões do pensamento estatístico estava descrita e exemplificada no material disponibilizado pela professora CI (Zampirolo, Scordamaglio e Cândido, 2000b), mas os professores não se interessaram pela leitura.

# 11 Raciocínio sobre variação/variabilidade na representação da distribuição de freqüência simples

Como a coleta de dados demandaria duas semanas, a análise dos resultados do ciclo investigativo foi iniciada em 15 de Abril, antes mesmo dos professores aplicarem os questionários em suas escolas.

Para isto, foi escolhido trabalhar com a pesquisa desenvolvida por Nifoci e Silva (2004), cujo objetivo foi identificar se o aluno ingressante no ensino médio teve uma aprendizagem significativa do Teorema de Pitágoras[32].

Como foi observado no diagnóstico e na fase de sensibilização, o conteúdo estatístico mais conhecido e desenvolvido em sala de aula por estes professores era representação gráfica e, por este motivo, decidiu-se dar início à fase de análise dos resultados do ciclo investigativo pelos gráficos, inicialmente a partir de distribuição de freqüência simples. A tarefa solicitada aos professores está descrita na Figura 62.

| |
|---|
| PSQ: Diante da compreensão da pesquisa realizada por Nifoci e Silva (2004), vamos analisar as primeiras variáveis do banco de dados (sexo, idade e disciplina que mais gosta) tal como vocês acham que o seu aluno faria. |

Figura 62 – Tarefa 1 da análise dos dados do ciclo investigativo – representação de distribuição de freqüência simples

Para realizar esta tarefa, os professores dividiram-se em duplas e foi disponibilizado papel milimetrado, quadriculado, lápis de cor, canetas coloridas, etc..

## 11.1 Resultados do Grupo 1

Este grupo foi composto pelos professores RN, LH e por uma professora que só participou deste encontro e está identificada como EVE.

A professora RN acreditava que seus alunos elaborariam uma tabela para cada variável e o professor LH sugeriu que verificassem a relação entre as variáveis.

---

[32] O banco de dados (Apêndice 6) da pesquisa de Nifoci e Silva (2004) foi entregue aos professores participantes.

> LH: O que a gente tá discutindo aqui é o seguinte. Dá pra misturar? <u>Sexo masculino ou feminino e a pergunta gosta ou não gosta</u>? Sei lá, acho que fica um negócio, né?
>
> RN: Não! São três grandezas.
>
> LH: Então, eu acho que não daria. Ou dá? Deixa eu ver… Sexo feminino... O que os meninos fariam? O que as meninas fariam? Certo? Como eles responderiam?
>
> RN: Tem muitas possibilidades aí.
>
> LH: Tem várias possibilidades. Aí entra a turma da probabilidade. Quantos responderiam? Não... porque já tá definido: Masculino. As meninas responderiam uma coisa. Os meninos outras. <u>Mas daí daria pra ver a tendência de meninos e meninas, quais matérias que elas mais gostam.</u> Entendeu? É... é isso mesmo, quando fala sexo é isso, é pra saber, é pra saber o que as meninas e os meninos não gostam. Tem que fazer três tipos de... três tipos de gráfico. É um prá, em cima disso aí, tem que passar pra eles...
>
> RN: Dá as três de uma vez. Dá as três situações junto?
>
> LH: <u>Se as idades, a gente deve saber, são todos iguais aí, não vai, não vai</u>…
>
> EVE: O que você tá querendo dizer assim (inaudível) qual a matéria que ele mais gosta? É isso?

É possível notar que o professor LH estava olhando o todo (no sentido de Ben-Zvi, 2004) e buscando traçar tendências das meninas e dos meninos a partir da observação da disciplina que mais gostavam, o que indica novamente o uso da moda e a manutenção do raciocínio verbal de variação.

Foi possível observar, também, que o professor LH percebeu que a idade dos alunos variava pouco e, possivelmente por este fato, começou a relacionar gênero com disciplina preferida. Isso mostra seu raciocínio de variabilidade, comparando uma idade com outra idade.

De acordo com a sugestão do professor LH, o grupo contou o número de meninos (e o número de meninas) que gostava de cada disciplina, elaborando uma tabela cuja porcentagem foi calculada em relação ao total de alunos por gênero.

Tabela 25 – Representação elaborada pelos professores RN e LH das variáveis gênero e disciplina que mais gosta *

| Gênero | Total | Disciplina que mais gosta | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Port. | bio | mat | geo | his | qui | fis | Ed.fis | ing | EA |
| M | 17 | 1 | 1 | 2 | 4 | 2 | 2 | 3 | 2 | --- | --- |
| F | 40 | 6 | 12 | 7 | 1 | 8 | 3 | 1 | --- | 1 | 1 |
| %M | 70 | 5,9 | 5,9 | 11,8 | 23,5 | 11,8 | 11,8 | 17,6 | 17,7 | --- | --- |
| %F | 30 | 15,0 | 30,0 | 17,5 | 2,5 | 20,0 | 7,5 | 2,5 | --- | 2,5 | 2,5 |

*A única diferença entre esta tabela e a original (elaborada pelos professores) é a formatação.

Dada a conclusão da tabela (Tabela 25), os professores escolheram elaborar um gráfico de setores para representá-la. A observadora MM explicou que um gráfico de setores tem o objetivo de representar o todo (100%) e que, por este motivo, seria necessário que eles elaborassem dois gráficos de setores, um para cada gênero.

A professora RN sugeriu ao grupo que fizessem o gráfico de “barrinhas” (ela rascunhou no caderno um gráfico de barras múltiplas), mas o grupo não deu atenção e começaram a calcular novamente as porcentagens, agora em relação à cinqüenta e sete alunos (toda a amostra).

A representação gráfica deste grupo está parcialmente reproduzida na Figura 63, pois o gráfico de setores que eles elaboraram tinha dez partes de tamanhos iguais, em que cada uma foi dividida em meninos e meninas.

O fato de dividirem o gráfico de setores em dez partes iguais pode ter duas explicações: esboçaram rapidamente o tipo de representação gráfica que pretenderiam fazer, pois já estava no final da sessão ou realmente eles não conheciam a técnica para a elaboração deste tipo de gráfico.

No *feedback,* durante a discussão sobre a elaboração do gráfico de setores, foi salientado a importância da clareza dos setores e que o fato de terem dividido cada setor em feminino e masculino dificultaria a leitura.

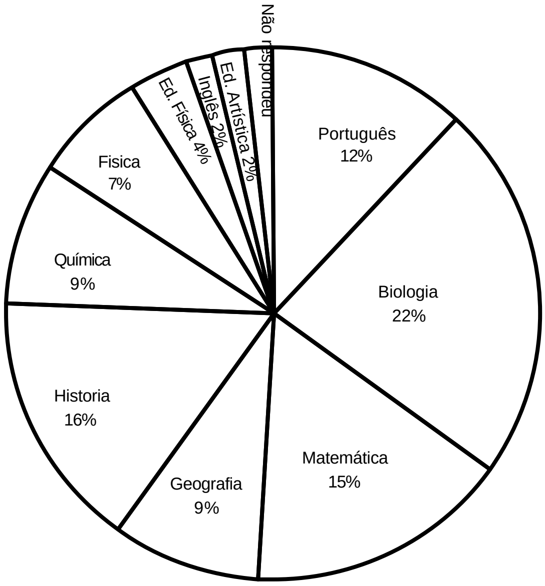



Figura 63 – Representação gráfica elaborada pelos professores RN e LH das variáveis gênero e disciplina que mais gosta

## 11.2 Resultados do Grupo 2

Este grupo (professores AM e LF) preocupou-se o tempo todo com a interpretação das representações que elaborariam, inclusive inferindo sobre os possíveis motivos pelos quais as meninas gostavam mais de Biologia.

Em outras palavras, além de perceberem a preferência da maioria das meninas, estavam buscando explicações para tal ocorrência. Este fato corrobora o que Meletiou (2000) identificou em seu trabalho, que quando se utiliza contexto do mundo real é necessário lidar com as crenças e valores dos sujeitos.

Os professores elaboraram uma distribuição de freqüências simples para disciplina de preferência e quatro distribuições, também simples, para apresentar a freqüência de meninos e meninas em cada disciplina preferida, o que mostra o não conhecimento destes professores de uma distribuição de freqüências conjunta, também denominada de tabela de contingência.

Quando foi questionado aos professores porque escolheram quatro disciplinas para elaborar a distribuição de freqüências por gênero, a professora LF explicou que “as disciplinas escolhidas por poucos alunos teriam uma barra tão baixa no gráfico, que não permitiria ser enxergada”.

Quadro 29 –Representação elaborada pelos professores AM e LF das variáveis gênero e disciplina que mais gosta *

| Sexo | Disciplina + gosta | Física | Química | Biologia | Matemática |
|---|---|---|---|---|---|
| 40M - 70%<br>17H - 30% | Português – 7<br>Biologia - 13<br>Química – 4<br>Física - 4<br>Inglês - 1<br>Matemática - 9<br>História - 10<br>Geografia - 5<br>E.A. – 1<br>E.F. – 2<br>Nenhuma - 1 | Fem 40 1<br>Masc 17 3 | Fem 40 2<br>Masc 17 3 | Fem 40 12<br>Masc17 1 | Fem 40 7<br>Masc17 2 |

* Exatamente como os professores apresentaram.

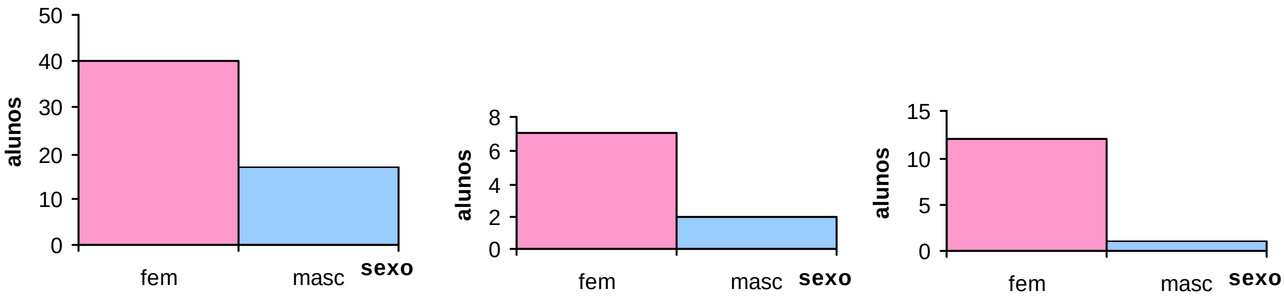



Figura 64 – Gráficos elaborados pelos professores AM e LF para representar as variáveis gênero e disciplina de preferência.

Como é possível observar na Figura 64, este grupo tinha a intenção de representar a variação das respostas, mas não tinham o conhecimento estatístico para realizar a tarefa que idealizaram.

A preocupação com a interpretação da distribuição de freqüências foi ratificada na verbalização dos dois professores: " Elas gostam mais de Biologia" (LF) e "a maioria das meninas gostava de biologia" (AM).

Embora ambos estivessem preocupados com a análise da distribuição, o professor AM estava levando em consideração a variação no momento que retratava sobre a maioria, enquanto que a professor LF apresentava uma resposta determinística.

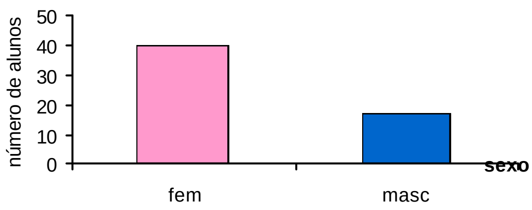
A professora OB questionou o grupo em relação à apresentação com freqüência, haja vista que os grupos tinham tamanhos diferentes, o que ilustrava a utilização de um raciocínio proporcional. Ela explicou que se refere a um aluno (entre 17) que preferia Biologia e doze alunas (entre 40) que preferiam a mesma disciplina, mas o grupo não se sentiu incomodado.

## 11.3 Resultados do Grupo 3

Diferentemente dos grupos anteriores, estas professoras (OB e SB) se preocuparam com os aspectos formais de representações gráficas, sem qualquer menção sobre a interpretação dos resultados obtidos. Elas elaboraram os gráficos como consideravam que seus alunos de ensino médio o fariam.

Elas se preocuparam em elaborar os gráficos em papel quadriculado, com escala demarcada, utilizando cores para as barras e mencionaram que, possivelmente, os alunos não se preocupariam com os títulos nos eixos, escalas, legendas e iriam preferir freqüência ao invés de porcentagem, gráfico de barras ao invés de setores e se interessariam em colorir.

As representações do grupo estão apresentadas na Figura 65.



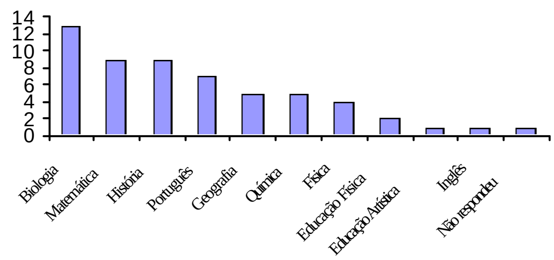




Figura 65 - Gráficos elaborados pelas professoras OB e SB para representar as variáveis gênero e disciplina de preferência.

## 11.4 Resultados do Grupo 4

Este grupo, composto pelos alunos AG e LJ, teve uma postura semelhante ao Grupo 3, pois se preocuparam com a forma e não com a interpretação. Eles elaboraram a tabela de distribuição de freqüência para cada variável e, em seguida, apresentaram o respectivo gráfico. A apresentação da variável gênero está na Figura 66 e a representação da variável disciplina preferida foi igual à do Grupo 3 e, por esse motivo, não foi reproduzida aqui.

Variável Gênero

| Sexo | f. absoluta | Porcentagem |
|---|---|---|
| 0 | 40 | 70,2% |
| 1 | 17 | 29,8% |
| | 57 | 100% |

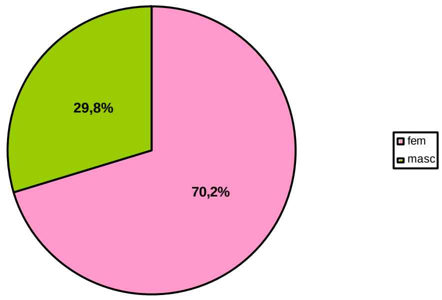



Figura 66 – Representações elaboradas pelos alunos AG e LJ para representar a variável gênero.

LJ explicou que o aluno do ensino médio teria condições de elaborar esta representação, pois nesta etapa da escolaridade já teriam condições de identificar grandezas diretamente proporcionais e relacioná-las com ângulos e este último com porcentagem.

A apresentação do gráfico de setores com 70,2% e 28,9% gerou novamente a discussão sobre as porcentagens. Os professores questionaram se o arredondamento não seria mais adequado para este gráfico.

## 11.5 *Feedback*

Durante a realização das atividades, foi possível observar que os professores desconheciam, mesmo que parcialmente, as técnicas para representação de distribuição de freqüências simples e apresentaram muita dificuldade para elaborar uma distribuição de freqüência conjunta.

Por este motivo, o grupo de pesquisa decidiu sobre o esclarecimento de diversos conceitos estatísticos, dando prioridade à formalização destas distribuições, o cálculo de porcentagem e as possibilidades de representação gráfica em ambos os casos, que foi feito em 29 de Abril.

Os professores se interessaram muito pelas explicações de procedimento, possivelmente por já terem sido questionados pelos seus alunos.

O assunto que causou maior espanto foi distribuição de freqüências conjunta e as diferentes possibilidades de calcular a porcentagem. Para exemplificar, foi analisado que 30% do total das mulheres gostavam de Biologia e que 92% do total de alunos que gostavam de Biologia eram mulheres. Ou seja, que eles poderiam calcular a porcentagem por linha (por gênero), por coluna (por disciplina) ou pelo total dos alunos (57).

Os conteúdos trabalhados e exemplificados foram: variável quantitativa e qualitativa, freqüência absoluta, freqüência relativa, porcentagem, distribuição de freqüências, gráficos de setores, barras simples e barras múltiplas.

Para terminar a sessão, os professores fizeram alguns exemplos com a utilização de planilha eletrônica e perceberam que esta ferramenta não trabalhava sozinha, como pensavam, mas que eles precisariam elaborar a distribuição de freqüências e solicitar o gráfico.

### 11.6 Análise de uma de uma distribuição de freqüência simples.

O momento do *feedback* foi marcado pelas discussões das técnicas para resolver a tarefa (no sentido de Chevallard, 1995) de elaborar a representação gráfica e não foi explorada a tarefa de analisar a representação gráfica, que permitiria compreender o raciocínio de variação/variabilidade. Isto pode ser explicado pela dificuldade da autora em administrar adequadamente a interpenetração de papéis, neste momento priorizando o papel de formadora e não de pesquisadora. Como relatado no capítulo de Metodologia, Franco (2005) explica que a grande questão da OPC (observação participante completa) é administrar o equilíbrio entre o papel de pesquisador e o papel de observador participante da pesquisa.

Por este motivo, no penúltimo dia da formação (19 de Agosto), foi solicitado aos professores que realizassem a tarefa apresentada na Figura 67.

Você encontra abaixo uma tabela que contém o número de letras de sobrenomes de 35 alunos israelenses e 35 alunos norte-americanos que pertencem a uma mesma escola.

1) Pense nos sobrenomes de seus alunos e faça uma estimativa da distribuição do tamanho de 35 sobrenomes.
2) Elabore uma frase comparando sua distribuição com os sobrenomes israelenses e norte-americanos;

Figura 67 – Tarefa 2 da análise dos dados do ciclo investigativo – representação de distribuição de freqüência simples

Essa tarefa foi baseada no estudo de Ben-Zvi (2004), cujo objetivo era perceber o raciocínio de variabilidade dos professores e verificar se eles utilizariam as medidas de tendência central e dispersão para fazer a análise, pois embora fossem duas distribuições de freqüências simples, dizia respeito à variável discreta.

Nesta etapa da formação continuada, os professores já tinham vivenciado situações de elaboração da distribuição de freqüência simples e com dados agrupados e sua relação direta com o tipo de variável. Também já tinha sido discutido sobre as medidas de tendência central e dispersão, tanto no que se refere aos cálculos quanto as suas interpretações.

Foi fornecida a Tabela 26 (apenas com as cinco primeiras colunas) e foi explicado que se referia a uma distribuição de freqüência simples de uma variável quantitativa discreta – número de letras do sobrenome. Foi solicitado que inserissem uma coluna para apresentar a distribuição do número de letras dos sobrenomes de seus alunos, que hipoteticamente representariam os alunos brasileiros matriculados na escola de Ben-Zvi.

Os professores envolveram-na na tarefa e começaram a lembrar de sobrenomes de seus alunos: Vieira, Gomes, Oliveira, Silva, Souza, Mendes, Queiroz, Barbosa, Ferreira, Nascimento, Santos, Jesus, Carvalho, Pinto, Lopes, Sá, entre outros e as seis últimas colunas da Tabela 26 apresentam as distribuições elaboradas pelas duplas de professores.

Tabela 26 - Número de letras de sobrenomes de alunos israelenses, norte-americanos e brasileiros (de cada dupla de professores).

| Número de letras dos nomes | Freqüência e porcentagem de alunos | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | israelenses | | americanos | | LH e RS | | OB e CI | | AM e SB | |
| | f | % | f | % | f | % | f | % | f | % |
| **2** | 1 | 3 | | | 1 | 3 | | | | |
| **3** | 7 | 20 | | | | | | | | |
| **4** | 11 | 31 | 4 | 11 | | | | | | |
| **5** | 4 | 11 | 2 | 6 | 18 | 51 | 13 | 37 | 12 | 34,3 |
| **6** | 4 | 11 | 10 | 29 | 8 | 23 | 6 | 17 | 6 | 17,2 |
| **7** | 6 | 17 | 4 | 11 | 6 | 17 | 7 | 20 | 8 | 22,8 |
| **8** | 2 | 6 | 9 | 26 | 2 | 6 | 4 | 11 | 3 | 8,6 |
| **9** | | | 2 | 6 | | | 3 | 9 | | |
| **10** | | | 1 | 3 | | | 2 | 6 | 4 | 11,4 |
| **11** | | | 3 | 9 | | | | | 2 | 5,7 |
| **Total** | **35** | **100** | **35** | **100** | **35** | **100** | **35** | **100** | **35** | **100** |

Os professores LH e RS encontraram um sobrenome com duas letras (Sá), pois o professor LH entendeu que precisariam encontrar sobrenomes brasileiros para igualar as duas distribuições.

Os resultados e a análise de cada grupo estão apresentados nos subcapítulos seguintes.

### 11.7 Resultados do Grupo 1

A PSQ solicitou que os professores elaborassem uma análise dos sobrenomes, tal como se fosse elaborar uma nota no jornal. Imediatamente, os professores LH e RS observaram e verbalizaram a moda de cada distribuição, como ilustra o diálogo seguinte.

> PSQ: Então, como é que vocês vão comparar?
>
> LH: Os americanos... com oito letras.....e tem dois judeus, oito brasileiros e nove americanos.
>
> RS: <u>Não sei nem se é isso que estou pensando (risos). Aqui, olha. É melhor pegar os israelenses 31%, com quatro letras, 29% os americanos com seis letras, e a de brasileiros com cinco letras, 51%.</u>

> LH: Brasileiro está parecido com americano.

Ambos estavam usando a moda, porém com estratégias diferentes. A professora RS observou a moda em cada distribuição, enquanto o professor LH pegou a moda na distribuição de americanos (mas ele ainda não tinha observado que a moda era seis letras e não oito) e comparou com as porcentagens das outras duas distribuições. Deve-se ressaltar que nenhum deles comentou que a maior freqüência era a moda.

Essa estratégia de análise pode ser categorizada como raciocínio verbal de variação, pois apresenta uma primeira idéia da variação da distribuição, mas necessita de algum complemento.

Houve um momento de silêncio no diálogo dos dois professores até que o professor LH apresentou sua nova estratégia de análise, que se restringia à descrição de cada distribuição, o que fora rapidamente negado pela professora RS.

> LH: Vamos lá. Brasileiros não têm nenhum com quatro letras. Com cinco letras...
>
> LH: Em pesquisa realizada sobre os sobrenomes israelenses, americanos e brasileiros, observamos que... né?... Observamos que com quatro letras, os sobrenomes israelenses... É, mas daí... Vai ficar um montão. [mas continua o raciocínio] 31% de israelenses, 11% de americanos e zero de brasileiros [referindo-se a 4 letras], né? Com cinco letras: 11% israelense, 6% de americanos, e brasileiros com 51%. Mas daí, eu estou descrevendo!
>
> PSQ: Em cima dessa descrição que o senhor já identificou, vamos comparar os três.
>
> RS: Olhei na tabela a freqüência dos três: Israelense, americano e brasileiro, tem que ver qual que é o maior. Tanto com cinco e seis letras. Olha: Cinco letras os israelenses tem 11%, americanos 6% e brasileiros 51%. Com seis letras: 11, 29, 31. Então, a gente concentra aqui. <u>Entre as três nacionalidades está variando entre cinco e seis letras os sobrenomes.</u> <u>Agora se quiser especificar qual o que está mais distribuído, aí seria de seis, tá vendo? Por que 11, 29, 23. Aqui ficou 11, por 51 tá vendo, e aqui seis, uma diferença.</u>
>
> Homem<u>: Um desvio grande</u>.

A estratégia de raciocínio da professora RS foi identificar o tamanho de nome que tinha maior freqüência (cinco e seis letras) em apenas dois grupos, ou seja, pegou apenas uma parte da distribuição, estratégia já relatada nos estudos de Ben-Zvi (2004), Makar e Confrey (2005) e Hammerman e Rubin (2004).

A partir da escolha da parte da distribuição, cinco e seis letras, ela observou a categoria que tinha menor variação entre as freqüências. Para facilitar a compreensão do raciocínio dessa professora, a tabela a seguir apresenta apenas a parte da distribuição observada.

Tabela 27 – Parte da distribuição analisada pela professora RS

| Número de letras do sobrenome | % de israelenses | % de americanos | % de brasileiros | Variação entre as freqüências |
|---|---|---|---|---|
| Cinco letras | 11 | 6 | 51 | maior |
| Seis letras | 11 | 29 | 23 | menor |

Isto mostra que ela está utilizando dois raciocínios de variação. O primeiro é a quebra da distribuição a partir da observação da moda em duas categorias, que pode ser classificado como raciocínio verbal. No entanto, quando ela observa a variação entre as freqüências nas categorias de maior freqüência, indica um raciocínio idiossincrático de variação.

Os alunos de Ben-Zvi (2004) elaboraram a frase "A tendência é que a freqüência de nomes curtos é maior em Israel do que nos Norte-americanos, mas a freqüência de nomes longos é maior nos norte-americanos", o que o autor denominou de comparação de freqüências de dois sub grupos.

Quando as duas estratégias de análise são comparadas, percebe-se que os alunos de Ben-Zvi quebraram a distribuição em duas ( nomes curtos e nomes longos) enquanto os professores restringiram-se à moda de cada distribuição, escolhendo aquela que consideraram mais uniforme (seis letras), sem levar em consideração que 51% dos brasileiros tinham sobrenomes com cinco letras. Outro fato que merece destaque é que os alunos de Ben-Zvi não tinham aprendido o conceito de desvio padrão.

Houve, aproximadamente, cinco minutos de silêncio e eles começaram a perceber os valores máximos e mínimos das distribuições.

> RS: Uma pessoa com duas letras!
>
> LH: Dois empatou, tanto brasileiro quanto israelense. Com três só tem israelense.

> RS: Mas a gente tem que detalhar tudo? A gente poderia responder, então, que gira em torno de cinco, seis, letras, né. E aí, fazer algum comentário sobre israelenses com menos letras no nome e brasileiros com mais..... Mais? Americanos, maior número de letras dos americanos são seis. Maior número de letras dos americanos não é de seis letras?

A interpretação da utilização da moda e dos valores máximos e mínimos pode ser classificada como um raciocínio de transição sobre variação, pois apresenta duas dimensões do conceito, mas não permite identificar a densidade das observações.

Essa estratégia de utilização da moda e dos valores máximos e mínimos já foi constatada nos estudos do próprio Ben-Zvi (2004) e também em Makar e Confrey (2005) e Reading (2004).

O diálogo entre os dois professores prossegue.

> LH: Não, não! O número de israelenses e brasileiros com seis letras no sobrenome são oito, os americanos são 10, tá? Agora, o que é maior, o que se destaca mesmo é o número de brasileiros com cinco letras.
>
> RS: Mas, então eu acho que a gente tem que comparar as três e pronto. As três giram em torno de cinco, seis letras.

A professora RS apresenta um raciocínio de transição de variação, pois percebe a moda em duas distribuições e cria um intervalo de cinco a seis letras no sobrenome, que poderia ser entendido como a representação da maioria.

> LH: Mas ......Os israelenses 11% e os brasileiros 51%.
>
> RS: A melhor distribuição está no cinco. Que aqui é 51 e aqui é só seis. Tem um monte de coisa. A maior freqüência foi de brasileiros com cinco letras.
>
> LH: Isso.
>
> RS: A menor freqüência foi de israelenses com duas letras, no caso pessoas.
>
> LH: Olha, brasileiros coincide com os israelenses. Nove letras só têm americanos, dez letras só têm americanos. Brasileiros e judeus não têm. Isso aqui é um dado importante.

Quando a professora RS relatou sobre a categoria de menor freqüência, ela estava observando a minoria que, segundo Ben-Zvi (2004), é uma pré-concepção de *outliers* (discrepantes) e, foi categorizada como um raciocínio de transição de variação, pois percebeu a maioria e a minoria.

É possível observar que os professores estavam apresentando raciocínios distintos para o problema proposto e, depois de um breve momento de silêncio, o professor LH permaneceu na interpretação da menor variação entre as freqüências, como pode ser observado no diálogo seguinte.

> LH: Quem está mais homogêneo nessa história toda?
>
> RS: Com seis letras.
>
> LH: <u>Não. Americanos. Tá mais bem distribuído</u>.
>
> RS: Você vai colocar aqui que está observando que o menor número de letras está no nome dos alunos israelenses, e o maior número de letras do povo americano?.... E a maior concentração dos brasileiros está no cinco.

Este diálogo, que na verdade não é um diálogo, mas sim reflexões feitas em voz alta, mostrou que os professores tinham diferentes raciocínios de variação para o mesmo problema proposto e isto se reflete na frase elaborada: "Baseado em dados coletados em pesquisa, verifica-se que a variação na quantidade de letras dos sobrenomes entre os povos israelenses, norte-americanos e brasileiros <u>gira em torno de 5 e 6 letras</u>, sendo que a <u>distribuição mais uniforme é 6 letras</u>; em destaque, podemos citar que o <u>menor número de letras</u> está com o povo israelense e o maior com o provo norte-americano, já o brasileiro tem em sua maioria 5 letras".

Para elaborar essa análise, eles fizeram um novo recorte na distribuição fornecida, observando as porcentagens nos três números de letras que tinham a maior freqüência, embora deveriam ter observado quatro letras e não sete.

Tabela 28 – Recorte da distribuição de freqüência elaborado pelos professores RS e LH

| Número de letras | israelense | Norte americano | brasileiro |
|---|---|---|---|
| 5 letras | 11% | 6% | 51% |
| 6 letras | 11% | 29% | 23% |
| 7 letras | 17% | 11% | 17% |

### 11.8 Resultados do Grupo 2

Esse grupo era composto pela dupla OB e CI e pela dupla SB e AM, que elaboraram distribuições diferentes, mas as discussões foram feitas conjuntamente.

Tal como o Grupo 1, a primeira idéia foi verificar a moda em cada grupo.

SB: Norte-americano com seis.

AM: Quantidade de letras?

SB: 31. Com quatro letras?

AM: Dois com duas letras, né?

OB: A gente tinha pensado assim, olha: Chegar ao maior, por exemplo, chegar a maior freqüência, tá? Maior quantidade de nomes israelitas, tem apenas quatro letras, sendo que, os americanos, a maior freqüência é com 29, possui seis letras, e os brasileiros, freqüência de 20 e possui sete letras, né?

PSQ: Cinco, né? 37%!

AM: Pegando a maioria, dez e cinco, e dez e seis. Então vamos lá. Se nós pegarmos a quantidade de alunos 12, 10 e 11, que mais aparece, a quantidade de letras fica próxima [referindo-se à maior freqüência dos brasileiros, americanos e israelenses, respectivamente]. E se a gente pegar a quantidade de alunos que mais aparece... não, quantidade de alunos que tem...Peraí! Tem alguma coisa errada aqui. Quantidade de alunos: Dois alunos com 11 letras, é, quantidade de alunos número maior de... olha, é assim: A quantidade de alunos para o número maior de letras. Está todos próximos, né? E a quantidade de alunos para...

SB: Quem aparece mais!

As professoras SB e OB estavam observando a moda de cada distribuição, já categorizada como raciocínio verbal de variação e o professor AM estava procurando complementar a análise delas observando, também, os valores máximos e mínimos, raciocínio de transição de variação. Devido à resposta da professora SB que tinha uma conotação de reprovação à sua interpretação, o professor AM mudou sua estratégia de raciocínio.

AM: Maior quantidade de alunos para um número maior de letras. Menor quantidade de alunos para um número menor de letras. Eu quero chegar no seguinte: Quando você reduz a quantidade de alunos, diminui a quantidade de letras...

O professor AM estava fazendo uma análise de covariância, na medida em que percebia que a variação da quantidade de alunos determinava a diminuição do número de letras do sobrenome. Esse raciocínio pode ser categorizado como idiossincrático, pois a freqüência de alunos não é uma variável.

Como nenhum professor se pronunciou a respeito de sua análise, ele continuou a verbalizar o que estava pensando.

> AM: E você reduz a quantidade de alunos do brasileiro com americano, ou seja, a minoria de brasileiros e americanos tem a quantidade de letras maior, né? Depois a gente melhora<u>. Minoria dos brasileiros e americanos tem um número de letras entre 11 e 10, certo</u>? Enquanto isso, no caso dos israelenses...

Este último diálogo mostra a mudança de raciocínio do professor AM, deixando de apresentar a maioria para apresentar a minoria, já apresentado pela professora RS e categorizado como raciocínio de transição de variação.

Depois de um longo trecho observando e verbalizando as porcentagens da tabela, o professor AM faz a seguinte análise:

> AM: Olha, <u>brasileiros e americanos aparecem entre quatro e onze letras</u>, né? <u>Ficam entre quatro e oito</u>. Os brasileiros não aparecem entre dois e três. <u>Os três aparecem entre cinco e oito</u>.

Esse raciocínio do professor AM mostra sua tentativa de elaborar um intervalo de variação das três distribuições conjuntamente. Sua estratégia de raciocínio foi a visualização dos valores máximos e mínimos das três distribuições conjuntamente, de maneira que esses valores da variável tivessem, pelo menos, freqüência um em todas as distribuições, que indica um raciocínio de transição.

Como a professora OB só conseguia pensar na maior porcentagem, a PSQ perguntou se o fato de ela publicar que 31% dos israelenses têm quatro letras não poderia permitir a um leitor imaginar a existência de um nome israelense com 45 letras. Ou seja, a PSQ explicou que ela estava descrevendo a maior porcentagem, não permitindo ao leitor compreender como seriam os outros 69% dos nomes. Intrigada com esta intervenção, OB dá continuidade ao diálogo.

> OB: <u>Nem espremendo sai mais. Vamos tirar a média, né?</u>
>
> CI: Vamos lá.
>
> OB: Pega a calculadora.
>
> CI: É muita informação.

A professora OB sugeriu a utilização da média como última alternativa em seu *ranking* de conceitos. Como salientado por Hammerman e Rubin (2004), raramente utiliza-se a média para comparar dois grupos graficamente, o que pode ser estendido para a tabela.

A continuação do diálogo indica a percepção dos professores sobre o objetivo da atividade.

> OB: Alexandre você terminou a frase aqui? [Referindo-se à frase da idade que seria colocada no relatório]
>
> AM: Por isso que ela voltou com isso aqui (risos).
>
> OB: Média ponderada, pra contar as médias aqui, entendeu? 2x1 + 3x7 + 4x11 +...
>
> CI: Se a gente tirar a média daqui é mais fácil.
>
> AM: Para que?
>
> OB: Para comparar.
>
> OB: Ô Alexandre! Tô me sentindo um dos alunos...
>
> AM: Mas olha a diferença, lá, conta todo mundo faz, o problema é pensar sobre ela. <u>O problema não é fazer a conta, é pensar sobre ela</u>.
>
> OB: Ô Claudia, eu acho que tô burrinha, mesmo. Pra tirar a média ponderada aqui é dois vezes um, mais três vezes sete...
>
> PSQ: Isso. Então você acha que é melhor comparar com a média! É uma decisão, não precisa calcular. O objetivo era só saber a estratégia que vocês iam usar pra fazer a manchete, aí você decidiu pela média.

Esse momento de reflexão dos professores não é analisado, pois não se está trabalhando com formação dos professores, mas sim com o seu raciocínio sobre variação, que confirma a explicação de Bakker (2004) sobre a falta de entendimento conceitual para analisar os dados.

O professor AM lê sua frase, que depois foi apresentada ao grupo todo:

> AM: Aí, entre todas aquelas coisas lá, dos cento e cinco alunos entrevistados, <u>setenta têm nomes entre cinco e oito letras</u>, dos três grupos, aí, sendo que apenas os israelenses, <u>oito, aparecem entre duas e três letras</u>, e apenas brasileiro e norte-americanos, <u>doze, aparecem entre nove e onze letras.</u>

É possível observar que, a partir da observação da moda em cada distribuição e da criação de um intervalo composto pelos valores da variável que tinham, pelo menos, freqüência um em todas as distribuições (entre cinco e oito letras), o professor contou o número de alunos nesse intervalo (70 alunos que representa 67%).

A conclusão do professor AM assemelha-se com o resultado que obteria se tivesse encontrado o intervalo de um desvio padrão da média, que neste caso, era $[6{,}23 \pm 2{,}08]$, permitindo estimar a porcentagem de observações.

Seu raciocínio pode ser classificado como de procedimento, pois utilizou as dimensões do conceito, mas não fez nenhuma relação com outras medidas de variação que permitiriam análises semelhantes.

### 11.9 *Feedback*

No dia 26 de Agosto, último dia da formação continuada, alguns aspectos observados pela PSQ foram discutidos: diferença entre descrição e análise, entre a observação da maioria e da minoria e as medidas estatísticas utilizadas pelos professores para a elaboração da análise.

Para discutir sobre a observação da maioria ou da minoria, a PSQ lembrou a atividade realizada sobre a disciplina de preferência e gênero.

> PSQ: Vocês lembram o dia que a gente tava fazendo uma atividade aqui, que eu trouxe aquele banco de dados da minha aluna sobre o teorema de Pitágoras? E aí eu me lembro da LF e do AM que estavam bem naquele canto, fazendo aqueles gráficos, lembram-se disso? E o AM falava assim pra LF: “Bom, a gente pode falar que a maioria, nanananana....” a LF falava: “Então tá bom, as meninas gostam mais de Biologia” só que eles nem perceberam que eles estavam discutindo coisas diferentes. Eles estavam tão preocupados em montar o gráfico... O AM estava dando uma idéia bem bonitinha de “Vamos dar a cara da maioria”, e a LF estava generalizando que todos eram assim. O grande cuidado pra estatística é que não é todo mundo, é a maioria.
>
> LH: Nenhum dos dois se tocou que eram coisas diferentes.

A PSQ fez uma analogia com o relatório que eles iriam elaborar: “Imagine vocês fazendo um relatório pra escola, por que eu estou falando isso? Vocês vão olhar e dizer assim... Vocês vão dar alguma informação pra escola se vocês disserem assim: “A minoria dos professores tem mais de 50 anos”. Que vocês estão informando pra escola?

A partir desta intervenção, foi discutido que maioria referia-se à moda, mas que outras medidas estatísticas poderiam ser utilizadas, tais como a média e o desvio padrão, haja visto que se tratava de uma variável quantitativa discreta.

> PSQ: Eu poderia tirar o desvio padrão e...Lembra que a gente discutiu? Se a gente fizer a média, mais ou menos um desvio, a gente já tem mais que a metade, a gente tem entre 60% e 80% dos dados, quer dizer, já dou uma grande maioria também. Em geral, o apelo era visual mesmo, não era pra medida, mas se identificou que poderia calcular a medida,

né...e claro que eu não dei tempo pra isso, mas se identificou, poderia ter feito isso.

Não houve questionamento nem espanto dos professores sobre o fato da possibilidade de utilizar a média e o desvio padrão. Isto pode ter duas explicações. A primeira, otimista, é que seria uma alternativa à análise que fizeram e que eles conseguiam mobilizar estes conceitos. A segunda explicação, pessimista, é que eles não se lembravam da possibilidade de elaborar um intervalo de *n* desvios padrão da média.

A professora RN verbalizou sua dificuldade em elaborar o relatório para a escola e apresentou a análise que tinha feito sobre as notas de importância da Estatística para o cotidiano do professor, para sua área de formação, para sua disciplina e para o seu aluno, reproduzidas nas Tabelas 29 e 30.

As tabelas são semelhantes à tarefa dos sobrenomes, mas a professora não fez a analogia entre as tarefas, que pode ser explicado pelo fato de não ter participado do desenvolvimento da atividade, somente do *feedback.*

Tabela 29 – Representação elaborada pela professora RN para as notas de importância da Estatística para o cotidiano do professor e para sua área de formação.

| Nota | F.A. | F.R.% |
|---|---|---|
| 0 | 2 | 2,30 |
| 2 | 1 | 1,15 |
| 3 | 1 | 1,15 |
| 4 | 4 | 4,60 |
| 5 | 14 | 16,09 |
| 6 | 9 | 10,34 |
| 7 | 13 | 14,94 |
| 8 | 17 | 19,54 |
| 9 | 4 | 4,60 |
| 10 | 19 | 21,84 |
| Não resp. | 3 | 3,45 |
| Total | 87 | 100% |

| Nota | F.A. | F.R.% |
|---|---|---|
| 0 | 1 | 1,15 |
| 3 | 1 | 1,15 |
| 4 | 1 | 1,15 |
| 5 | 12 | 13,79 |
| 6 | 6 | 6,90 |
| 7 | 6 | 6,90 |
| 8 | 15 | 17,24 |
| 9 | 8 | 9,19 |
| 10 | 34 | 39,08 |
| Não resp. | 3 | 3,45 |
| Total | 87 | 100,1 |

Tabela 30 – Representação elaborada pela professora RN para as notas de importância da Estatística para a disciplina do professor e para seu aluno.

| Nota | F.A. | F.R.% |
|---|---|---|
| 0 | 1 | 1,15 |
| 3 | 2 | 2,30 |
| 4 | 1 | 1,15 |
| 5 | 15 | 17,24 |
| 6 | 5 | 5,74 |
| 7 | 8 | 9,19 |
| 7,5 | 1 | 1,15 |
| 8 | 16 | 18,40 |
| 9 | 5 | 5,74 |
| 10 | 30 | 34,48 |
| Não resp. | 3 | 3,45 |
| Total | 87 | 99,99 |

| Nota | F.A. | F.R.% |
|---|---|---|
| 0 | 2 | 2,30 |
| 1 | 1 | 1,15 |
| 3 | 3 | 3,45 |
| 5 | 15 | 17,24 |
| 6 | 17 | 19,54 |
| 7 | 5 | 5,74 |
| 8 | 20 | 22,99 |
| 9 | 3 | 3,45 |
| 10 | 27 | 31,03 |
| Não resp. | 4 | 4,60 |
| Total | 87 | |

As Tabelas 29 e 30 merecem alguns comentários. A professora RN colocou os títulos das colunas, o que mostra a clareza sobre os conceitos de freqüência absoluta e porcentagem. Colocou os totais, que não dizem respeito aos cento de dez entrevistados, pois ela estava com um banco de dados incompleto. Mas permaneceu com sua dúvida sobre a soma das porcentagens não produzir 100%, assunto discutido num encontro que ela faltou.

Outro aspecto importante a ser observado é que permanecia a tendência de se trabalhar com distribuição de freqüência simples e não com dados agrupados, o que sugere uma repulsão aos trabalhos com variáveis contínuas, mesmo havendo na amostra um professor que tenha atribuído nota sete e meio.

### 11.10 Análise do raciocínio de variação/variabilidade na representação de uma distribuição de freqüência simples.

Na primeira tarefa houve a intenção dos professores LH, RN, AM e LF em analisar a representação que elaborariam, enquanto que os outros participantes restringiram-se em realizar a tarefa tal como imaginavam que seu aluno a faria.

O professor AM identificou que a maioria das meninas gostava de Biologia e, esta maneira de elaborar sua análise leva em consideração a variabilidade envolvida, pois a utilização do termo maioria não exclui a possibilidade de que outras alunas gostem de outras disciplinas, o que já foi relatado por Ben-Zvi (2004).

A análise elaborada pela professora LF (as meninas gostam mais de Biologia) poderia indicar que esta professora levava em consideração a variabilidade, mas quando se observa atentamente a maneira como foi denominada a variável "disciplina que mais gosta", entende-se que o termo "gostam mais" não se refere à possibilidade de alunos gostarem de outras disciplinas, mas que as alunas gostam de Biologia.

Quanto à segunda tarefa realizada, diferentes estratégias de análises foram observadas e a síntese da classificação do nível de variação está apresentada nos Quadros 30 e 31.

Quadro 30 – Raciocínio idiossincrático e verbal sobre variação utilizado na análise de três distribuições de freqüência simples

| Raciocínio sobre variação | Prof. | Verbalização | Nível de raciocínio sobre variação |
|---|---|---|---|
| Percepção da variação das freqüências na categoria da moda | LH | LH: Quem está mais homogêneo nessa história toda?<br>RN: Com seis letras.<br>Desvio grande [em relação as freqüências dos nomes com 5 letras] | Idiossincrático |
| Percepção da variação das freqüências em cada número de letras | LH | Não. Americanos. Tá mais bem distribuído. | Idiossincrático |
| Covariância | AM | Quando você reduz a quantidade de alunos, diminui a quantidade de letras... | Idiossincrático |
| Utilização da moda em cada distribuição | LH, RS, AM, OB | É melhor pegar os israelenses 31%, com quatro letras, 29% os americanos com seis letras, e a de brasileiros com cinco letras, 51%. | Verbal |
| Utilização de uma parte da distribuição | RS, LH e AM | Escolha de uma parte da distribuição. Entre cinco e oito. Entre cinco e seis. | Verbal |

Quadro 31 – Raciocínio de transição e de procedimento sobre variação utilizado na análise de três distribuições de freqüência simples

| Raciocínio sobre variação | Prof. | Verbalização | Nível de raciocínio sobre variação |
|---|---|---|---|
| Moda mais a observação dos valores máximos e mínimos | RS<br>AM | Você vai colocar aqui que está observando que o menor número de letras está no nome dos alunos israelenses, e o maior número de letras do povo americano? | De transição |
| Percepção intuitiva de *outliers* e da moda | AM<br>RS | Minoria dos brasileiros e americanos tem um número de letras entre 11 e 10, certo | De transição |
| Elaboração de um intervalo | AM<br>RS | AM: Olha, brasileiros e americanos aparecem entre quatro e onze letras, né? Ficam entre quatro e oito. Os brasileiros não aparecem entre dois e três. Os três aparecem entre cinco e oito.<br>RS: As três [distribuições] giram em torno de cinco, seis letras. | De transição |
| Contagem de observações em um intervalo | AM | Aí, entre todas aquelas coisas lá, dos cento e cinco alunos entrevistados, setenta têm nomes entre cinco e oito letras, dos três grupos, aí, sendo que apenas os israelenses, oito, aparecem entre duas e três letras, e apenas brasileiro e norte-americanos, doze, aparecem entre nove e onze letras. | Procedimento |

Nenhum professor sugeriu a utilização do desvio padrão, conceito já discutido e exemplificado na formação continuada. A utilização da média só foi sugerida como uma última alternativa, quando nada mais resolvia o problema. Ou seja, a média e o desvio padrão não foram conceitos mobilizados para fazer a comparação de três grupos.

Em vez disso, a partir da moda de cada grupo, das categorias com freqüência em todos os grupos ao mesmo tempo e dos valores máximo e mínimo eles criaram um intervalo de variação, que continha a maior quantidade de pessoas (setenta dentre cento e cinco) e que pode ser considerado uma quebra da distribuição, que também foi observado no raciocínio da professora RS.

Vale ressaltar que a estratégia utilizada pela professora RS para quebrar a distribuição mobilizou menor quantidade de conceitos do que o professor AM.

Outro destaque sobre a análise dos professores é que nenhum sugeriu ou elaborou um gráfico, conteúdo tão enfatizado por eles.

A seqüência do raciocínio dos professores foi semelhante à dos alunos de Ben-Zvi (2004) e os termos mais utilizados se assemelham aos do estudo de Makar e Confrey (2005), exceção feita à média e ao desvio padrão.

Esses resultados permitem relacionar a dificuldade apresentada pelos professores com o conteúdo priorizado pelos livros didáticos. A tabela de distribuição de freqüências é utilizada pelos livros como uma estratégia para a realização do gráfico, que não foi feito pelos professores, e não para realizar análises, como pode ser observado nas Tarefas 1 e 2 da análise da organização matemática dos livros (Capítulo 3 deste trabalho).

## 12 Raciocínio sobre variação/variabilidade na representação de variáveis contínuas

De posse do banco de dados da pesquisa realizada pelos próprios professores (Apêndice 7) e de acordo com os resultados obtidos na representação de variáveis qualitativas, em que foi observado que a maioria dos professores tinha pouco conhecimento das técnicas para a elaboração de gráficos, no dia 20 de Maio foi iniciada a discussão sobre distribuição de freqüências com dados agrupados e outras estratégias para representar variáveis contínuas. A tarefa solicitada está apresentada na Figura 68.

| |
|---|
| PSQ: Pensando no relatório que será realizado, vamos elaborar uma representação gráfica das variáveis *idade* e *tempo de magistério* dos professores. |

Figura 68 – Tarefa 1 da análise dos dados do ciclo investigativo – representação de variáveis contínuas

Devido à dificuldade encontrada pelos professores em relação à representação das variáveis gênero e disciplina de preferência, não houve tempo para discussão sobre a representação da idade dos participantes da pesquisa de Nifoci e Silva (2004), que seria discutida neste encontro, mas com a utilização da idade dos professores que responderam ao questionário.

Para a realização da tarefa solicitada, os professores dividiram-se em quatro grupos.

### 12.1 Resultados do Grupo 1

O Grupo 1, composto pelos professores AM e SB, iniciou a atividade observando a menor e a maior idade e a grande diversidade das observações, o que levou-os, imediatamente, à elaboração de faixas etárias e a contagem das pessoas em cada uma dessas faixas. Os resultados deste grupo estão apresentados no Quadro 32.

Quadro 32 – Representação elaborada pelos professores AM e SB para idade e tempo de magistério

| Tempo de Magistério | |
|---|---|
| 1 – 5 | 24 |
| 6 – 10 | 28 |
| 11 – 15 | 24 |
| 16 – 20 | 21 |
| 21 – 25 | 6 |
| 26 ou mais | 5 |
| brancos | 2 |

| Idade | |
|---|---|
| 21 – 25 | 12 |
| 26 – 30 | 12 |
| 31 – 35 | 15 |
| 36 – 40 | 22 |
| 41 – 45 | 21 |
| 46 – 50 | 9 |
| 51 – 55 | 11 |
| 56 – 60 | 5 |
| Não respondeu | 3 |
| | 110 |

Os professores não tiveram tempo de elaborar a representação gráfica e, apenas verbalizaram que elaborariam o gráfico de barras.

Quando questionado pela PSQ sobre a justificativa pela utilização de classes de amplitude cinco, o professor AM explicou que sua intenção era deixar, aproximadamente, vinte professores em cada classe, que refere-se à um raciocínio idiossincrático de variação. Raciocínio semelhante já foi observado no capítulo anterior, quando os professores LH e RS buscavam uma distribuição mais uniforme de freqüências em uma categoria.

### 12.2 Resultados do Grupo 2

A estratégia utilizada pelas professoras OB e IS levou muito tempo para ser realizada, pois elas contaram o número de professores com cada idade, ou seja, fizeram uma contagem para elaborar a distribuição de freqüências simples.

Quando terminaram a contagem, perceberam que os professores AM e SB haviam elaborado uma distribuição de freqüência com dados agrupados. Em dúvida sobre qual técnica era mais adequada, permaneceram contando o número de professores para cada ano diferente de magistério.

### 12.3 Resultados do Grupo 3

O Grupo 3, composto pelos professores RN, RS e LH, fez a contagem da idade e do tempo de magistério por faixas, calculando a porcentagem de cada grupo, cuja estratégia foi sugerida pela professora RN. A representação deles está apresentada no Quadro 33.

Quadro 33 - Representação elaborada pelos professores LH, RS e RN para idade e tempo de magistério.

| Idade | | | |
|---|---|---|---|
| 20 - 25 | | 12 | 11% |
| 26 – 30 | | 12 | 11% |
| 31 – 35 | | 15 | 14% |
| 36 – 40 | | 22 | 20% |
| 41 – 45 | | 21 | 19% |
| 46 – 50 | | 9 | 8% |
| 51 – 55 | | 11 | 10% |
| 56 – 60 | | 5 | 7% |
| | | 108 | 100% |
| **Tempo de magistério** | | | |
| 1 – 5 | | 24 | 22% |
| 6 – 10 | | 28 | 26% |
| 11 – 15 | | 24 | 22% |
| 16 – 20 | | 19 | 18% |
| 21 – 25 | | 7 | 7% |
| 26 – 30 | | 5 | 5% |
| | | 107 | 100% |

Com a intenção de relacionar as variáveis idade e tempo de magistério, o professor LH sugeriu elaborar uma tabela de dupla entrada (sem utilizar este termo), mas o grupo rejeitou imediatamente, justificando que a experiência anterior (com o gráfico de setores) havia sido frustrante.

Vale ressaltar que estratégia semelhante foi utilizada pelos professores do estudo de Hammerman e Rubin (2004), quando relacionaram a idade mediana dos eleitores e a porcentagem de votos no Bush observados nos estados americanos.

O grupo explicou que fariam um gráfico de barras para cada variável e a PSQ perguntou como analisariam os gráficos. O diálogo seguinte apresenta a resposta da professora RN.

> RN: Pegava o maior pico de idade, aí estudaria qual a faixa etária dali: qual é o pico de magistério, e ia estudar. Olharia os picos. Porque o jeito que o LH quer fazer é uma loucura!
>
> LH: Não dá assim no manual, mas você joga no computador...e faz tudo!

É possível observar que a professora RN apresentava uma idéia de analisar o pico, ou seja, a classe modal, enquanto o professor LH acreditava que o computador resolveria o seu problema.

Assim como os grupos anteriores, é possível observar que estes professores fizeram a contagem, mas não colocaram no cabeçalho da "tabela" a que se referia, levantando a dúvida se haviam adquirido o conceito de distribuição de freqüências.

Os professores dos Grupos 1, 2 e 3 que elaboraram uma distribuição de freqüência com dados agrupados, fizeram-na como se idade e tempo de magistério fossem variáveis discretas e não contínuas, o que foi discutido no momento de *feedback*.

### 12.4 Resultados do Grupo 4

Este grupo foi composto pela professora CI e pelos alunos AG e LJ, que tinham aprendido este conteúdo na graduação, no mês anterior.

Os alunos começaram a resolver a tarefa, sem questionar se havia outra maneira. A professora ocupou a posição de observadora dos alunos. A representação da variável idade elaborada pelo grupo está apresentada no Quadro 34 e na Figura 69.

Quadro 34 - Representação elaborada pelos alunos AG e LJ e pela professora CI para idade

| idade | Freqüência | Freqüência relativa | Densidade | Porcentagem (%) |
|---|---|---|---|---|
| 19 ⊢ 25 | 9 | $\frac{9}{108} = 0{,}0833$ | $\frac{9}{648} = \frac{9}{108} \cdot \frac{1}{6} = \frac{9}{108} \cdot \frac{5}{30} = \frac{45}{3240}$ | 8,3 |
| 25 ⊢ 30 | 15 | $\frac{15}{108} = 0{,}1388$ | $\frac{15}{540} = \frac{15}{108} \cdot \frac{1}{5} = \frac{15}{108} \cdot \frac{6}{30} = \frac{90}{3240}$ | 13,8 |
| 30 ⊢ 35 | 14 | $\frac{14}{108} = 0{,}1296$ | $\frac{14}{540} = \frac{14}{108} \cdot \frac{1}{5} = \frac{14}{108} \cdot \frac{6}{30} = \frac{84}{3240}$ | 12,9 |
| 35 ⊢ 40 | 22 | $\frac{22}{108} = 0{,}2037$ | $\frac{22}{648} = \frac{22}{108} \cdot \frac{1}{5} = \frac{22}{108} \cdot \frac{6}{30} = \frac{132}{3240}$ | 20,3 |
| 40 ⊢ 45 | 20 | $\frac{20}{108} = 0{,}1851$ | $\frac{20}{540} = \frac{20}{108} \cdot \frac{1}{5} = \frac{20}{108} \cdot \frac{6}{30} = \frac{120}{3240}$ | 18,5 |
| 45 ⊢ 50 | 10 | $\frac{10}{108} = 0{,}0925$ | $\frac{10}{540} = \frac{10}{108} \cdot \frac{1}{5} = \frac{10}{108} \cdot \frac{6}{30} = \frac{60}{3240}$ | 9,2 |
| 50 ⊢ 55 | 12 | $\frac{12}{108} = 0{,}1111$ | $\frac{12}{540} = \frac{12}{108} \cdot \frac{1}{5} = \frac{12}{108} \cdot \frac{6}{30} = \frac{72}{3240}$ | 11,1 |
| 55 ⊢⊣ 60 | 6 | $\frac{6}{108} = 0{,}0555$ | $\frac{6}{648} = \frac{6}{108} \cdot \frac{1}{6} = \frac{6}{108} \cdot \frac{5}{30} = \frac{30}{3240}$ | 5,5 |
| | 108 | | | 100% |

O aluno LJ calculou as frações equivalentes (quarta coluna do Quadro 34) e a aluna AG não conseguia entender o que e nem porque ele estava fazendo. Ele explicou para PSQ que isto era muito fácil e ajudava a elaboração do histograma.

Quando questionado sobre as cores diferentes em cada barra, o aluno LJ disse que isto era “obra da AG” e que ele preferiria não colocar cor nenhuma.

Ao término do encontro, este grupo havia elaborado a distribuição de freqüências com dados agrupados e o histograma tanto para a idade quanto para o tempo de magistério, mas em momento algum se preocuparam em analisar o que estavam fazendo.

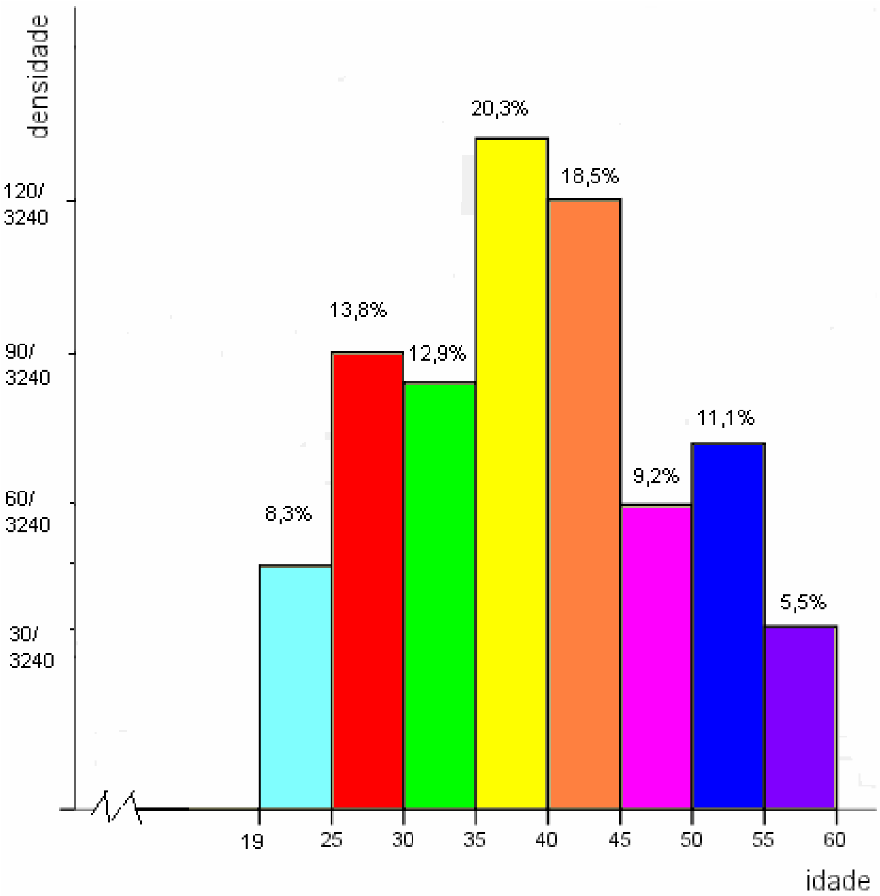



Figura 69 – Representação gráfica elaborada pelos alunos AG e LJ e pela professora CI para idade

## 12.5 *Feedback*

O *feedback* aconteceu no encontro seguinte (03 de Junho), em que foram reproduzidas na lousa todas as representações elaboradas pelos grupos.

A primeira discussão foi sobre os diferentes gráficos e distribuições de freqüências com dados agrupados que poderiam ser elaborados para representar a idade.

Os professores explicaram que elaboraram as faixas de uma maneira que consideraram mais fácil, sem levar em consideração a variação.

Vários aspectos sobre a elaboração da distribuição de freqüências com dados agrupados foram formalizados:

- adequada para variável contínua – barras contíguas;
- diferentes notações de intervalo - $[a,b[$ ou a $\vdash\!\!\!-\!\!-$ b;

- classes com amplitudes iguais ou diferentes;
- sugestões para a obtenção do número de classes (Sturges e regra empírica);
- amplitude total como uma medida de variação necessária;
- a não necessidade de colorir as barras;
- diferentes escalas nos eixos das ordenadas e das abscissas;
- utilização de freqüência absoluta, porcentagem ou densidade de freqüência;
- quebra de escala no eixo das abscissas;
- normas da ABNT e do IBGE para a elaboração de tabelas e gráficos, etc.

Devido à insistência do professor LH sobre a possibilidade do Microsoft Excel elaborar "sozinho" o histograma, foi salientado que isto não acontece, sendo necessário que o usuário elabore a distribuição de freqüências, digite-a na planilha, selecione a opção gráfico de barras e realize os ajustes necessários ao gráfico.

Foi discutido sobre o fato de que a pesquisa idealizada e realizada por eles (professores participantes) não suscitava a necessidade do histograma, pois eles mesmos estavam satisfeitos com a distribuição de freqüências simples para a idade e tempo de magistério, que daria origem ao gráfico de barras.

Por este motivo, foi reforçada a questão do monitoramento da elaboração do questionário, variável didática importante quando da utilização de pesquisa como metodologia de ensino de Estatística.

Na lousa, a PSQ elaborou o gráfico de barras para a idade de acordo com a distribuição de freqüências elaborada pelos professores e reproduziu o histograma elaborado pelo Grupo 4.

A professora IS preocupou-se com a precisão dos resultados. O diálogo seguinte ilustra esta preocupação.

> IS: <u>Quanto menor o intervalo, mais preciso será os resultados</u>? Se fôssemos contar de dez em dez, íamos perder um pouco dessa informação, não é?
>
> PSQ: Qual seria a pergunta que vocês fariam para ter o histograma como resposta da idade? Vocês não vão apresentar um relatório lá para

> escola? A gente quer saber o quê? Qual a idade dos professores: será que são jovenzinhos, de meia idade, etc?

Possivelmente a professora IS estava se referindo à precisão das medidas que seriam calculadas a partir da distribuição de freqüências, mas isto não foi perguntado.

Foram elaboradas diferentes distribuições de freqüências da idade para que os professores pudessem observar que não existia uma única maneira de representação.

Mesmo tendo todas as distribuições na lousa, nenhum professor fez menção para analisar. Eles estavam muito preocupados com o aspecto técnico da tarefa “elaborar a representação” em vez de se preocuparem com a tarefa “analisar a idade dos professores”.

Na formalização de diferentes tipos de variável, alguns professores confundiram variável quantitativa com a freqüência (ou a porcentagem) de uma variável qualitativa. Para eles, “era tudo número”! Esta dificuldade se repetiu na tarefa do penúltimo dia da formação (já apresentada o capítulo anterior).

Para sanar esta dificuldade, foi simulado um exemplo de pesquisa sobre as taxas de juros cobradas pelos bancos, variável quantitativa contínua, cujas respostas seriam em porcentagem e o exemplo da variável gênero, qualitativa nominal, cujas respostas seriam homens e mulheres.

Como o grupo de alunos havia trabalhado a densidade de freqüências, a PSQ elaborou outro exemplo para formalizar este assunto com os professores, que está apresentado na Tabela 31.

Tabela 31 – Distribuição de freqüência com dados agrupados da variável idade

| Idade dos professores | freqüência | h | Densidade de freqüência |
|---|---|---|---|
| 19 $\vdash$ 29 | 21 | 10 | $\frac{21}{10} = \frac{147}{70}$ |
| 29 $\vdash$ 34 | 10 | 5 | $\frac{10}{5} = \frac{140}{70}$ |
| 34 $\vdash$ 39 | 22 | 5 | $\frac{22}{5} = \frac{308}{70}$ |
| 39 $\vdash$ 44 | 22 | 5 | $\frac{22}{5} = \frac{308}{70}$ |
| 44 $\vdash$ 49 | 14 | 5 | $\frac{14}{5} = \frac{196}{70}$ |
| 49 $\vdash$ 54 | 11 | 5 | $\frac{11}{5} = \frac{154}{70}$ |
| 54 $\vdash\dashv$ 60 | 8 | 7 | $\frac{8}{7} = \frac{80}{70}$ |
| Total | 108 | | |

Com a distribuição de freqüência da Tabela 31, foi elaborado o histograma apresentado na Figura 70 e discutido que cada barra representa uma área cuja base é a amplitude do intervalo de classe e a altura é a densidade de freqüência (absoluta ou relativa). Foi explicado que se a freqüência for absoluta, a soma das áreas deve ser igual ao total de freqüências e se a freqüência for relativa, a soma das áreas deve ser igual a um.

Esta relação do histograma com o conceito de área foi novidade, pelo menos, para a professora SB, que em conversa informal após o desligamento do gravador afirmou: “Eu nunca tinha relacionado o histograma com área”.

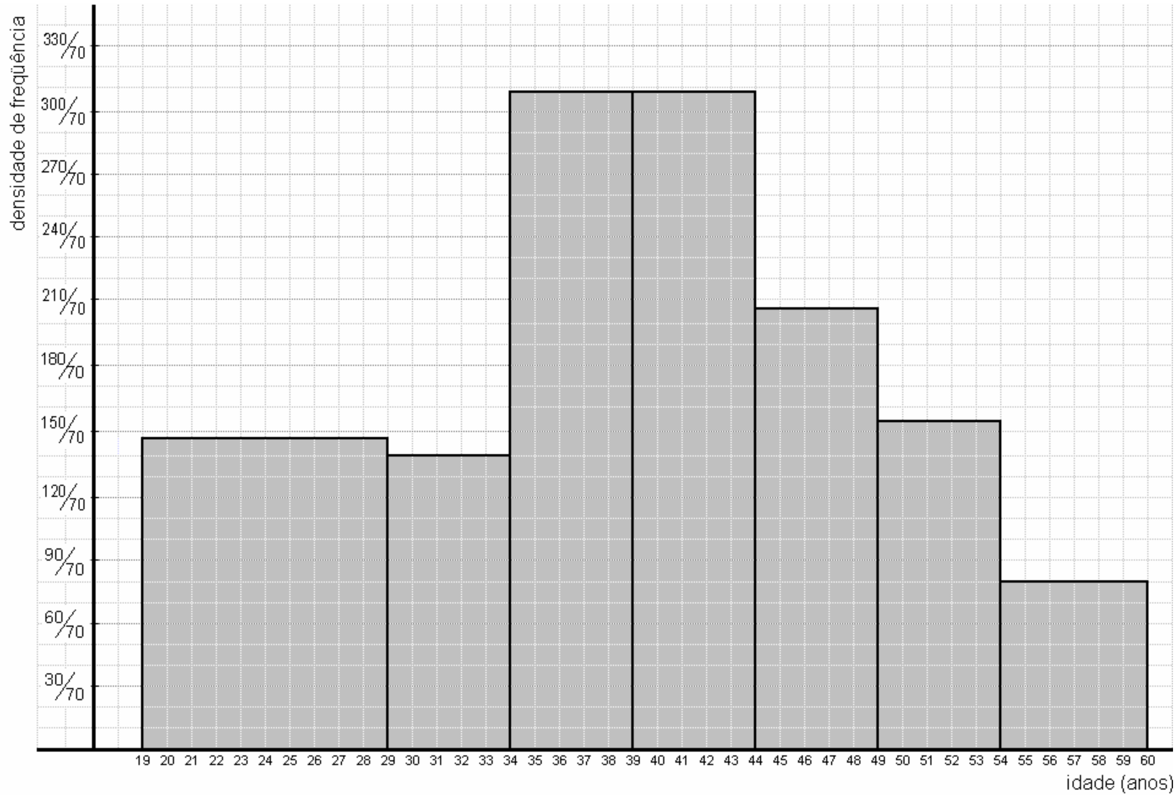



Figura 70 – Histograma elaborado para idade dos professores entrevistados, em que as classes tinham amplitudes diferentes.

Ao término da discussão sobre as questões técnicas de elaboração do gráfico, a PSQ perguntou aos professores qual é o objetivo deste gráfico e não houve nenhuma resposta.

Devido ao tipo de reação do professores, que ficaram quietos e com uma postura de indignação, a PSQ teve a impressão de que todas as questões discutidas eram novas para eles.

A professora IS perguntou sobre a possibilidade de representar a idade dos professores com um gráfico de segmento (linha) e, na discussão, a PSQ explicou sobre a adequação deste tipo de gráfico para variável acompanhada ao longo do tempo. Aproveitando essa oportunidade, a PSQ esboçou o diagrama de ramo-e-folhas e o gráfico de pontos (dispersão) para a idade dos professores pesquisados e, devido ao grande número de observações, solicitou que fizessem em casa essa tarefa.

Percebendo esta novidade para os professores, a PSQ solicitou duas atividades para serem desenvolvidas durante a semana:

Atividade 1: Foi solicitado, novamente, uma pesquisa sobre diferentes maneiras de representar uma variável quantitativa contínua.

Atividade 2: Foi solicitado que priorizassem esta pesquisa: “como escrever uma análise sobre a idade dos professores em apenas uma linha?

## 12.6 Análise de uma distribuição de freqüências com dados agrupados

Tal como aconteceu com a distribuição de freqüências simples, o momento do *feedback* da distribuição de freqüências com dados agrupados priorizou sua elaboração e não sua interpretação.

Logo, no penúltimo dia da formação continuada (19 de Agosto), foi apresentada a tarefa descrita na Figura 71 e disponibilizados os gráficos da Figura 72, os mesmos utilizados por Meletiou (2000).

| |
|---|
| PSQ: Os gráficos a seguir representam as notas na disciplina Matemática de duas turmas. Qual turma apresenta maior variação? |

Figura 71 - Tarefa 2 da análise dos dados do ciclo investigativo – representação de variáveis contínuas

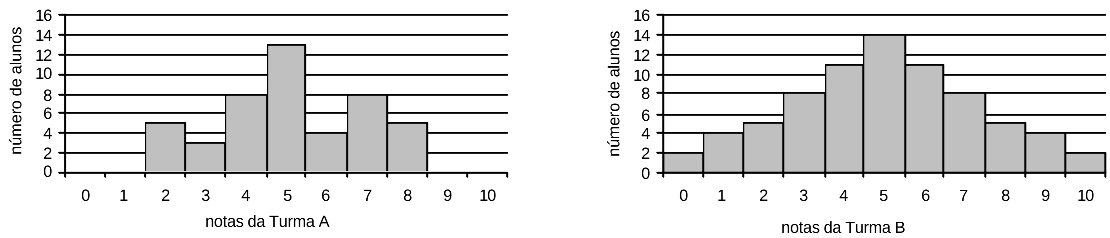



Figura 72 – Gráficos fornecidos para a realização da Tarefa 2 da análise dos dados do ciclo investigativo – representação de variáveis contínuas

## 12.7 Resultados do grupo 1 – professores RS e LH

A professora RN não pode participar deste encontro e isto desestruturou os professores RS e LH, que apresentaram dificuldades de diferentes ordens com a tarefa.

O grupo preferiu elaborar a distribuição de freqüências para cada distribuição, a partir das quais calcularam a média aritmética da seguinte maneira: somaram os limites inferiores de cada classe e dividiram pelo número de classes.

Tanto o conceito quanto o cálculo já tinham sido discutidos em encontros anteriores, mas as dificuldades também se apresentaram na compreensão da mediana e da moda.

### 12.8 Resultados do grupo 2 (AM e SB) e grupo 3 (OB e CI)

Como eles já tinham calculado a média no exercício das notas dos meninos e das meninas (apresentado no capítulo seguinte), eles imediatamente tomaram a decisão de usá-la novamente.

O professor AM verbalizou sua dificuldade em ler o histograma e realizou a tabela de distribuição de freqüências, a partir da qual calculou a média, de maneira correta e rápida.

Como o objetivo da PSQ era verificar a estratégia de raciocínio que utilizariam para decidir a distribuição com maior variação, fez uma intervenção.

> PSQ: Sem fazer conta [média], qual tem maior variação?
>
> AM: No visual acho que a média aqui deve cair um pouco pra direita [referindo-se à Turma A].
>
> OB: No visual, o de cima [Turma A].
>
> AM: Por causa desse aqui, desses dois, por isso que a média cai um pouquinho pra cá.
>
> PSQ: Com a média vocês conseguiriam detectar qual é a maior variação? Só com a média?
>
> OB: Não. A moda aqui é quatro e meio. Entre quatro e meio e cinco e meio, né?
>
> PSQ: É.
>
> OB: E aqui embaixo é quatro e meio. Eita! [descobriu que a classe modal era a mesma nos dois histogramas]

Eles perceberam que tanto a média quanto a moda eram muito semelhantes e ficaram em silêncio por alguns minutos para decidir sobre a distribuição. Ou seja, a estratégia utilizada na análise da distribuição dos sobrenomes não seria útil nesta atividade.

> PSQ: Então vocês sabiam que medidas iam calcular?
>
> AM: Sim. Logo de cara é... o desvio vai ter que aparecer, aliás, nesse aqui o desvio vai ter que aparecer, porque senão...
>
> PSQ: E aí você poderia trabalhar com seu aluno, o desvio padrão e mostrando os gráficos pra detectar maior variação. Porque a gente, invertendo a ordem dos fatos, a gente vê o primeiro gráfico e pergunta qual tem maior variação. Vocês calculariam o desvio, não é? O que

> interfere no desvio? O que torna o desvio muito grande ou muito pequeno?
>
> OB: <u>A variação.</u>
>
> AM: <u>A variação da média</u>

Novamente a ansiedade da PSQ em que eles analisassem as distribuições fazendo o uso do desvio padrão atrapalhou o diálogo e, conseqüentemente, a observação do raciocínio que utilizariam. As respostas “variação” e “variação da média” foram automáticas, pois sabiam que era a resposta esperada, mas possivelmente não apreenderam o que estavam verbalizando. A continuação do diálogo ilustra este fato.

> PSQ: Muito bem, então vocês só têm esse recurso gráfico. Não é isso? Sem calcular, vocês vai observar o desvio cada vez maior, enquanto os dados estão mais diferentes, mais distantes da média, não é isso? Se vocês olharem aí, usando só esse recurso...[essa discussão tinha acontecido minutos antes, na realização da atividade das notas dos meninos e das meninas – que estão apresentadas no capítulo seguinte].
>
> SB: AM, se você pensar em distante da média, se a média deu cinco e nós temos aqui 0,5, esse tá mais distante da média.
>
> AM: Nunca parei pra pensar... eu penso sempre na média, mas não penso no desvio padrão, na conta de luz. Nunca parei pra pensar no desvio padrão. Então, o que nós observamos aqui é que eles estão mais próximos da média, lá de cima é cinco e meio, e aqui como se faz o zero até o 10, então, o desvio padrão, aqui [Turma A], é menor do que aqui [Turma B].
>
> PSQ: Aqui você não tem os extremos e sim na hora que você vai ficando cada vez mais longe da média, maior é o desvio padrão.
>
> AM: Nunca abri a boca pra falar sobre isso, fazia o cálculo pra depois fazer o gráfico, não parava pra pensar nos dois.

A professora SB ficou afastada devido ao nascimento de sua filha e não participou dos dois últimos encontros de Junho, quando foram discutidas as medidas de variação.

Mas, a intervenção da PSQ foi suficiente para que ela olhasse a amplitude da variável e percebesse que a Turma A tinha a amplitude menor que a Turma B, uma das dimensões desse conceito.

## 12.9 *Feedback*

Embora esta atividade tenha sido realizada no final do encontro do dia 19 de Agosto, seria necessário dar um retorno aos professores, porque eles elegeram-na a mais difícil.

A PSQ resolveu abordar os seguintes aspectos: leitura do histograma e/ou da tabela de distribuição de freqüências, o relacionamento das medidas de tendência central e dispersão com o histograma, a relação entre este gráfico e o formato da distribuição.

Foi feita uma síntese do artigo de Meletiou e Lee (2002) e foi explicado que o histograma é um gráfico que gera dificuldade para o leitor e que, por este motivo, ele não pode ser negligenciado no ensino. Foi verbalizado que eles, professores, enfatizavam muito o gráfico de barras e pouco trabalhavam o histograma.

Os professores verbalizaram que se calculassem o desvio padrão, teriam condições de identificar a distribuição que apresentava maior variação, demonstrando uma intenção de não analisar o conceito, apenas calcular.

> PSQ: Ah é, eu falei: “Então calcula aí”, foi a hora que terminou [referindo-se ao encontro anterior]
>
> OB: Isso.
>
> PSQ: Então calcula pra ver se bate à percepção de vocês com o desvio padrão. E aí, qual é o objetivo do exercício? Quando a gente vê esse histograma, que é um gráfico que não estamos acostumados a olhar, a gente usa bem menos do que os outros gráficos, então, a tendência que a gente tem é olhar, é dizer que o gráfico um [Turma A] tem maior variação. Por quê? Porque as colunas são de tamanhos... não dá aquela harmonia que... para gente é ....

Neste momento, a PSQ apresentou os resultados obtidos na revisão bibliográfica, em que havia uma tendência dos alunos de escolher distribuições simétricas em relação à média (semelhante à curva normal) como distribuições de menor variação e que nem sempre isto é verdade, como no exercício realizado.

Foi feita uma síntese do trabalho de Delmas e Liu (2005) e explicado os conceitos que precisariam estar disponíveis para a análise da variação em um histograma. O professor AM aproveitou a oportunidade para tirar uma dúvida.

> AM: Você falou em amplitude, cabe uma coisa, quando a palavra “amplitude” aparece pra mim já penso... Você tá falando amplitude,

amplitude, mas estamos olhando na horizontal! Isso nunca passa pela minha cabeça!

Neste momento da discussão, discutiu-se sobre a importância dos títulos dos eixos, fator fundamental para a leitura do gráfico, e foi retomado o exemplo utilizado pela professora SB no encontro anterior.

A professora SB explicou sua dificuldade em ler o histograma, pois ela estava acostumada a ler o gráfico da conta de luz, em que o eixo das abscissas continha o consumo (em kw/h) e o eixo das ordenadas continha os meses, dando condições de apresentar as diferenças entre um histograma e um gráfico de linhas.

E o debate sobre a leitura do histograma prossegue.

AM: Então, preciso ler o eixo...

PSQ: Isso mesmo. E aqui poderia ser a mesma situação, poderia ter um histograma com amplitude pequena, de um e meio até cinco e meio, mas todo mundo nas duas pontas. Não foi o que aconteceu aqui. A gente tem a amplitude menor e a concentração no centro, estão vendo que a concentração está no cinco, quatro?

AM: Posso falar uma coisa que me passou pela cabeça, agora? Se estiver errado, me corrija, por favor? O fato… As medidas… Funções, né? A gente padroniza “x” e “y”, e agora você começa a falar isso aqui, até que ponto eu tenho que necessariamente está assim... porque a gente fala... porque o aluno pergunta assim: “Mas professor, a variável dependente sempre na horizontal e independente sempre na vertical?” Eu digo: “Bom, é o padrão adotado”...

Devido à intervenção da orientadora desta pesquisa em um dos primeiros encontros, em que ela associou a distribuição de freqüências com uma função, o professor AM associou que, em cada eixo havia uma variável (x e y) e gostaria de saber a relação entre o gráfico de função e o histograma.

Para explicar este ponto, tanto a PSQ quanto a observadora VG apresentaram suas interpretações.

PSQ: Em estatística, a variável independente está no eixo das abscissas e a dependente está no eixo das ordenadas, porque se não, não faz sentido a análise, você tem duas variáveis: Uma variável dependente, e uma independente que a gente costuma chamar de “x”, “y”. Mas aqui você só tem uma variável, você tem uma variável que é nota. O outro eixo é uma contagem que é a freqüência, embora venha a idéia dessa contagem, a função. Você pode trocar os eixos aqui, isso não faz a menor diferença. Porém, se você trocar os eixos numa análise de regressão linear, ou seja, trocar x por y, quem está por trás é a função de primeiro grau e faz toda a diferença, dá o resultado trocado. Então aqui não! Você tem um gráfico que a gente chama de unidimensional, porque você tem uma única variável.

OBS VG: Se fizer um gráfico de barras assim, né? Cada nota vai ter exatamente quantos alunos tiraram aquela nota, eu não sei qual é qual.

PSQ: A resposta freqüência não é variável qualitativa, não é quantitativa, não é discreta, não é contínua. É freqüência, mas tem o conceito de função, porque dado uma nota, você tem uma freqüência.

OBS VG: Não tem uma expressão algébrica, mas você tem pra cada elemento do domínio uma única imagem.

Como esta foi a última discussão da formação continuada, talvez essa relação do professor AM do gráfico estatístico com o gráfico de função tenha induzido-o a confundir variável com freqüência, como relatado no capítulo anterior.

E o diálogo volta-se novamente para a análise.

PSQ: Então, o que tem maior variação é o "b" [Turma B], e depois façam como exercício o desvio padrão, variação em torno da média e confirmem o resultado.

OB: <u>Mais distante da média maior variação</u>.

PSQ: Eu perguntei pra eles o seguinte: "Tudo bem, como é que você calcula o desvio padrão?" Eles disseram: "Até quanto cada valor está distante da média". Então eu falei: "Ótimo, então comecem olhando as distâncias em relação à média", que é o que a gente tava fazendo, dando a amplitude de notas. O problema é que aí a gente não pode olhar só a amplitude de nota, porque podem ocorrer os casos extremados. Eu posso ter uma amplitude pequena, mas tá todo mundo nos extremos.

OB: <u>Pode ter uma amplitude menor e não ter um desvio padrão menor</u>.

Ambas verbalizações de OB parecem uma forma de raciocínio na busca de *links* mentais para a compreensão do conceito. Utilizando a teoria dos Campos Conceituais, é como se ela buscasse outras dimensões do conceito para que esta verbalização pudesse ser significativa.

PSQ: Sim, porque eu posso ter uma distribuição de dois a seis, e está todo mundo no dois e todo mundo no seis. Então, a amplitude é pequena, e tá todo mundo nos extremos, como é o caso do dez e sessenta [exercício apresentado no capítulo seguinte] . Agora eu posso ter uma amplitude pequena, e estarem todos concentrados no centro, e aí, certamente, o desvio padrão vai ser pequeno, porque é uma medida de variação em relação à média.

AM: Na estatística está dizendo que esse tem maior variação em relação ao outro? [ele estava intrigado, pois o número de alunos em cada distribuição era muito diferente]

PSQ: O que é cada coluna? O que é altura da coluna?

AM: Não, não, é exatamente, a altura da coluna. Se você comparar...

PSQ: Sim. Então, ele já tinha identificado que seria até difícil avaliar, porque aqui você tem menos alunos no A do que no B. E é por isso que eu não posso olhar só o "desenho", tenho que analisar tudo: a amplitude da nota, que é a variável que a gente está olhando, e como estão as alturas das colunas que essa é a contagem.

AM: É. Nós contamos tudo. Nós olhamos a quantidade de alunos dos dois gráficos, né?

É interessante observar a tendência do professor AM para o raciocínio proporcional, algo que ele não tinha observado na análise de uma distribuição de freqüência simples, pois estava preocupado com a diferença entre a quantidade de alunos na Turma A e B.

## 12.10 Análise do raciocínio de variação/variabilidade na representação de variável contínua

Esta atividade foi realizada após a comparação da variação das notas dos seis alunos e das seis alunas (apresentadas no capítulo seguinte) e, por este motivo, a estratégia de utilização da média aconteceu tão naturalmente.

Tal como aconteceu no estudo de Meletiou (2000) e Meletiou e Lee (2002), a grande dificuldade na tarefa foi a leitura do histograma, que nem mesmo a tarefa complementar foi suficiente para identificar o raciocínio sobre variação, pois as questões ainda estavam em um nível técnico, de leitura e não interpretação da informação.

Como explicam Delmas e Liu (2005), compreender a variação numa distribuição de freqüências requer a mobilização de muitos conceitos: amplitude da magnitude da variável (no histograma apresentado, a amplitude do eixo das abscissas), a altura das barras (freqüência de pessoas com cada valor), disposição das barras ao longo do eixo das abscissas, ou seja, a densidade de freqüência em torno da média, a visualização (imaginária) da média aritmética e dos desvios da média.

São muitos conceitos a serem mobilizados ao mesmo tempo, tornando a avaliação da variação uma tarefa extremamente complexa.

Como já apresentado no capítulo anterior, a moda (ou classe modal) é a medida que aparece intuitivamente e este fato serve como diagnóstico para o ensino de variação, que deve privilegiar situações em que a moda em duas (ou mais) distribuições sejam a mesma e situações em que sejam diferentes.

Tal como anteriormente, esse tipo de raciocínio foi categorizado como verbal, pois é correto, necessário, mas não suficiente.

As poucas verbalizações que permitiram inferir o raciocínio de variação estão apresentadas no Quadro 35, em que foram analisadas segundo o nivelamento proposto por Garfield (2002).

Quadro 35 – Nível de Raciocínio sobre variação na fase de análise de uma distribuição de freqüências com dados agrupados

| Raciocínio sobre variação | Professor | Verbalização | Nível do raciocínio sobre variação |
|---|---|---|---|
| Variação pequena entre as freqüências das classes de uma distribuição | AM | o professor AM explicou que sua intenção era deixar, aproximadamente, 20 professores em cada classe. | Idiossincrático |
| classe modal representa a maioria | RN | Para analisar o histograma, observaria o “maior pico” | Verbal |
| Desvio padrão grande devido à variação | OB | Respondendo à questão da PSQ | Transição |
| Desvio padrão grande devido à variação da média | AM<br>OB | Desvio padrão grande devido a variação em torno da média<br>Mais distante da média maior variação | Transição |
| Amplitude como um indicativo das distâncias em torno da média. | SB | se você pensar em distante da média, se a média deu cinco e nós temos aqui 0,5, esse tá mais distante da média. | Transição |

Um dos objetivos desta tarefa era relacionar o tipo de raciocínio de variação que os professores apresentariam com as justificativas dos alunos de Delmas e Liu (2005) para um desvio padrão maior.

Os professores utilizaram duas justificativas também utilizadas pelos alunos de Delmas e Liu: barras distantes da média e observação dos valores extremos da barra numérica.

É interessante observar que a tarefa anterior, que dizia respeito ao número de letras do sobrenome, proporcionou muito mais debate e utilização de medidas estatísticas do que essa tarefa, o que indica o nível de dificuldade com a leitura e compreensão do histograma.

A PSQ observou que o momento da formação continuada reservado para esse assunto foi exclusivamente para a discussão da tarefa "de elaborar o gráfico" e que, nem mesmo a inserção da segunda tarefa "de analisar o gráfico" foi suficiente para verificar o nível do raciocínio sobre variação, pois os professores estavam ainda com dificuldades relacionadas ao procedimento.

Este fato ratifica a observação de Gal (2002), sobre a necessidade de conhecer o algoritmo da média para compreendê-la. Estendo sua afirmação, pode-se dizer que conhecer como se elabora um histograma e, principalmente, ter a apreensão que é uma representação de variável contínua, é essencial para que se possa analisá-lo e relacioná-lo com outros conceitos estatísticos, como a média e o desvio padrão.

Outro fato que merece destaque é o possível obstáculo didático criado a partir da analogia de distribuição de freqüências com o conceito de função, sugerindo ao aprendiz a existência de duas variáveis, quando na verdade só existe uma.

## 13 Raciocínio sobre variação/variabilidade na compreensão do conceito de média aritmética

No dia 10 de Junho deu-se início à discussão das medidas de tendência central. O objetivo desta atividade era verificar se os professores envolveriam raciocínio sobre variação ou variabilidade quando analisassem uma média ou outra medida de tendência central. A primeira tarefa apresentada está descrita a seguir.

> Eu tinha pedido para vocês fazerem a pesquisa sobre uma maneira de representar a idade e o tempo de magistério dos professores, sem a utilização gráfica, imaginando que nós escreveríamos uma pequena reportagem em jornal, em que só nos disponibilizariam uma ou duas linhas.
>
> Para isto, eu trouxe alguns livros didáticos de Estatística de maneira que possam complementar a pesquisa que vocês fizeram e ajudar a elaborar uma frase para divulgar a idade e o tempo de magistério dos professores pesquisados.

Figura 73 – Tarefa 1 da análise dos dados do ciclo investigativo – compreensão do conceito de média aritmética

Foram disponibilizados os livros: Barbetta (1998), Costa Neto(2002), Bussab e Morettin (2003) e Levin e Fox (2004) levados pela autora e o livro de Magalhães e Lima (2004), que os alunos AG e LJ estavam utilizando na graduação.

Os participantes dividiram-se em dois grupos, sendo um formado pelos professores e o outro pelos dois alunos.

### 13.1 Resultados e discussão da análise elaborada pelo grupo dos professores

Apenas a professora OB fez a pesquisa sobre uma maneira de representar a idade dos professores sem a utilização de gráficos. Os professores AM, LH e SB ouviram atentamente seu raciocínio. Ela disse:

> OB: <u>Eu achei que era só a média</u>, né? Eu achei a média, né? <u>Depois eu ia achar o desvio padrão</u>. Eu fiz no trem.
>
> SB: Quando falou em representação sem gráfico, eu pensei assim. No jornal vem <u>sempre com porcentagem</u>.

OB: É, daí, depois, como eu tenho aqui... Veja, a gente pode calcular o desvio padrão ainda. Porque como <u>está muito disperso assim</u>...

Pôde-se perceber que o desvio padrão não surgiu de maneira intuitiva. Esta medida apareceu na discussão devido ao fato de a professora OB ter realizado a pesquisa solicitada e perceber que havia um capítulo nos livros denominado Medidas de Dispersão.

A professora OB apresenta sua primeira idéia do desvio padrão, que seria usado para medir a dispersão, sem se referir à média, o que pode ser considerado com um raciocínio verbal, pois o fato do desvio padrão ser alto não implica, necessariamente, em grandes diferenças entre os valores, mas a possibilidade da existência de um único valor discrepante.

Além disso, sua frase foi “está muito disperso”, que segundo Makar e Confrey (2005), refere-se à variação com um predicado e que este conceito só pode ser apreendido à medida que é interpretado como um substantivo, tal como “a variação é grande”.

SB: Eu pensei uma tabela. Não sei.

OB: Dá média de 36,2 e o desvio padrão... Como é que você ia fazer?

SB: Ah, eu ia fazer o levantamento dos dados e ia calcular a média. Vai, dentro do, como o jornal faz, através de uma frase.

Essa tentativa da professora OB em compreender o que a professora SB faria pode ser um indicativo de que ela não tinha certeza sobre a necessidade de utilizar o desvio padrão.

Devido ao fato que a professora OB estava muito interessada em calcular e verificar se a média da idade estava correta, o grupo envolveu-se nesta atividade, esquecendo a análise.

A professora OB ensinou à SB como calcular a média com dados agrupados, que fora a estratégia utilizada por OB a partir da Tabela 31.

Devido ao fato de que o cálculo da média aritmética a partir de dados agrupados permite uma aproximação de seu valor, a PSQ fez a seguinte intervenção:

PSQ: Vocês fizeram a média da idade?

AM: Idade média 39 anos. Pegamos por faixa e vimos a freqüência.

PSQ: Antes de vocês pensarem no tempo de magistério, vocês fizeram a média usando as faixas, não é?

OB: Sim.

> PSQ: Quando vocês fazem as faixas, quando vocês estão pensando na média, entre 19 e 29, vocês pegaram o ponto médio?
>
> AM e SB: Isto.
>
> PSQ: Será que todas as vinte e uma pessoas tinham esta idade? [referindo-se ao ponto médio da primeira classe]
>
> SB: Não. Poderíamos ter todas com 19 ou 20 anos.
>
> PSQ: Então, uma primeira pergunta para vocês pensarem e discutirem. E se.... É que a gente tem 110, e na sala de aula não vai trabalhar com tudo isto. A gente vai calcular a média usando as faixas, que facilita os cálculos e em contra partida...perde informação. O que vocês acham de calcular a média com todos os dados, sem agrupar?
>
> SB: Vamos tentar calcular rapidinho para ver qual a diferença que dá.
>
> PSQ: Qual a diferença que dá para a gente poder discutir.
>
> SB: Deixa eu pegar minha tabela aqui.

A professora OB calculou a média dos 108 professores e obteve 38,48 anos de idade e a professora SB calculou de 107 professores, pois excluiu o professor que tinha 19 anos de idade e 15 anos de magistério e obteve 38,66 anos de idade.

A professora OB não se surpreendeu que a média que calculara era diferente da professora SB, mas não entendeu porque era diferente da média que calculara com dados agrupados. A PSQ preferiu deixar esta discussão para o momento do *feedback* e explorar a análise dos resultados, a partir da tarefa descrita na Figura 74.

| |
|---|
| PSQ: Então, eu vou propor outra coisa para vocês: 38,6, 38,4, 38,9 são muito próximos. Como vamos analisar esse número? Para calcular, de uma maneira ou de outra, não tem muita novidade. Mas como é que vamos olhar esse número? O que significa esse número? Quando eu vou olhar 38,9, se eu pedir para vocês analisarem esse número, o que passa pela cabeça de vocês? Como vocês analisariam? Imaginem que eu sou o ET! |

Figura 74 – Tarefa 2 da análise dos dados do ciclo investigativo – compreensão do conceito de média aritmética

> OB: Mas eu não diria 38,9. Eu diria aproximadamente 40 anos. Aproximadamente 39 anos.
>
> AM: O que veio na minha cabeça agora. <u>Um monte de gente</u>...
>
> OB: Que <u>a maioria tem</u>....
>
> SB: Que eu chegaria na escola e os <u>professores vão ter essa característica.</u>

> OB: <u>A maioria. A maioria. Lógico que nem todos vão ter 39, pode ter 38, pode ter até um de 19 anos, mas a maioria tem 39 anos, a característica é 39 anos. Igual a característica do aluno, jeans, tênis, óculos, camiseta, mas entre eles pode ter um que gosta de social, outro que gosta daquela saia de rock, aquelas coisas todas, né.</u>

A compreensão de que nem todos os valores são iguais à média foi considerado um raciocínio de transição, pois apresenta um entendimento correto de variação em torno da média. Em contrapartida, o fato de compreender a média como maioria, como moda, impede o surgimento da necessidade do desvio padrão e reflete um raciocínio idiossincrático.

Outro raciocínio idiossincrático é a compreensão que todas as observações da amostra são iguais à média, como apresentado pela professora SB.

> PSQ: E como é que vamos dizer para o ET que ele não deve acreditar nisto fielmente?
> SB: Por que só com este número fica muito .... vago, né? Não dá idéia do que realmente é. Só com um número você não consegue falar para ele que isso é só uma média.
> PSQ: E ai, o que você vai fazer para ele ….
> AM: <u>Daí você vai mostrar o gráfico</u>.
> OB: Mas o gráfico não pode. Ela falou que é só uma frase.
> PSQ: Suponha que você não tenha espaço para colocar o gráfico. Suponha que é apenas uma nota de jornal, que nem venda de carro, que você pode fazer até tantas letras e não tem espaço para colocar o gráfico tudo bonitinho. Como é que a gente poderia dizer para o ET, a gente já sabe que poderia pensar que todo mundo tem 38, 39....
> LH: <u>Mas se olhar o gráfico</u> ai você já vê as faixas
> OB: Mas não pode. É só uma frase!

Os professores perceberam que a média necessitava de um complemento, que poderia ser o gráfico, demonstrando um raciocínio de transição de variação. Porém, deve-se ressaltar que o uso do gráfico deixaria para o leitor a análise da variação, excluindo-se da responsabilidade.

> PSQ: Como diria para o ET que .. todo mundo não tem 39 anos?
> LH: <u>Você pega uma faixa menos, uma faixa mais. Vê o numero que tem de pessoas, a idade que começa, e sei lá.</u> Calcula três faixas, e calcula.

Quando o professor LH se refere às três faixas, poderia ser pensado que ele estava se referindo em faixas a partir da média. Porém, no transcorrer do discurso, foi possível inferir que ele estava se referindo às faixas do histograma, mesmo porque ele estava muito entusiasmado com a pesquisa que havia feito sobre este assunto. Outro fato ratifica esta inferência. A professora RS não pode comparecer ao encontro e deixou com ele a frase que elaborara: “Após estudos estatísticos realizados com base em pesquisa feita por um grupo de professores de Matemática da escola pública de ensino fundamental e médio que participam de um projeto sobre ensino aprendizagem na PUC-SP – Pontifícia Universidade Católica – chegou-se a conclusão de que a idade dos professores varia de 19 a 60 anos, onde a maioria estava na faixa de 34 a 44 anos de idade”, que referiam-se às faixas da distribuição de freqüência com dados agrupados ( de 34 a 39 e 39 a 44 anos de idade).

> PSQ: Vamos pensar assim. A gente tem uma pessoa que vem para cá para distribuir uma verba para ajudar famílias carentes, ou alguma coisa assim. A primeira coisa que ele quer saber é quantos filhos as famílias tem. E alguém diz para ele assim: Em média as famílias tem 2,3 filhos. Será que isto para ele é suficiente?
>
> OB: Não
>
> PSQ:Como é que ele vai olhar esse número? Assim como nós vamos olhar a idade média dos professores, como é que a gente vai interpretar isto? Como é que a gente vai ajudar essa pessoa a interpretar? A gente sabe da nossa realidade. O nosso vizinho não tem 2 filhos. Nosso vizinho tem 5 filhos. Então como é que a gente vai mostrar para ele que nem todo mundo tem 2 ou 3 filhos? Eu estou jogando a pergunta para vocês discutirem agora.
>
> OB: Então só se colocar assim: entre os professores pesquisados, eles tinham de 19 a 60 anos

A professora ampliou seu raciocínio de variação, podendo ser classificado como de transição, pois estava utilizando a média e os valores máximo e mínimo da distribuição. Vale ressaltar que a estratégia de utilizar esses valores já tinha sido empregada na tarefa dos sobrenomes e também nos estudos de Reading (2004), Meletiou (2000) e do próprio Ben-Zvi (2004).

> PSQ: Isso, já começou a dar uma outra idéia
>
> OB: Sim, ele já sabe que tem.... e se a média é 39 anos, a maioria tem 39 anos, mas nem por isso todos tem 39. Daí agora, como a média é 39, já vai logo saber que de 19 anos tem pouquinho, 60 anos tem pouquinho, 40, 43, 35, já dá para saber mais ou menos.

Novamente a professora OB apresenta um raciocínio idiossincrático de variação, pois interpreta a média como moda e infere a densidade de freqüência da distribuição, o que não é possível apenas com a média e os valores máximo e mínimo. O diálogo seguinte apresenta a intervenção da PSQ e da observadora para a explicação sobre o raciocínio incorreto.

> OBS RL: Pode Ter um monte de 20 e um monte de 60. Soma e a média vai dar 40. E?
>
> PSQ: Isso! Eu ia fazer com nota de aluno. Suponha que tenha um aluno, vou pegar aqui: 0, 0, 0, 10, 10, 10. Trinta dividido por seis é cinco. Ele tem nota média cinco. Vamos supor outro aluno que tenha tido notas: 5, 5, 5, 5, 5, 5. Ele terá média cinco.
>
> LH: É! Mas, aí então.......
>
> PSQ: Eles não são iguais né? Como é que nós vamos dizer para o professor que vai entrar nessa sala e vai encarar esses dois alunos aqui. Como é que a gente vai preparar esse professor e dizer assim, olha, todos esses alunos tem nota média cinco, isso não quer dizer que todos são iguais. Então, como é que nós vamos preparar.... Vamos supor que metade dessa amostra tivesse 20 e metade tivesse 60? E a média seria em torno de 40. Como é que a gente vai dizer para o leitor, vamos dizer para o cara que vem distribuir a verba, vamos dizer no relatório da escola como é a idade dos professores, assim como é o número médio de filhos.

Além dessa explicação, foi feita a média do número de filhos das pessoas que participavam do encontro e percebeu-se que a moda era zero filho e a média era dois filhos.

> LH: Então, e agora?
>
> OB: A moda é a maioria, né?
>
> PSQ: Isto. Se a moda é a maioria, como é que a gente vai falar que a maioria tem 39 anos [que era a média] ? É para você pensarem, esse é o objetivo!
>
> LH: Então a gente tem que pegar por faixa e....
>
> OB: Achar o desvio?
>
> PSQ: Será que quando você disse para mim que a idade mínima é 19 e a máxima é 60, você já começa a contar alguma coisa? A gente viu na semana passada que isso é a amplitude total, valor máximo menos valor mínimo. Já é uma medida de desvio. Mas que outro desvio você tinha pensado?
>
> OB: A moda também poderia entrar nesse caso. Nesse caso ai, do primeiro menino moda zero e moda dez e no segundo aluno, moda 5 que é igual a média. E ai o professor dá para saber direitinho que tipo de aluno ele tem ai.

> PSQ: Vamos supor um outro aluno. Um aluno que tenha 1, 4, 5, 10, 10 e que a média é 6 e a moda é 10. E?
>
> LH: Então vamos fazer por faixa.
>
> LH: <u>Então temos encontrar outra coisa aí que explique</u>.
>
> PSQ: Boa pergunta! Meu objetivo é... minha pergunta é exatamente esta!
>
> OB: <u>Trabalhar com dispersão. Vamos achar o desvio padrão e</u> ...

É possível notar que o professor LH permaneceu com a idéia das faixas e a professora OB pretendia complementar a média com a moda, que indica um raciocínio de transição de variação. A sugestão do desvio padrão, possivelmente, surgiu devido à leitura dos livros didáticos e não pela compreensão de sua necessidade.

> PSQ: Isto! Minha pergunta é.., a média é... A gente viu que a média pode ser cinco quando ele tem nota zero e dez, pode ser cinco quando só tem nota cinco e.... Então, a minha pergunta é: o que eu posso fazer para ajudar a pessoa a interpretar esse número? Pois, a gente já viu que não é a maioria, porque a maioria é a moda. E aqui a moda é dez e a media é cinco e aí a moda nem sempre vai ajudar a interpretar. A maioria é a maior quantidade de casos, que é a moda. Então a gente precisa achar uma outra medida
>
> AM: Agora você deu um nó!
>
> OB: <u>É para trabalhar com a dispersão. Desvio médio e desvio padrão. É isto que eu quero discutir!</u>

Essa verbalização da professora OB permite confirmar a inferência anterior. Ela tinha interesse em compreender o significado das medidas de dispersão.

> PSQ: Se o objetivo é trabalhar com dispersão, será que a gente não pode considerar uma medida de dispersão a primeira coisa que você disse [referindo-se à professora OB] idade mínima 19, já dá uma idéia para a pessoa. Esse aluno tá indo de 5 a 5, dispersou alguma coisa?
>
> OB: Não!
>
> PSQ: Esse aluno tá indo de 0 a 10!
>
> OB: Nossa senhora! Tenho que pensar com carinho nestas coisas!
>
> PSQ: Esse cara tá indo de 0 a 10 também! Esse aluno tem hora que entendeu tudo que eu disse e tem hora que ele não entendeu nada.
>
> OB: <u>Ou ele falta demais</u>!

A professora OB já estava se preocupando com as causas da variação da nota do aluno, o que já havia sido alertado por Meletiou (2000), quando voltou para pensar na frase que elaboraria acerca da idade dos professores pesquisados.

> OB: Então, a gente fala a quantidade de professores, que é 110, de 19 a 60 anos...com idade média de 38,6 e ...Vamos achar a dispersão! Vamos ver como é que se faz. Variância..... Medidas de dispersão.....[procurando nos livros didáticos].
>
> PSQ: Então vamos lá.

A evolução das discussões permitiu a alguns professores identificar a necessidade de um complemento para a média aritmética, que foram os valores máximo e mínimo, a moda e a própria representação gráfica. Isto corrobora os estudos de Lehrer e Schauble (2002) e Ben-Zvi (2004), no que diz respeito ao não surgimento da necessidade de uma medida de variação e de Makar e Confrey (2005) no que diz respeito à menor incidência do uso do termo desvio padrão, quando comparado com outros termos estatísticos.

Embora o desvio padrão tenha sido citado mais de uma vez pela professora OB, isto não significa que ela utilizaria o desvio padrão para complementar a média, mas sim que sabia que precisava usar o desvio padrão para esta finalidade e não sabia como calcular e nem interpretar.

A partir deste momento, os professores motivaram-se para a leitura dos livros didáticos, buscando compreender o que eram as medidas de variação.

> OB: Aí, [referindo-se ao Professor AM], <u>eu só queria entender o que é desvio médio, variância e desvio padrão</u>. A altura média é 49, e depois vem o desvio padrão 1,8, a variância é 4,3 e o desvio padrão 2. O que significa? [estavam observando um exemplo no livro didático de ensino médio]
>
> AM: Até um pouco de tempo atrás eu achava que sabia o que significava. Mas agora.... Eu já trabalhei com este assunto nos últimos dois anos e ai a Claudia falando aqui, comecei a pensar quanta coisa eu nunca falei e nem sabia que era para falar . Eu posso até chegar para meu aluno e perguntar o que é moda, o que é desvio, só que, olhar para isto e saber quando usar um ou outro...Nunca discuti isto.
>
> OBS VG: Não é só saber o que é, mas para que serve!
>
> AM: Em que momento aplicar. Eu nunca parei para pensar nisto.
>
> OB: Aqui ó [olhando o livro]. <u>É só para saber se estão muito dispersos ou não!</u>
>
> SB: Você está falando do desvio?

> OB: De todos, né! [referindo-se ao desvio médio, variância e desvio padrão] E o que a gente faz com isto?

A professora OB permaneceu com um raciocínio verbal de variação.

> OB: Desvio médio….
>
> SB: Então esse x com tracinho em cima é a média, certo? Desvio médio será que é a média do desvio?
>
> OB: E a variância, quer ver. A dispersão dos dados também pode ser calculada considerando-se os quadrados do desvio médio. A média aritmética desses quadrados chamamos de variância. Considerando como exemplo a altura teremos o seguinte calculo da variância.
>
> AM: Desvio padrão é a raiz quadrada da
>
> OB: Variância.
>
> AM: E a variância, por sua vez se utilizou do
>
> OB: Do desvio médio. E o desvio médio se utilizou da altura. Então, o mais .., o mais..., o que mais dá a idéia é o desvio padrão.

Os professores estavam buscando nos livros as respostas para as indagações feitas pela PSQ .

> AM: Daí, você pode olhar em algum lugar que, quanto mais isto se aproximar de zero, mais perto de.....Se os professores [referindo-se à pesquisa realizada], a gente faz os cálculos e se estiverem próximos de zero, significa que a maioria está dentro desta faixa aqui. Se não se aproxima de zero, vai mostrar que está muito disperso. Então vamos fazer?
>
> OB: Olha Claudia! Eu falei para ele, então a gente precisa ver qual é a dispersão, a medida de dispersão. Aí eu leio no livro, desvio médio, variância e desvio padrão. Aí eu falei para ele que o desvio padrão é a que dá a idéia mais correta da dispersão. Mas, e agora? Se a média é, no caso, aqui acho que é 49, eu tenho um desvio médio 1,8, variância 4,3 e desvio padrão 2. Eu não estou entendendo nada. Ai ele falou....
>
> AM: Eu disse o seguinte. Esse número aqui, o dois, quanto mais próximo ele estiver do zero significa que aqueles dados dos professores está mais homogêneo, se estiver mais longe do zero e eu não sei quanto longe do zero. Ai tem que fazer uma outra análise. Então muita gente com vinte, muita gente com trinta, e assim por diante.

O professor AM retomou sua interpretação de homogeneidade, apresentada em sua apostila e reforçada pelos livros didáticos de Matemática, já classificada como um raciocínio verbal de variação, e que ele mesmo pôde perceber a fragilidade dessa análise. Além de não fazer menção à média, é subjetivo dizer maior e menor.

O diálogo seguinte ilustra o surgimento da variação em torno da média.

> LH: Recebe o nome de dispersão tendo como referência o quê? A média? <u>Se afasta da média?</u>
>
> PSQ: Isto! Exatamente! Então, a primeira coisa é: esta dispersão que a gente está chamando de desvio padrão, variância ou desvio médio, todos eles estão vendo quanto fica distante da...
>
> OB: da média.
>
> PSQ: Então, quanto mais distante do zero, mais distante os dados estão da...
>
> LH: Da Média. <u>O zero seria a média</u>?

Esta dúvida do professor LH permite retomar a frase do professor AM na fase de sensibilização da pesquisa, em que confunde o desvio padrão próximo do zero com a média próxima do zero, classificado com um raciocínio idiossincrático de variação, que pode surgir se o conceito não estiver bem formado.

> PSQ: O que seria um desvio padrão zero?
>
> OB: Que não existe, quer dizer, que não existe uma
>
> LH: Está bem concentrado!
>
> PSQ: Dá um exemplo de desvio padrão zero.
>
> SB: É o exemplo do aluno 5, 5, 5....
>
> PSQ: Isto. Não tem nenhuma variação em torno da média, pois se a media foi 5 e todas as notas foram 5, quanto as notas estão distante da média? Nada!

Esse raciocínio de transição permitiu à professora SB compreender quando o desvio padrão é zero, os valores são iguais. Porém, ainda assim não foi considerado igual à média, que foi reforçado pela PSQ.

A discussão sobre a interpretação do desvio padrão continuou, pois o objetivo da PSQ era despertar a percepção para possíveis intervalos em torno da média.

> LH: Vamos pegar aquele exemplo que a OBS RL falou. Nos extremos,
>
> PSQ: Vamos supor que metade das pessoas tivesse 20 anos e metade tivesse 60 anos, esse desvio padrão seria perto ou longe de zero?
>
> SB: Não ia ter nenhuma pessoa com 40 anos. A gente ia ter uma média de 40 anos e nenhuma com 40.
>
> PSQ: O que acontece com cada idade? O que acontece quando você tem mais dados longe da média?
>
> OB: Mais dispersão!
>
> PSQ: Então o que acontece com o desvio padrão?
>
> OB: Vai ser alto!

Para o professor LH ainda não estava clara a relação entre o valor do desvio padrão e as distâncias da média, enquanto as professoras OB e SB estavam buscando fazer essas conexões entre os conceitos.

> PSQ: Então, vamos voltar ao nosso exercício. A gente não vai dizer para um ET que a idade média é 38,8, 38,6. Se ele tiver um desvio padrão de 1 ano, então....
>
> OB: Então eles tem 37, 40.
>
> PSQ: Se eu falar que estes professores têm uma idade média de 38 anos com uma variação de 30, um desvio padrão de 30. Então, tem a gente .....
>
> OB: Bem novinha ou muito velha.
>
> PSQ: Então, a gente está começando a mexer com esta idéia.

A partir da intervenção da PSQ, surgem os primeiros indícios de um intervalo em torno da média, que foi considerado um raciocínio de procedimento, pois apreende a média, os desvios da média, um intervalo a partir desses desvios, mas não permite compreender que existe uma grande maioria nesse intervalo.

> OB: Que ótimo, viu! Eu não sabia o que era isto. Eu peguei, eu ia fazer isto lá em casa, eu ia achar a média, tudo, né. Mas ai eu vim pensando no trem: que droga é esta de desvio padrão, o que é este desvio médio, o que é a variância? Aí eu pensei: nem vou fazer mesmo, pois eu não sei mesmo.

O desabafo da professora OB permite confirmar as inferências anteriores, que ela estava em busca da compreensão do desvio padrão e, por esse motivo, havia verbalizado-o. Isto confirma o que Makar e Confrey (2005) alertaram, que o fato do professor usar o termo desvio padrão não significa que ele esteja compreendendo o que está falando.

Os professores deram início ao cálculo do desvio padrão com dados agrupados e comprometeram-se a calcular com todas as observações em casa. O professor AM verbalizou sua dificuldade para compreender o significado do desvio padrão alto e a PSQ explicou que voltaria a esta discussão.

### 13.2 Resultados e discussão da análise elaborada pelo grupo dos alunos

Eles tinham aprendido, recentemente, as medidas de tendência central na graduação e, inclusive, teriam prova deste assunto. O aluno LJ buscou em seu caderno o "modelo" de cálculo da média com dados agrupados. Eles decidiram pela média, mas sem ter conhecimento do motivo. O diálogo a seguir ilustra este fato.

> LJ: Bom, a moda não vai ser útil, ou vai? A gente vai ter o maior número de dados, e daí? A mediana, meu professor não falou.
>
> AG: Nossa que absurdo!
>
> LJ: A mediana não serve. Então vamos calcular a média. X barra.

Os dois alunos verbalizaram que não tinham aprendido as medidas de variação e recorreram ao livro de Magalhães e Lima (2004) para ler sobre o assunto. Quando a autora questionou sobre a análise da média, o diálogo seguinte ilustra o raciocínio dos alunos.

> LJ: O professor tinha falado, na verdade, uma vez ele até propôs na sala. Eram duas faixas de empregos. Ele dava a média. Eram dois estágios e você tem que escolher. Ai você vê né? Se a média tá muito... se tem um pico muito alto ou muito baixo, lógico que a média dá uma aliviada e você não consegue sentir que tem muita gente embaixo. Aí você teria que ir pela mediana ou pelo cálculo do desvio, da variância. O outro é quando é mais ... eu acho...que a média infelizmente é um dado muito... se a gente tivesse a moda, por exemplo, que aparece mais gente, ai pode te ajudar. Que nem, vai pensar o que? A média dos professores é 38, mas as pessoas são mais idosas. <u>Mas, às vezes muitos são jovens e um cara muito velho ai a média sobe.</u>
>
> PSQ: Isto. Então, vamos analisar esta média.
>
> LJ: Pode olhar os outros dados para ajudar a interpretar?
>
> PSQ: Claro!
>
> LJ: Moda, mediana.
>
> PSQ: Isso. É para vocês mexerem mesmo com os dados.

Este diálogo ilustra o raciocínio de transição do aluno LJ, quando utiliza a média e a moda para analisar uma distribuição. A PSQ questionou novamente como fariam a análise da média para explicar que nem todos os professores tinham 39 anos.

> LJ: Calcular a variância?
>
> PSQ: É uma das alternativas. E como a gente vai analisar essa variância?
>
> LJ: <u>Quanto maior a variância, mais dispersos estão os dados</u>.

> C: Isto. Esse é o objetivo. Porque aí você dá a dimensão para a pessoa. Você fala: essa aqui é a média, mas...
>
> LJ: Então, quanto maior for a variância. A gente até fez o exemplo. A pessoa disse 2, 2, 2,… A variância é zero e é tudo homogêneo.
>
> PSQ: Isso. Então vamos lá.

O aluno LJ apresenta dois raciocínios diferentes, ambos verbais. Também apresenta a variação como um predicado e não como um substantivo, tal como os professores já o tinham feito.

### 13.3 *Feedback*

As Tarefas 1 e 2 consumiram todo o encontro do dia 10 de Junho e, antes do término da sessão, foi salientado a não necessidade de trabalhar com uma amostra grande para estimular a compreensão da média aritmética e do desvio padrão, mas que após a discussão inicial deste conteúdo, o professor poderia explorar a utilização de planilha eletrônica e realizar pesquisas com os alunos.

O *feedback* desta atividade foi realizado no encontro seguinte (17 de Junho) e tinha como objetivo formalizar os conceitos de média aritmética simples, moda e mediana, priorizando a interpretação destas medidas e a conseqüente necessidade das medidas de variação. A tarefa solicitada está apresentada na figura a seguir.

| PSQ: O que significa média aritmética? |
|---|

Figura 75 – Tarefa 3 da análise dos dados do ciclo investigativo – compreensão do conceito de média aritmética

A interpretação predominante de média aritmética foi o ponto de equilíbrio do conjunto de observações, mas também houve a interpretação da professora RN como referência de um conjunto e a interpretação do aluno LJ como divisão eqüitativa. Não se pretende aqui explorar o tipo de raciocínio sobre a compreensão da média, mas sim como eles levavam em consideração a variabilidade/variação nesta interpretação.

> LF: Quando ela fala em ponto de equilíbrio, eu acho que o aluno tem essa coisa de equilíbrio mais fácil na mente dele: que é o centro, né? O centro de quê? O ponto de equilíbrio de vários dados. Mas, que este centro, ele pode variar, tá! Não quer dizer que ele tá demonstrando que

> tudo está aqui no centro. Existe uma variação anterior e posterior dos dados.
>
> PSQ: Quem varia é o centro ou são os dados?
>
> LF: São os dados. Quem varia são os dados.
>
> PSQ: A média é a média e ponto.
>
> LF: Lembra que a gente tinha discutido as quatro notas de três alunos? [Aluno A: 0, 0, 10, 10; Aluno B: 5, 5, 5, 5 e Aluno C: 1, 6, 6, 7].
>
> PSQ: Nos três casos, a média é cinco. Foi o que a gente colocou no caderno da LF.
>
> LF: Média não pode ser o meio, porque ali [referindo-se ao Aluno C] não está no meio.

Embora a professora LF estivesse levando em consideração a variação das observações em torno da média em sua interpretação, que permitiria classificar seu raciocínio como de transição, havia uma confusão se esta variação era dos dados em torno da média ou a variação da própria média, que poderia ser classificado como raciocínio idiossincrático, pois, pelo contexto da atividade, infere-se que a professora estava se confundindo e não que ela estivesse se referindo à  distribuição amostral da média.

Este diálogo suscitou a formalização dos conceitos de media aritmética simples, moda e mediana a partir do exemplo das notas dos três alunos, que foi realizado. Diferentemente das representações gráficas, os professores sabiam calcular as medidas, mas desconheciam sua interpretação.

Houve a discussão sobre a maneira de explicar ao aluno o que seria equilíbrio. Surgiram sugestões como centro de massa, utilização de uma barra com pedras e a utilização de uma régua de trinta centímetros com o dedo centrado no quinze. A observadora RL fez quatro bolas de papel e as colou numa régua de trinta centímetros e, o ponto em que o seu dedo equilibrava a régua identificava a média aritmética.

Os professores gostaram muito da experiência concreta elaborada pela observadora e foi discutido sobre a importância de distribuir as bolas de papel na régua de maneira que não haja a mesma quantidade de bolas de papel acima e abaixo da média aritmética (semelhante ao exemplo do Aluno C).

Para complementar esta discussão e estimular a percepção da variação, a PSQ elaborou um gráfico de pontos para o aluno A, um para o aluno B e um para C, todos numa mesma escala.

> LF: Então, ela é o ponto de equilíbrio, mas <u>não retrata</u> a realidade. A realidade..... não! Vamos falar, é.... a variação das distâncias, não! <u>Variação da distribuição</u>.
>
> CI: Eu ensino velocidade média. Então, saindo de um ponto x, eu calculo a variação no intervalo. Então, eu posso dar.... nem sempre.... falando de velocidade.... eu posso andar a cem quilômetros por hora, essa é a média, mas nem sempre.... pode haver variação.
>
> LF: Eu acho que ela [a média] não demonstra a variação das medidas, né?

A professora LF apresenta um raciocínio verbal de variação, pois compreende que a média não retrata a variação da distribuição, mas não sugere como poderia ser complementada. Além disso, esse diálogo ilustra o momento em que inicia a construção do conceito de variação.

Pela fala da professora CI é possível observar que não era uma tarefa simples para os professores explicar que as observações variavam em torno da média.

Devido à intervenção da professora RN sobre a necessidade de entender se sua interpretação de média como referencial estava correta, esta discussão foi interrompida, mas em breve retomada.

> PSQ: Como a gente explica para o aluno que a média é um referencial?
>
> LF: Uma pessoa boa e uma pessoa má. Um referencial de bom e mau.
>
> RN: A partir daí eu vou.....
>
> PSQ: Quando você fala: ele passou e a nota média dele é 5. Ou, ele passou e a nota média dele é 7. Ou ainda, ele passou e a nota média dele é 8. Aquele 5, aquele 7, aquele 8....
>
> RN: É uma referência do aproveitamento do aluno.
>
> PSQ: Daí, eu, professor novo, entro lá para dar aula para esse aluno, esse aluno e esse aluno [referindo-se ao exemplo dos alunos A, B e C] e eu imagino que, como ele é nota 5, que é a referência dele, eu estou imaginando que parte do que eu vou ensinar para ele, ele não vai entender. Parte ele vai entender. Agora, eu não estou olhando as notas dele. A hora que eu verificar as notas dele, aí eu vou falar: ele não vai entender muita coisa mesmo [referindo-se ao aluno B] ou tem coisa que ele vai entender tudo e outra nada [referindo-se ao aluno A]. Por isso que, se eu olhar só a média, não me dá essa idéia. Eu só olho e vou esperar que ele é assim. Daí, eu dou a primeira prova, espero que ele é cinco e ele tira oito.

> LH: Mas, acontece o seguinte. Ele pode ter decorado tudo para a prova, tira nota alta e depois esquece tudo.

Este diálogo ilustra a retomada da discussão da variação em torno da média e, foi possível observar que o professor LH estava buscando explicações para a variação, tal como já tinha acontecido com a professora OB. Vale destacar também, ratificando o que Meletiou (2000) alertou, que o fato de utilizar dados reais em exemplos, motiva os participantes a buscar as causas de variação, trazendo a tona suas crenças e valores. Este ponto não foi explorado nesta discussão.

O diálogo prossegue:

> LF: Então, o que faz a diferença aí? É a variação que se tem entre um número e outro? É o desvio padrão!

Novamente uma interpretação do desvio padrão com a variação entre uma observação e outra, raciocínio verbal, o que Loosen, Lioen e Lacante (1985) denominaram de *unalikeability.*

> PSQ: Então, o que a gente tá vendo é que a hora que a gente olha só a média, você tem um olhar mais ou menos, por quê? Um olhar que pode ter uma interpretação equivocada porque existe variação.
>
> LH: Então, o cara que não tem todos os dados, o que ele tem que ter? Desvio, mediana, estas coisas todas?
>
> PSQ: A gente ta vendo aqui que, se olharmos só a média, vamos pensar que ele [referindo-se aos alunos A e C] é sempre cinco, isto se eu não levar em consideração a variação. Porque a hora que eu olho esta variação, eu já sei que, se este aluno tiver uma nota bem abaixo de 5, eu não devo me surpreender e se tiver uma nota bem acima de cinco, também não. Porque a variação das notas dele é alta.
>
> LF: Mas, quando você pega a moda, você já consegue ter uma noção um pouco diferente da sua média.

Novamente o surgimento da moda como uma substituta de uma medida de variação

> PSQ: Verdade. Então, o que a gente ta vendo? É que só a média, sozinha, você não vai conseguir tirar grandes conclusões. E aí, sai todo mundo tirando conclusão em cima da média. E aí, a gente escuta coisa assim: a Estatística é mentirosa! Eu não acredito na Estatística! Na verdade, o objetivo que a gente tem é mostrar que a Estatística não é só a média e que a média sozinha não dá conta do recado. Então, a gente tem que mostrar para o nosso aluno, alfabetizar o nosso aluno em Estatística e dizer para ele assim: Olha, quando você olhar para a média,

> chame o desvio padrão. Por quê? Porque a média sem o desvio padrão, eu não sei se estes dados variam muito ou variam pouco [...] Se o desvio padrão for pequeno, vou esperar sempre em torno de cinco [referindo-se ao exemplo dos três alunos]. Se o desvio padrão for grande,...A média de idade dos professores foi 38,9 [voltando à idade dos professores pesquisados]. Eu vou esperar que todos estes professores tenham em torno de 39 anos? Não! Porque o desvio padrão foi nove [calculado pelo professor AM]. Então, é um desvio padrão alto. Eu tenho, pelo menos, nove anos para baixo e nove anos para cima.

Para encerrar este *feedback*, a PSQ sintetizou o artigo de Rodrigues (1999), que criticava a reportagem do jornalista que dizia: "Estatisticamente, eu já morri há cinco anos" e solicitou duas tarefas para serem realizadas em casa: a) elaborar a frase para o ET sobre a idade dos professores pesquisados, usando a média e o desvio padrão e b) relacionar a média e o desvio padrão com o histograma (já realizado) da idade dos professores.

## 13.4 Análise da média aritmética de observações apresentadas em gráfico de barras.

No penúltimo dia da formação continuada foi solicitado aos professores comparar as notas das alunas e dos alunos, que estavam apresentadas no gráfico reproduzido na Figura 76 e decidir sobre a distribuição com maior variação (Tarefa 2).

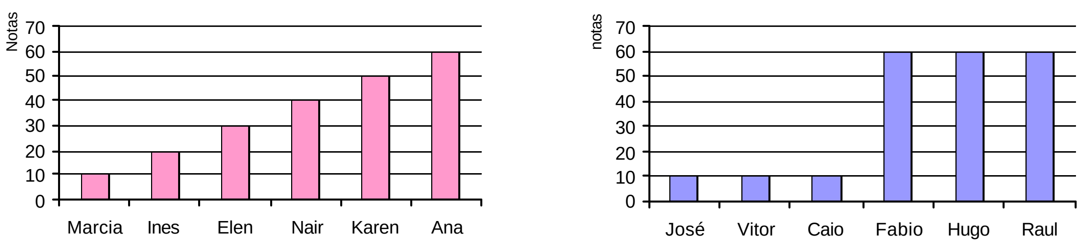



Figura 76 – Notas de alunos e alunas em uma determinada disciplina - Adaptado de Loosen, Lioen e Lacante (1985) -

## 13.5 Resultados e discussão da análise elaborada pelo Grupo 1 – LH e RS

Como já constatado por Loosen, Lioen e Lacante (1985), esta apresentação pode induzir o aprendiz a perceber a variação de uma observação para a outra, denominado pelos autores de *unalikeability*, e não a variação em torno da média. E foi exatamente o que aconteceu com os professores LH e RS, como pode ser observado no diálogo seguinte.

> PSQ: Quero saber que turma que vocês acham que teve maior variação de notas.
>
> LH: <u>Essa aqui não é igual a essa, não é igual a essa, não é igual a essa, não é igual a essa</u>, ....tá variando! Isso é igual a velocidade, tá acelerando, acelerando, acelerando! Variação da velocidade.
>
> RS: Olhando rapidamente.... [concordando, mas desconfiando]
>
> LH: Bom, <u>aqui a gente vai ter que fazer a média</u>, somar tudo. Você vai ver quantos tem.... A média vai ser igual.

Além da constatação da variação das observações, raciocínio verbal, o surgimento natural da média aritmética pode ter sido em função do exemplo utilizado, notas em uma disciplina, que faz parte do cotidiano do professor e ele calcula a média sempre. Essa hipótese é levantada, pois o mesmo fato não foi observado na tarefa do número de letras do sobrenome.

> RS: A média é 35. E aí?
>
> LH: Aqui a gente vai ter que trabalhar os desvios. Agora, <u>aqui tá mais concentrado e aqui também está concentrado aqui [referindo-se ao grupo de 10,10,10 e 60,60,60]</u>
>
> RS: E?
>
> LH: A mediana, quer dizer a moda, né ?, que é aqui e aqui também. Só que aqui, a moda está bem forte do que aqui.

Novamente observa-se a confusão que o professor LH fez com os conceitos de mediana e moda.

> RS: A média é 35, ok?
>
> LH: A moda é a que tá mais concentrada, mais alta.
>
> RS: Moda aqui é três. Aqui é três e aqui é três. A moda não é a que tem maior freqüência? Maior quantidade de vezes? Aqui três alunos tiraram seis, três alunos tiraram um.
>
> RS: O que ela quer saber mesmo? Ela quer saber a variação (silêncio) Cada um tirou uma nota e aqui....

A professora RS confunde o valor da moda com a freqüência. A moda era dez e sessenta, pois a freqüência de cada uma era três.

> LH: A moda está mais para o lado de cá ...... <u>Maior número de alunos é seis.</u>
>
> RS: Não! Cada barra representa um aluno. A Karen tirou cinco, a Ana tirou seis. Não é dez pessoas com a nota...
>
> LH: Então, aqui também. Quais os alunos que tiraram acima de 40?

É possível notar que este tipo de representação, além de induzir o aprendiz a observar a variação entre as observações, induz ao erro de pensar numa distribuição de freqüências. É possível observar que o professor LH olhou a barra mais alta para encontrar a moda, mesmo sendo alertado pela PSQ que isto não era uma distribuição de freqüência.

> LH: Será que, individualmente.... aqui nós estamos trabalhando o que? As notas ou o número de alunos?
>
> RS: É tudo igual. Quantos alunos tiraram 10, um. Quantos alunos tiraram 20, um. E assim por diante.
>
> LH: Aqui houve variação na nota e aqui houve variação no número de alunos. .
>
> RS: Cada um aluno tirou uma nota!
>
> LH: Então, se a média desse aqui...(silêncio) Help! O que você quer saber mesmo?
>
> PSQ: Quero saber qual é o que tem maior variação de notas.
>
> RS: Aqui, o que variou foi a nota e aqui o que variou foi a quantidade de alunos?
>
> PSQ: Não. Aqui a quantidade é a mesma. Tem seis mulheres e seis homens.
>
> RS: Maior número de alunos...Aqui, cada aluno tirou uma nota diferente. Aqui, três tiraram uma nota e aqui três tiraram outra. E aí, o que nós falamos mesmo?
>
> PSQ: Isso. Qual que é a maior. Qual dessa aqui é a maior?
>
> LH: Da nota?
>
> PSQ: É, da nota. Variação de notas. Vocês tem duas turmas: De menino e menina e quero saber qual variação é maior, variação de notas. Têm seis alunos e seis alunas. Aí, vocês calcularam a média e deu tudo igual. Então, eu quero saber o que tem mais variação: o grupo dos meninos ou o grupo das meninas?

Os professores permanecem olhando a distribuição, tentando entender, mas não utilizando nenhuma medida estatística.

> LH: Aqui varia de 10 em 10. Aqui varia de 10 pra 60. Aqui tem um delta...(silêncio) Então, aqui eu tenho que usar aquela história de desvio
>
> PSQ: Você pode usar a estratégia que você quiser. Poderia usar o desvio. Como vocês já calcularam a média, né?
>
> LH: A gente pode usar o desvio mesmo. Daí da para comparar. É! Vamos tentar. A gente fazendo o desvio, a gente vai ter uma visão melhor. Daí a gente vai ver quem desvio mais aqui e quem desvio mais aqui.

> PSQ: Desvio em relação?
>
> LH: Á média.

A decisão do professor LH em usar o desvio, possivelmente não se relacionava com o desvio padrão, mas a ansiedade da PSQ era tanta, que induziu o professor LH a dizer que era em relação à média, mas ele não se apropriou disto. Pois ele relata os desvios em cada distribuição, sem se referir à média.

> PSQ: Mas, antes de calcular o desvio, visualmente, qual grupo tem maior variação de notas?
>
> LH: Fica difícil saber. Só olhando assim?
>
> LH: Essa daqui está variando mais [meninas]
>
> PSQ: E você, RS?
>
> RS: Não sei. Tava tentando fazer aqui com esses dados diferentes, vê se conseguia ver alguma coisa... 10, um, 20, um, 30, um. Agora eu não sei como contar isto aqui.
>
> PSQ: Como contar, como assim? Você tem seis alunos, você tem três com 10 e três com 60.
>
> RS: Então, daqui pra cá eu tenho que contar esse intervalo aqui?
>
> PSQ: Você não tem ninguém ai!
>
> RS: A diferença como seria? A diferença seria de 40? A variação...
>
> LH: Ahhh, maior é o de cima? [meninas]
>
> RS: Fica uma variação de 10, 20, 30, 40, não é isso?
>
> LH: O aluno permanece igual, porém a nota de 10 pulou pra 60. Nesse intervalo só houve um patamar que passou de 10 pra 60. E aqui foi de 10 em 10 em 10 em 10...<u>Houve variação, mas de um por um</u>. O número de variação nesse aqui, pra mim, foi maior, agora...
>
> PSQ: Eu quero a variação de notas.
>
> LH: Então, você quer variação de notas, né? <u>De 10 pra 20, houve uma variação, de 20 pra 30 outra variação, de 30 pra 40... e assim até chegar no 60. Então, número de notas, aqui houve cinco variações e aqui houve uma variação.</u>

A variação foi compreendida pelos professores como o número de alterações das observações, raciocínio verbal. Tal como alertado por Loosen, Lioen e Lacante (1985), existe uma tendência natural de interpretar o desvio padrão como uma medida de variação entre as observações e não uma medida de variação em torno da média.

Portanto, eles não analisaram a magnitude das diferenças entre as observações, apenas o número de notas diferentes.

### 13.6 Resultados e discussão da análise elaborada pelo Grupo 2

O outro grupo de professores, composto por OB, CI, AM e SB, também, iniciaram a análise com a média aritmética. A professora SB, assim que olhou a Tarefa 2 já percebeu a variação em torno da média. Calculou-a e percebeu que nos dois grupos a média era 35. O diálogo que segue ilustra isto.

> SB: A média nos dois é a mesma. Só que os que estão mais distantes dessa média são esses aqui.
>
> AM: Os meninos?
>
> SB: É.
>
> AM: Normalmente é isso, mesmo. Se você avaliar elas são mais... é, menos displicentes em sala de aula, ou <u>eles são bons ou ruins</u>. Na verdade eles são... ou eles são largados ou dedicados, <u>as meninas não, são mais homogêneas</u>, eu acho...

Mais uma verbalização, agora do professor AM, em busca de causas para a variação.

> SB: Porque é o seguinte. A impressão que dá é que só a média de 35 é diferente para este caso do que para este.
>
> AM: A média 35 aqui, é isso mesmo que você tá falando, é mais fácil pensar nisso aqui [meninas] ... do que nisto aqui [meninos].
>
> SB: Mas, parece que este número não consegue demonstrar isso por causa da freqüência.
>
> AM: Você tá falando em média? Você tá mostrando um lado que está equilibrando, o que não é verdade, né. Não sei nem se é isso que a gente pode falar. <u>Quando você fala em média, você pensa todo mundo próximo de um ponto</u>, o que não é verdade aqui. Embora...

A diferença na variação das duas distribuições incomodou os professores AM e SB, questionando a validade da média. E o professor AM percebeu que sua interpretação da média como moda não serviria sempre, raciocínio idiossincrático, como ocorrido na atividade.

> AM: Vamos dizer assim, os bons alunos da sala, os que estão na sala tiram que nota, os alunos que tiram mais notas da sala qual é a média deles? 60, e os menos, 10. Aí se a gente equilibrar dá 35. Agora, aqui já é diferente, né?
>
> AM: Será que tem que fazer desvio padrão? Desvio padrão mostra bem isto para a gente. Você falou em média 35 né? Se a gente for analisar o desvio padrão, o que vai dizer para a gente?

> AM: Aqui, os meninos, você tem até 60, né? Três com 10 e três com sessenta. Nas meninas, você também tem 60 e 10, mas também tem 20, 30, 40, 50.

O professor AM percebeu que as notas dos alunos pulavam de 10 para 60, ficando muito longe da média 35. Sua intuição era que o desvio padrão fosse divulgar isto, mas estava em dúvida, o que ilustra seu raciocínio de transição. E o diálogo entre eles prossegue.

> SB: Na variação já vai sair uma média, só que a média vai ficar distante do...
>
> AM: Se eu tivesse que aplicar em sala de aula pegaria esse aluno trocaria com um...
>
> CI: Que você vai fazer?
>
> AM: Tem uma pergunta assim...na minha visão, tá claro que esse grupo não está trabalhando como esse grupo: Eu tiraria ele daqui e colocaria aqui e no lugar do Vitor e tornaria a sala mais homogênea, certo. O Vitor iria aprender um pouco mais e o José e o Caio também podiam aproveitar. E aqui, nesse caso, trocaria a Inês com a Karen.

Essa explicação do professor AM indica sua propensão de "alterar o processo" (na visão de SNEE, 1990) para que a variação diminua, o que tornaria "a sala mais homogênea", raciocínio verbal de variação. Essa explicação também pode demonstrar a preferência pela utilização de desvio padrão pequeno, pois ainda não estava claro para esse professor o que significava o desvio padrão grande.

> PSQ: Qual é a que teve mais variação?
>
> AM: Olha, qual tem maior variação, olha, variou, variou, variou, e variou. Que é maior variação pra você?
>
> CI: Tá, mas quando você olhou, você falou logo esse, né?
>
> SB: Aqui, a cada aluno varia a nota, né, tem mais variações, são seis variações, seis notas diferentes. E aqui são três notas iguais, três notas iguais, é maior variação em relação à média.

É possível observar que a professora SB percebeu que se fosse falar de variações entre as observações, as meninas variavam mais, mas se fosse falar de variação em torno da média, o grupo dos meninos variava mais. Ela enxergava a distância de 10 para 35 e a distância de 60 para 35 e comparava com as outras meninas, distância de 30 para 35. Ela pensava nas distâncias em relação à média. Seu raciocínio pode ser considerado completo, pois faz a distinção da linguagem e suas conotações.

Tal como enunciado por Makar e Confrey (2005), quando a professora SB percebeu as variações entre as observações, ela disse ”tem mais variações”, variação como predicado, e quando ela se referiu às variações em torno da média, que a levaria ao desvio padrão, ela disse “é maior a variação”, substantivo.

## 13.7 *Feedback*

O *feedback* desta atividade foi realizado no último dia da formação continuada (26 de Agosto).

> LH: O que varia mais? O que quer dizer essa palavra “varia mais”? Variou mais em que sentido? Teve uma variação maior? Um patamar maior? Foi nesse sentido?
>
> PSQ: Eu deixei cada um pensar na variação como queria, né? Não falei assim: “Olha, qual tem maior desvio padrão?”, né? Eu perguntei: “Qual tem a maior variação?”, e de cara quando a gente bate o olho, a gente é levado a achar que é o azul [meninos] né? Pode ver! Eu me lembro bem que a hora que vocês bateram o olho, você falaram: “É o azul, né? Mas vamos pensar”, e foi isso que a maioria falou: “Eu acho que é o de baixo, mas vou pensar”, e aí olha, olha. E a Sandra falou uma coisa... lembra que eu brinquei, anota, anota, anota, que vai pra minha tese. E acho que ela resumiu muito melhor, do que o trabalho que eu li. Ela disse assim: “Se eu olho aqui, todos estão variando, no rosa, então, tem mais variação. No azul, se eu pensar em relação a média, a variação é maior, ela enxergou variação de tipos diferentes. Ela enxergou entre as pessoas, e a variação em relação a média, que são duas maneiras de a gente pensar em variação.

Para discutir a indução de interpretação devido à esta forma de apresentação das observações, a PSQ fez na lousa o gráfico de pontos das notas dos meninos e das meninas e explicou que este gráfico é mais adequado para despertar a variação em torno da média, tal como alertado pelos autores Loosen, Lioen e Lacante (1985).

### 13.8 Análise do raciocínio sobre variação/variabilidade na discussão da análise da média aritmética

As tarefas propostas permitiram observar diferentes raciocínios de variação, que foram classificados de idiossincrático até completo.

Em cada nível de raciocínio, foi possível observar diferentes estratégias e, por esse motivo, foram sintetizadas nos quadros seguintes.

Quadro 36 – Raciocínio idiossincrático sobre variação

| Raciocínio sobre variação | Prof. | Verbalização |
| --- | --- | --- |
| Ausência de variação em torno da média | SB | "Eu chegaria na escola e os professores <u>vão ter</u> esta característica" |
| Média com significado de moda | OB<br>AM | "A maioria. Lógico que nem todos vão ter 39, pode ter 38, por ter até um de 19 anos, mas a maioria tem 39 anos" |
| Inferência da densidade a partir da média e dos valores máximo e mínimo | OB | Se a média é 39 anos, a maioria tem 39 anos, mas nem por isso todos tem 39. Daí, agora, como a média é 39, já vai logo saber que de 19 anos tem pouquinho, 60 tem pouquinho, 40, 43, 35, já dá para saber mais ou menos. |
| Distância do zero: médio ou desvio padrão? | LH | "Quanto mais distante do zero, mais distante os dados estão da média. O zero seria a média?" |
| Variação da média | LF | Média é o ponto de equilíbrio, o centro, e que este centro pode variar |

A ausência de variação em torno da média já havia sido constatada por Ben-Zvi (2004), quando seus alunos estavam elaborando uma hipótese para o tamanho dos nomes.

O significado de média como moda já tinha sido observado por Batanero (2000), porém deve-se ressaltar que esse equívoco prejudica o surgimento natural de uma medida de variação.

Nesta etapa da formação, os professores já tinham elaborado a distribuição de freqüência para a idade dos professores e também já tinham representado-a com um histograma. Porém, essa experiência não foi suficiente para permitir o relacionamento dos conceitos de média, valores máximo e mínimo e formato da distribuição, inferido a partir do histograma. Isto pode ser notado na fala da professora OB, que simplesmente imaginou a densidade da freqüência sem nenhuma referencia à essas relações.

Quanto aos dois últimos raciocínios apresentados no Quadro 35, acredita-se que devido ao fato do campo conceitual de variação estar em desenvolvimento, é natural os deslizes de linguagem.

Os raciocínios de variação considerados verbais estão apresentados no Quadro 37.

Quadro 37 – Raciocínio verbal de variação

| Raciocínio sobre variação | Prof. | Verbalização |
|---|---|---|
| Tendência de eliminar a variação | AM | Trocaria os alunos de turma, para ficar mais homogêneo. |
| Desvio padrão é a variação entre as observações | LF | Desvio padrão é a variação que se tem entre um número e outro |
| Desvio padrão como indicativo da dispersão dos valores | OB | “é só para saber se estão muito dispersos” |
| Maior variação devido valores diferentes | LH | De 10 pra 20, houve uma variação, de 20 pra 30 outra variação, de 30 pra 40... e assim até chegar no 60. Então, número de notas, aqui houve cinco variações e aqui houve uma variação. |
| Se Desvio zero, as observações são iguais | SB | Quando o desvio padrão é zero, as observações são iguais. |
| Variância zero indica homogeneidade | LJ | A variância é zero, é tudo homogêneo |
| Maior variância, mais dispersão das observações | LJ | Quanto maior a variância, mais dispersos estão os dados |
| Desvio padrão próximo do zero indica homogeneidade | AM | Quanto mais próximo o desvio padrão estiver do zero, mais homogêneos os dados |
| Média não ilustra a variação | LF | Admite que a média não retrata a variação da distribuição |

Diferentemente do raciocínio idiossincrático, o raciocínio verbal apresenta corretamente algum componente de variação, mas que necessitaria ser complementado ou relacionado com outros. Podem ser considerados como uma primeira tentativa correta, incompleta, rudimentar sobre a medida de variação.

O Quadro 38 apresenta o raciocínio classificado como transição, pois, pelo menos dois componentes de variação precisariam ser declarados de maneira correta.

É possível notar que os professores apresentaram diferentes raciocínios de transição, dependendo do que estavam tentando compreender, ou seja, com os conceitos disponíveis que tentavam fazer relações.

Quadro 38 – Raciocínio de transição de variação

| | | |
|---|---|---|
| Admite a variação, mas não percebe a necessidade de uma medida | OB<br>AM<br>LF | “A maioria. Lógico que nem todos vão ter 39, pode ter 38, por ter até um de 19 anos, mas a maioria tem 39 anos” |
| Média complementada pela moda | OB<br>LJ<br>LF | Moda zero e dez – média cinco;<br>Moda cinco - média cinco. |
| Média complementada pelo gráfico | AM<br>LH | O professor AM disse que apresentaria a média e o gráfico (possivelmente o histograma) |
| Média complementada pelos valores máximo e mínimo | OB | “entre os professores pesquisados, eles tinham de 19 a 60 anos” |
| Percepção das distâncias da média | AM | Observando a média, 35, analisou que uma distribuição tinha três com 10 e três com 60 enquanto a outra também tinha dez e sessenta, mas tinha 20, 30, 40 e 50. |

Apenas um raciocínio de procedimento e um raciocínio completo foram observados, ambos apresentados pela professora SB. No primeiro caso, ela apresentou um exemplo de uma análise, ainda rudimentar, de um intervalo em torno da média: “eles tem 37, 40”, quando a PSQ perguntou o que significaria um desvio padrão um para uma média 39.

O seu raciocínio completo sobre variação, que havia sido comentado na formação continuada que estaria na tese, foi: “Aqui, a cada aluno varia a nota, né, tem mais variações, são seis variações, seis notas diferentes. E aqui são três notas iguais, três notas iguais, é maior variação em relação à média.”

E este raciocínio da professora SB reflete exatamente o que já havia sido observado por Makar e Confrey (2005), a distinção entre a variação como um predicado: “está disperso” e a variação com um substantivo: “a variação é grande”.

## 14 Raciocínio sobre variação/variabilidade na interpretação do desvio padrão.

Até este momento da formação, os seguintes aspectos de variação possivelmente já estariam disponíveis: existe variação em um conjunto de observações, representar este conjunto só com a média não permite identificar a variação e esta variação é medida pelo desvio padrão.

No que tange à compreensão do desvio padrão, para alguns professores significava a medida de variação entre as observações e não em relação à média e alguns entenderam que quanto maior o valor do desvio padrão, mais variação existe no conjunto de observações.

Muitos professores verbalizaram que, realmente, queriam entender o que eram estas medidas de variação. No término da sessão de 10 de Junho, foi deixada a tarefa apresentada na Figura 77.

1) Calcule o desvio padrão da idade (com todas as observações), utilizando o Microsoft Excel para realizar os cálculos parciais;

2) Calcule o desvio padrão com dados agrupados, a partir da distribuição de freqüência utilizada para calcular a média;

3) Compare os dois valores obtidos para o desvio padrão;

4) Elabore uma análise do desvio padrão.

Figura 77 – Tarefa 1 da análise dos dados do ciclo investigativo – interpretação do desvio padrão – deixada em 10/06/2005 e discutida em 17/06/2005.

Três professores realizaram, parcialmente, a tarefa descrita na Figura 77 e, então, decidiu-se iniciar o encontro de 17 de Junho com a exploração das respostas obtidas. O professor AM, não apenas calculou o desvio padrão da idade dos professores, cujo valor obtido foi 9,8 anos, mas também encontrou um material na internet que explicava o cálculo da nota do aluno no vestibular a partir da padronização por disciplina; a professora LF consultou o livro do Levin (1977) cuja reedição foi Levin e Fox (2004) sobre a porcentagem de observações sob uma curva normal e o professor LH consultou uma apostila de professor (na Internet) sobre as medidas de variação, que apenas descrevia seu cálculo.

Diante do material consultado pelos professores, foi perguntado como eles explicariam para seu aluno o que era o desvio padrão e os professores AM, LF e os alunos LJ e AG verbalizaram que se tivessem que explicar o desvio padrão para os alunos, eles “pulariam” esta aula. O único professor que se aventuraria a explicar o desvio padrão foi LH, mas se limitaria a explicar os cálculos.

Então, a pergunta foi alterada e o diálogo seguinte ilustra as respostas de alguns professores.

> PSQ: Então, a gente está num ponto que é o seguinte: a gente precisa entender o que é o desvio padrão.
>
> LF: É, o que eu tive a maior dificuldade para entender é o que é o significado desse número final que você chega. O que representa na nossa pesquisa? Quando eu fui conseguir interpretar foi através daquele gráfico que eu vi aqui no livro [referindo-se à curva normal]. Ele dá a média, então, a partir dessa média, os valores positivos a direita, que ele chama de sigma, então uma unidade, duas unidades, então aí é que <u>eu consigo ver a amplitude</u>, aonde é que esse número está se encaixando. Acima da média ou abaixo da média. Como ele está plotado. Foi o que eu melhor consegui entender aqui. <u>Ainda tive muita dificuldade de interpretar o que significava esse número.</u>

A professora LF demonstra sua tentativa de compreender o intervalo em torno da média. Ela entendeu que havia a média e alguns valores acima e abaixo dela, dados em unidades de sigma, mas não fez relação com o desvio padrão, que permitiu categorizar seu raciocínio como verbal, pois não permitia que ela somasse e subtraísse o valor do desvio padrão à média.

> LH: Ele [referindo-se ao texto lido] fala que é a raiz quadrada da média aritmética....
>
> LF: O processo matemático a gente consegue entender, mas o que eu não estava entendendo é qual a representação desse bendito desse número dentro da pesquisa que eu estou fazendo.
>
> AM: Foi isso que eu falei [referindo-se à não compreensão do texto obtido na Internet]
>
> PSQ: O que você vai falar para o ET, não é?
>
> LF: Desvio padrão é o desvio padrão. Pronto! Daí, colocar isso numa frase, também fica mais complexo. Então, eu colocaria o gráfico, o cara olhava e conseguiria até definir alguma coisa.

Esse diálogo mostra que o professor LH estava se preocupando com a definição de procedimento enquanto que a professora LF e o professor AM estavam interessados em saber o que significava o número obtido pela fórmula do desvio padrão. A PSQ perguntou diretamente aos alunos e o diálogo seguinte ilustra a resposta deles.

> LJ: Eu, primeiro eu pensaria se eu deveria explicar. Não, eu não sei como eu explicaria!
>
> AG: Só no segundo semestre. Meu professor disse assim: tem o desvio padrão, mas vocês só vão saber o que significa no segundo semestre, quando vocês vão fazer Estatística 2.
>
> LJ: O que a gente tem idéia é que o desvio é....
>
> AG: A gente acabou tudo. Aí no próximo semestre a gente vai fazer inferência estatística e aí a gente vai saber tudo.
>
> LJ: A gente está no "achismo" ainda.
>
> PSQ: Tudo bem. Nesse "achismo", como você explica o que é o desvio padrão?
>
> LJ: Eu prefiro não falar. Achar por achar....Se eu tivesse que explicar hoje, eu não explicaria. Deixaria para outra pessoa explicar.
>
> LF: Da frase dela aqui já dá para iniciar. Eu explicaria que, a partir da média, do ponto base que é média, eu faria uma demonstração, por exemplo, que nem ela fez ali. Que existe a média de idade, por exemplo, 35 anos, mas existe maior quantidade de 30, não é? Eu não tenho os dados porque eu não participei. Seria tantos de 34, 39, então ficaria muito disperso se você se basear só nos 35. Então eu ia fazer essa demonstração com o aluno, e mostrar que dá para a gente aproximar mais esses valores das idades que persistem. Então, eu procuraria explicar o desvio padrão em relação à média e ia deslocando de um lado para o outro em relação a esses valores que estão se repetindo. Eu ia para esse lado.

É possível observar que a professora LF não compreendeu que, ao colocar a média no centro da distribuição normal, o deslocamento a que se referia era o aumento ou a diminuição de unidades de desvio padrão no valor da média, para a direita ou para a esquerda, respectivamente. Ela explicou que ia deslocando os valores que estão se repetindo, possivelmente se referindo as idades com maior freqüência, ilustrando sua permanência num raciocínio verbal.

Então, a PSQ preferiu retomar a interpretação da média aritmética, para fazer surgir, novamente, a necessidade do desvio padrão.

PSQ: Se vocês tivessem que explicar para o ET, só com a média, vai! Ou para um aluno de vocês. Vocês tem lá um monte de professores, 108 professores, essa aqui é a média de idade. Como vocês explicariam para o aluno ou para o ET o que significa isto? O que significa a média 38,8?

LJ: Eu gostei muito do exemplo que o professor falou na aula, né. Bolo, você pega, junta todo o bolo, e corta um pedaço, quer dizer, cada pedaço vai ter este valor. Vai ter, no caso, vai ter 38 cerejas. Então, deu uma homogeneizada no valor individual.

O aluno LJ explicou detalhadamente o exemplo num outro momento da discussão. Quando o bolo é fatiado, é como se todos os pedaços tivessem o mesmo número de cerejas, que é a média.

AM: Quando eu estou falando da média, eu não estou falando a verdade.

PSQ: Por quê?

AM: Porque é aquela história das idades. Estão muito dispersas, estão muito perto do 19 e muito perto do 60, lá.

PSQ: Você vai dizer para o ET que o que você calculou não é verdade?

AM: Aí eu penso assim. <u>A idéia de desvio padrão, tá dando uma variação grande lá</u>. Isto me faz refletir o que? Olha, pega, analisa todos os pontos, um por um, explica para seu aluno que a média pode ser real em casos específicos e o desvio padrão está te mostrando. Nesse caso aqui, você dá uma observada em todos os dados que você tem, ou neste caso aqui, <u>em que o desvio padrão está muito próximo do zero, você não precisa se preocupar porque a média é verdadeira.</u>

O professor AM faz uma interpretação do desvio padrão como um indicativo de veracidade da média e não dispersão dos dados em torno da média. Permanece com a mesma interpretação que tinha no início da formação continuada, raciocínio verbal, em que o desvio padrão menor apresenta informação mais “verdadeira” porque os valores são mais parecidos uns com os outros. E o professor complementa:

AM: Eu acho que está agilizando o trabalho dele. <u>Calcula o desvio padrão e o valor baixo significa que está tudo bem. Mas se o desvio padrão deu muito longe, tá mostrando que as idades estão muito dispersas, então ai eu já tenho que me preocupar com estas idades aqui</u>.

PSQ: O professor AM está começando a interpretar para o aluno dele o que é o desvio padrão. Que a média sozinha, a gente tem que tomar cuidado. Você está alertando o seguinte: eu tenho que olhar com um olhar desconfiado para a média. O que você está dizendo é que o desvio padrão vai dar uma idéia se os dados estão mais dispersos ou mais próximos. Você está começando a interpretar o desvio padrão.

> LF: Se você fosse explicar, por exemplo, para um aluno, por exemplo, vamos fazer a média da nota dele. Ele tirou 2 numa prova, 7 na outra e 6 na outra. Então, quando você soma tudo isto e divide por 3. Você vai ter uma média 5, mas você vai ter uma diferença muito grande de uma nota para as outras duas. Então, a representação está muito dispersa. Como é que eu vou mostrar para esse aluno que aqui ficou muito baixa e aqui ele está legal. E que a média ficou aqui.

A professora LF está construindo o conceito de desvio padrão. Ela começa a pensar na média aritmética, mas ainda centra sua percepção na diferença entre as observações, o que permite classificar seu raciocínio como verbal.

> AM: Ele tira a média 6 para ser aprovado. Mas nada garante que esse aluno tá bem. Ele pode estar vindo com boas notas e de repente tirar zero. Mas por que ele tirou zero? Agora, para saber, eu tenho duas formas de olhar isto. Ou eu olho aluno por aluno e nota por nota, bimestre por bimestre, ou eu tenho uma saída: desvio padrão vai apontar, vai acusar, vai ser um alarme, vai dizer: opa! Tem uma dispersão de notas aqui. Nesse aluno. Aí eu vou olhar o boletim dele e vou dizer ele tinha zero no primeiro bimestre, mas agora está com 7 e 7. Oba, então ele subiu e eu não vou ter que me preocupar. Porém, pode ser ao contrário. Ele tá com 7, depois 4 depois 1. Opa! O que está acontecendo com ele. Aí eu vou ter que me preocupar com este aluno e não com aquele que subiu a nota.
>
> LF: Ele vai ter nota, mas vai estar em queda.
>
> AM: Exato. Ele vai estar em queda. Você tem dois alunos aqui com a mesma média. Só que um está subindo de produção e o outro está descendo. Com qual que eu vou me preocupar? Com o que está caindo de produção.

O professor AM percebe que a ordem das observações não altera o valor do desvio padrão e que, em uma situação escolar, esta seqüência seria importante. Porém, percebe que o desvio padrão alto é um alerta e que, após este “aviso”, o professor pode recorrer às observações para identificar o que está ocorrendo com o aluno que teve um desvio padrão alto.

Essa explicação do professor AM leva em consideração a média, o desvio padrão, mas observa a diferença entre as notas e não a diferença entre as notas e a média, o que permite classificar seu raciocínio como verbal.

> PSQ: Mas o desvio padrão vai te dar isto?

> AM: Não. Ele vai me mostrar que tem uma dispersão de notas e isto vai me alertar para olhar para este aluno, mas ele não vai me dar a realidade. A realidade vai ser olhar para todas as notas.
>
> LH: Acho que mostra!
>
> AM: Não. O desvio padrão não mostra. <u>Ele mostra a dispersão</u>. Ele não mostra nota por nota. <u>Ele mostra a irregularidade</u>! A gente tem um aluno lá que tá com um no primeiro bimestre, cinco no segundo e nove no terceiro. Só que eu tenho o aluno B, aluno 2, esse aluno tá com 9, 5 e 1. <u>Os dois não estão com a mesma média? Ele está com mesmo desvio padrão?</u>
>
> LH: <u>Eles estão com mesmo desvio, só que invertido</u>.

A explicação do professor AM ratifica a classificação como verbal, pois ele torna explícito a diferença entre as observações.

Pode-se observar uma crescente necessidade de compreender o significado do desvio padrão, pois alguns professores já não tinham dúvidas quanto à maneira de calculá-lo.

Percebendo este tipo de dificuldade, a PSQ solicitou que os professores lessem a interpretação da média e do desvio padrão apresentadas pelos livros: Triola (1999), Martins (2001) e Soares e Siqueira (1999).

Estes livros trabalhavam a análise da variação a partir do coeficiente de variação, a partir do Teorema de Tchebichev e apresentavam o cálculo do escore padronizado *z*. No capítulo de Estatística Descritiva, apenas o livro do Triola (1999) apresentava uma análise do desvio padrão segundo a Distribuição Normal.

A realização desta tarefa, que foi dada conjuntamente com a Tarefa 3 do capítulo anterior, durou trinta minutos, e foi feita em duplas.

## 14.1 Resultados e discussão da análise elaborada pelo grupo dos professores

A maioria dos professores focou-se na leitura e interpretação da média aritmética, identificando questões que foram discutidas na formação continuada e estão apresentadas no capítulo anterior.

As professoras LF e IS discutiram, apenas, sobre as medidas de tendência central, alvo do interesse da IS que ministraria esta aula na semana seguinte.

Os professores RS, RN e LH leram Soares e Siqueira (1999) e cada um elaborou sua frase sobre a interpretação do desvio padrão:

> LH: Desvio padrão são os valores dos dados quantitativos que mostram a variação desses acima ou abaixo da média.
>
> RS: Desvio padrão é a variação em torno da média facilitando a interpretação dos dados em questão
>
> RN: Desvio padrão é a variação numérica de como esses valores estão se comportando em relação à média

É possível observar que estes três professores detiveram-se na definição conceitual, ou seja, estavam motivados para a compreensão dos cálculos para a obtenção do valor do desvio padrão, mas o nível de raciocínio ainda era idiossincrático. Isto pode ser explicado pelo fato destes professores não terem realizado a tarefa de calcular o desvio padrão.

Os professores AM e CI já tinham calculado o desvio padrão, cujo valor obtido foi 9,8 e estavam muito interessados em compreender seu significado. Detiveram-se muito tempo na leitura sobre a interpretação da média aritmética e, com o auxílio da observadora VG, perceberam que havia a oscilação de um desvio padrão para cima e para baixo da média, mas a leitura não permitiu apreender que seria possível estimar a porcentagem de dados neste intervalo.

A frase elaborada por eles permite classificar o raciocínio sobre variação como de procedimento, pois entendem que existe uma variação em torno da média e que essa variação é dada pelo desvio padrão, elaborando um intervalo, sem a compreensão que o intervalo contém a maioria das observações.

> Uma pesquisa realizada com 108 professores da rede estadual de Ensino Fundamental II e Ensino Médio constatou que a idade média é de aproximadamente 39 anos. A moda, ou seja, as idades que aparecem com mais freqüência são 39 e 41 anos (7 vezes cada uma) estão próximas da média.
>
> Já para o desvio padrão é calculado o valor 9,8, que significa que <u>sobre a média há uma variação de dez anos para cima e dez anos para baixo, ou seja, a faixa de variação é de 29 anos até 49 anos.</u>

## 14.2 Resultados e discussão da análise elaborada pelo grupo dos alunos

Os alunos AG e LJ foram os únicos a apreender a noção de intervalo de *n* desvios padrão da média aritmética. A discussão deles para compreendê-lo está apresentada a seguir.

LJ: Coloca ai: desvio é 9. Pega a média soma 9 e tira 9. Vamos achar o intervalo. Deixa o intervalo e vamos achar quantas.... achou. O delta.... Quanto que é a média?

AG: 38,8.

LJ: E o dp?

AG: 9,83.

LJ: E aqui deu 38,8. O intervalo.... 38,8 menos ..... aproximadamente 29. 48,6 [referindo-se à soma da média com um desvio padrão]. A variação no intervalo foi 19,6, certo? Quanto que era o valor? A variação no intervalo? Agora tem que fazer isto dividido por 108 vezes ...

O aluno LJ observou que o livro sugeria um intervalo em torno da média a partir da soma e subtração do desvio padrão à média, raciocínio de procedimento, mas pretendia verificar a razão da amplitude do intervalo pelo número total de observações e não a razão do número de observações no intervalo pelo número total de observações, que pode ser compreendido como raciocínio idiossincrático.

AG: Eu não entendi o que você está fazendo!

LJ: Ué! Calculando pela formulação empírica. Isto vezes 108, dividido por 100 deu 18%. Cadê o livro?

AG: [lendo no livro do Martins, 2001] A média mais ou menos um desvio padrão abrange mais ou menos 60% desta variação. A média mais ou menos dois desvios padrão...

LJ: Mas é outro jeito de chegar na mesma coisa! Cadê o livro?

AG: O desvio padrão não é esta diferença? [diferença em relação à media]?

LJ: Isto. A gente chegou a fazer um exercício deste! Eu não me lembro bem! Aqui, neste intervalo, tem 108 pessoas.

AG: Quem disse?

LJ: Eu acabei de contar! Não! Tem 90. Dividido por 108.

AG: Por quê?

Quando o aluno fez uma nova leitura, ele compreendeu a razão que deveria encontrar, raciocínio de procedimento. Porém, questiona-se se estava entendendo o que estava fazendo. A observação do diálogo seguinte esclarece essa dúvida.

LJ: Por quê é para comparar! De 60 a 80 para ser simétrico. Deu 83.

LJ: Mediana deu 37,5

AG: A média deu 38,8.

LJ: Então, tá ...

AG: Pode escrever aí...

LJ: Pode escrever aí, o caderno é seu.

AG: A gente tem que escrever para o ET.

LJ: Mas o que a gente tem que dizer? Agora, a gente calculou aquele negócio dos oitenta...

AG: Não adianta!

LJ: Adianta sim! Com um desvio, a gente calculou um intervalo. Certo?

AG: Certo!

LJ: A gente calculou que oitenta ..... a gente tinha calculado por probabilidade também? Não tinha?

AG: Não tinha não!

LJ: Preciso de um jornalista. Eu sou péssimo para escrever.

AG: Vamos começar assim: a idade média é....

LJ: Idade média é assunto de história. Eu não sei o que dizer! A gente calculou....

AG: O quê?

LJ: A gente não consegue fazer, cara! A gente precisa de mais informação. Claudia, vem cá, por favor! A gente calculou, pela regra empírica, e a gente chegou em 83,3%, se eu fiz as contas certas...

PSQ: Como você chegou nisto?

LJ: Tem o desvio, calculamos o intervalo...

PSQ: Como você chegou neste número?

LJ: Peguei a freqüência e dividi pelo total, vezes 100. Daí ele fala aqui que a porcentagem é aproximadamente 70 para distribuições simétricas e chegando a 90% para distribuições fortemente simétricas [regra empírica apresentada por Martins, 2001]. O nosso está mais próximo do 90.

A PSQ não estava acompanhando o raciocínio dos alunos e não percebeu a dificuldade. Então, propôs a discussão sobre simetria, que não era o foco da dúvida deles. Depois dessa discussão, o aluno LJ pergunta novamente:

LJ: Então, como eu vou escrever?

PSQ: Vocês montaram o intervalo com um desvio. Ele foi de quanto até quanto?

LJ: 29 a 49.

PSQ: Se a gente se prendesse só este intervalo, a gente está deletando quem tem abaixo de 29 e acima de 49. Então, a gente tem que dizer para ele [ET], não é todo mundo que está neste intervalo.

LJ: Para mim, isto é boa parte da amostra...

PSQ: Sim! E aí você poderia dizer isto mesmo!

LJ: <u>Boa parte da amostra está dentro do intervalo que a gente achou: de 29 a 49 anos.</u>

> PSQ: Então, com nosso aluno, a gente foi melhorando a nossa análise a medida que a gente vai aprendendo novas coisas de Estatística. Só temos que pensar sobre o ensino de Distribuição Normal para o ensino médio.
>
> LJ: Será que ele ia excluir? A gente tá pegando o intervalo ... Eu acho que eu fiz....[refaz as contas] dá 75%.

Os alunos tinham encontrado o intervalo de um desvio padrão da média e tinham contado o número de professores nesse intervalo, mas não estavam associando este número com a estimativa da porcentagem de observações apresentada pelo livro, o que mostra o raciocínio de procedimento.

Para compreender o raciocínio de LJ, foi reproduzida a distribuição de freqüências (Tabela 32) sobre a qual calcularam o desvio padrão. LJ contou o número de professores que havia no intervalo de 29 a 49 anos, a partir da soma das freqüências das seis primeiras classes da distribuição, que correspondia a 90 professores, 83% das observações.

Depois percebeu que havia contado a partir do 19 e não do 29, mas como os dados estavam agrupados, ele contou a partir de 25 até 50, correspondendo a 81 professores, 75% das observações.

Tabela 32 – Distribuição de freqüências das idades, elaborada por AG e LJ

| idade | Freqüência |
|---|---|
| 19 ⊢— 25 | 9 |
| 25 ⊢— 30 | **15** |
| 30 ⊢— 35 | **14** |
| 35 ⊢— 40 | **22** |
| 40 ⊢— 45 | **20** |
| 45 ⊢— 50 | **10** |
| 50 ⊢— 55 | 12 |
| 55 ⊢—⊣ 60 | 6 |
| Total | 108 |

Embora ele tivesse interpretado corretamente a regra empírica do livro de Martins (2001), ele não a recuperou na elaboração da frase, preferindo ficar com o termo “boa parte”, o que ilustra a não compreensão completa do conceito.

Eles foram os únicos que contaram o número de professores no intervalo de um desvio padrão da média, estratégia sugerida por Hart (1984) e ratificada por Loosen, Lioen e Lacante (1985).

### 14.3 *Feedback*

Devido ao fato de que as leituras foram, predominantemente, sobre a média aritmética, no encontro do dia 17 de Junho foi feita a formalização sobre as medidas de tendência central e solicitado que continuassem a pensar em suas frases para representar a idade dos professores pesquisados.

No dia 24 de Junho foi feito o *feedback* sobre as medidas de variação.

Com base no exemplo das notas dos alunos A: 0, 0, 10, 10 e Aluno C: 1, 6, 6, 7, foram discutidos os seguintes aspectos:

- Diferença entre desvio médio e variância (justificativa pela preferência do quadrado das diferenças ao invés do módulo);
- Soma das distâncias (absolutas) em relação à média é sempre zero;
- Arredondamentos e suas conseqüências;
- Diferença entre calcular o desvio padrão com as observações e com dados agrupados em classes;
- Diferenças entre população e amostra e sua conseqüência no cálculo da variância e do desvio padrão (parâmetro ou estimativa);
- O uso da fórmula do desvio padrão do parâmetro ou da estimativa e sua conseqüência na escolha da função “variância” e “desvio padrão” do Microsoft Excel.
- Noções sobre o formato de uma distribuição;
- Noções sobre a Distribuição Normal de Probabilidades;
- Elaboração do intervalo de um e dois desvios padrão da média e sua interpretação pela regra empírica de Martins (2001);
- Elaboração da frase para o ET, relatando sobre a idade dos professores pesquisados utilizando a média e o desvio padrão.

Enquanto a PSQ explicava, os professores apresentaram uma postura muito diferente dos demais encontros. Ficaram quietos, quase não fizeram intervenções e, ao término do encontro, relataram que desconheciam o assunto.

Por esse motivo, só estão transcritos alguns trechos da apresentação. Logo no início do *feedback*, os professores LH e RN, que haviam pensado na frase, verbalizaram:

> LH: É pra... quando escrever essa frase, essas frases, né, é para que qualquer um lesse e entendesse, mas se você começa a falar em mediana, desvio padrão, média, o leigo já vai saber? <u>Então, eu acho que teria aí de trabalhar com as porcentagens, tantos por cento estaria nessa...</u>
>
> RN: <u>Ou tantos professores, mesmo assim de 108, tantos professores</u>...
>
> LH: Mas olhando assim agora, a gente percebeu, eu percebi pelo menos que na faixa <u>de 39 anos pra baixo a concentração de professores é bem maior do que acima</u>. Dessa referência de 39, né?

Pôde-se perceber uma tentativa de encontrar uma porcentagem dos professores, mas sem entender em que intervalo, já que a média não significava maioria, o que permite classificar o raciocínio como verbal.

> AM: Ah, mas tem um detalhe também, né? Até a VG ficar cutucando aí (risos), até cair a ficha que se 9,8 tava variando, os 39 pra cima, os 39 pra baixo, levou um tempo aí, né?
>
> PSQ: Por quê? O que vocês tavam pensando?
>
> AM: Porque eu olhava pro desvio padrão, pra mim só tava mostrando, pra mim só tava mostrando a dispersão. <u>Mas não eu consegui olhar pra média e ver o desvio padrão ali, então tem um indicador ali de 10 pra cima, ou 9,8 pra cima, 9,8 pra baixo.</u> Foi aí que eu arredondei pra 10, aí pensando, bom, quem olhar pra 9,8 imagina que...

Essa verbalização do professor AM ilustra o obstáculo criado pela noção de regularidade, homogeneidade. Ele disse que enxergava a dispersão (entre as observações), mas não conseguia entender como se relacionava com a média. Tanto esse diálogo como o seguinte demonstram a passagem de um nível verbal de variação para um nível de procedimento.

> AM: Então, de novo, né, me senti idiota, né? Faço isso em sala de aula e a VG fala: “Mas e daí?” Tá, dispersando, e daí? Dispersando, e daí, dispersando? Vem mais e aí, e aí, e aí? Aí ela falou: “Não sei<u>, tem que fazer diferença em alguma coisa</u>”. (risos) Diferença do que? Caramba, né?
>
> CI: Isso aqui, eu havia pensado isso desde aquele sábado, <u>quando nós pegamos os valores, eu falei: “Olha, se a gente comparar um com o outro tem uma diferença, mas...”</u>

> PSQ: E naqueles livros que vocês leram na semana passada, que vocês leram lá, lembram que vocês tavam lendo, alguns livros que eu tinha trazido?
> AM: E mais: se você perguntar pros meus alunos desse ano ou do ano passado o que é desvio padrão, eles vão falar só da aproximação do zero lá, a primeira coisa que eu falei aqui: <u>próximo do zero  dados mais concentrados</u>... só isso que eles vão conseguir. Não, eu acho que alguns vão dizer isso aqui que eu falei! Agora... a parte dessa avaliação aí, pra cima pra baixo, vão falar, “porque eu não falei, porque eu nunca vi isso, porque eu não li, não tem nada nos apontamentos da apostila”...
> OB: Pra mim era só um cálculo...

É possível notar, também, que mesmo o professor AM identificando a variação em torno da média, a professora CI permanecia compreendendo a variação entre as observações. E esta interpretação do desvio padrão como regularidade ou homogeneidade foi observada em seis (dos sete) livros didáticos analisados.

Este aspecto também apareceu na frase elaborada pela professora IS, que embora tenha faltado deste encontro, enviou sua frase por intermédio da professora OB, que leu-a.

> OB: Em uma pesquisa realizada com 108 professores das redes particular, municipal e estadual de ensino, mostrou que a faixa etária é de 19 a 60 anos, com média de 39 anos, com predominância de 22 anos e <u>com variação de uma idade para outra de aproximadamente 9 anos</u>. Ela disse aqui ó: “Com predominância é a moda”, ela falou pra mim, né, e “A variação de idade é o desvio padrão”. Isso ela explicou pra mim por telefone. Ainda, né? ela falou: “Passa a frase lá”, né?
> LF: É a mesma da semana passada? [que fizeram juntas]
> OB: Ela disse que completou. Ela colocou aqui <u>a variação de idade que é o desvio padrão</u>. <u>Ela tava tentando descobrir uma palavra para desvio padrão</u>. Que ela diz que a variação de idade, ela diz que é desvio padrão.
> OBS VG: Tá. E 9,79 ela aproximou pra 9?
> OB: É... que tá... deveria ser pra 10, né? No caso, aproximaria para 10. Então falei pra ela...

O assunto de arredondamento já tinha sido discutido, mas é possível notar que permanecem os equívocos acerca do assunto.

A PSQ apresentou os cálculos para a obtenção do desvio padrão e foi possível discutir os aspectos sobre a tarefa de “calcular a medida de variação”, que estão brevemente apresentados nos diálogos seguintes.

A PSQ explicou que, devido à soma das distâncias dos valores em relação à média ser zero, faz-se necessário colocar essas distâncias em módulo ou elevá-las ao quadrado, que dão origem ao desvio médio ou à variância, respectivamente.

> OBS VG: É... eleva ao quadrado pra ficar positivo e também pra ficar derivável! Porque quando trabalha com módulo não dá... a função não é derivável.
>
> AM: Agora tá claro pra mim, a diferença que fez a palavrinha derivável aí, né? Não tinha passado, nunca havia passado pela minha cabeça e eu sempre questionei por que era o quadrado, por que era o módulo. Agora, agora você botou a raiz aí, agora eu falei, caramba, agora é módulo.
>
> LF: É 25 o que? 25? [referindo-se ao valor da variância do Aluno A]
>
> PSQ: Isso, então ó, 25 é uma medida ao quadrado, então a primeira coisa difícil que a gente tem de analisar na variância...
>
> LF: Tem que ter unidade, né?
>
> PSQ: É que você tem nota, e a variação tá em nota ao quadrado. Se você tivesse a altura, você tem centímetros e depois você tem a variação em centímetros ao quadrado.
>
> LH: Isso quer dizer que extrair o... dividir por partes aí pra chegar naquela...
>
> RN: Pra chegar em centímetros.
>
> LH: Variância e variação é a mesma coisa?

E fora explicado que a variância é uma das medidas de variação em torno da média. Foi calculado o desvio padrão e elaborado o intervalo de um desvio padrão da média para as notas dos alunos A e C $(5{,}0 \pm 5{,}0; 5{,}0 \pm 2{,}4)$.

> AM: E entende? (risos) Pois é! Quantas vezes eu falei isso aqui? Quantas vezes eu não perguntei isso aqui, essa história pra cima e pra baixo, quantas vezes, não entende. A hora que você acabar de falar, eu ia te falar isso, e daí, eles falam isso e daí? Qual é o propósito de falar isso aí?

Essa verbalização do professor AM ratifica as inferências anteriores. Ele sabia que tinha que encontrar o intervalo, mas não conseguia entender o motivo dessa necessidade.

A PSQ interpretou a regra empírica apresentada por Martins (2001), apenas para a porcentagem de observações a menos de um e dois desvios padrão e utilizou a idade dos cento e oito professores pesquisados para exemplificar. A média era 38,5 anos e o desvio padrão 9,8 anos que originou o seguinte intervalo de um desvio padrão da média: [28,7; 48,3].

> PSQ: A gente sabe que tem professor que tem menos dessa idade e a gente também sabe que tem professor que tem mais que essa idade. Tá? Então, o que é essa regra empírica? Que é a maneira mais frágil, que ela não tá levando em consideração o formato da distribuição, ela vai dizer assim: olha, pra qualquer conjunto de dados, eu devo ter em torno de 60 e 80% das observações, que no nosso caso são professores. Então, entre <u>60 e 80% dos professores deve ter idade entre 28,7 e 48,3</u>.

A PSQ sugere a contagem do número de observações no intervalo de um desvio padrão da média como estratégia didática com o aluno do ensino médio, explicando que, a partir da realização de algumas atividades como esta, o próprio aluno pode generalizar a concentração de observações no intervalo de um desvio padrão da média.

Devido ao fato que o conceito de Distribuição Normal não havia sido discutido formalmente com os professores, escolheu-se trabalhar com a regra empírica apresentada por Martins (2001). Foi calculado o intervalo de um e dois desvios padrão da média. Como poderia haver dificuldade em compreender a imprecisão da porcentagem (entre 60% e 80%), fez-se o intervalo de dois desvios e percebeu-se que apenas dois professores ficavam fora deste intervalo. Os professores calcularam que a porcentagem de dois em cento e oito (quase 2%) e disseram que este intervalo tinha 98% dos professores.

> PSQ: Então, se nós fôssemos redigir uma frase, sem usar o nome desvio padrão, sem usar o nome de média, nós poderíamos dizer lá pro relatório da escola que nós temos aproximadamente 95% dos professores com idades entre 18,9 e 58,1. A escola ainda pode esperar que ela tenha 5%, que pode ser uns 5% abaixo disso ou acima disso, mas nós sabemos que só está acima disso porque fomos nós que lidamos com os dados. Se nós dermos pra ela só a média, ela vai falar assim: “Ah, então todos os professores vão estar em torno de 40 anos”.

Após a discussão sobre a porcentagem de professores observada e estimada no intervalo de dois desvios padrão da média, percebeu-se a interferência do conceito de moda como impeditivo para o surgimento da necessidade do desvio padrão, ilustrado no diálogo seguinte.

> LH: É, mas aí, eu acho que a informação aí não é (inaudível). Porque a concentração, o maior número de professores não estão acima de 31.
>
> RN: 34 e 44 anos. É onde tem a maior concentração.
>
> PSQ: Não, esse aqui ele dá a maior parte dos dados, ela dá 95, em torno de 95% dos dados...
>
> RN: Como é que a gente considera esse fato de que o maior número de professores estão entre 34 e 44 anos?

A PSQ não entendeu que estavam se referindo às duas classes modais da distribuição de freqüências da idade dos professores, mas a observadora VG fez a intervenção, o que gerou um debate muito interessante sobre o conceito de maioria.

> OBS VG: Antes de você, é, eu tava pensando na... na... na preocupação da RN, quando você faz um desvio padrão, um pra cima e outro pra baixo, você já tem a maioria. E tá dando o mesmo...
>
> PSQ: Já tem a maioria.
>
> OBS VG: 28,7 e 48,3. Vocês falaram, né, de 30 a 44, não era isso? É que é assim: o que vocês tão interpretando como maioria? 50% mais 1? Então, aí, acho que não precisaria fazer todo o estudo estatístico. Passaria a fazer o levantamento das idades.
>
> OBS VG: Então! Mas é isso que tá dando entre o desvio padrão! Tá dando a concentração porque dá entre 60 e 80%. É isso aí. Você tá fazendo um estudo científico em cima dos dados. Porque, pra dizer só a maioria não precisaria fazer desvio padrão nem nada. Entendeu? Desvio padrão tá dando uma informação científica a respeito da distribuição.

A professora RN relatava sobre a maioria obtida a partir da observação das duas classes que tinham a maior freqüência (a classe de 34 a 39 anos tinha 22 professores e a classe de 39 a 44 também tinha 22 professores). Essa maioria representava quarenta e quatro professores entre os cento e oito pesquisados, ou seja, aproximadamente 41% da amostra. Se fosse utilizar o conceito de maioria como metade mais uma unidade, o intervalo obtido pela professora RN não retratava essa maioria.

Quando a PSQ falava de maioria, ela estava se reportando à maioria no intervalo de um desvio padrão da média (60% a 80% pela regra empírica) ou a maioria no intervalo de dois desvios padrão da média (95%).

Para encerrar a discussão sobre maioria, foi retomada a primeira análise realizada pela professora OB, quando disse que as idades estavam entre 19 e 60 anos, utilizando para isto os valores máximo e mínimo da distribuição.

> PSQ: Ou falaria assim: eu daria a média e daria o menor e o maior. Se você pegar o maior e o menor, você já tá comparando. Tá todo o mundo...
>
> OBS VG: Aí é 100%.
>
> PSQ: Aí é 100%. Não tem ninguém fora. Quando você faz o desvio padrão você escolhe a porcentagem, né? E isso a gente tá vendo a regra empírica. A gente tem outras regras, né, pra gente discutir.

É possível observar que o diálogo aconteceu somente entre a PSQ e a OBS VG, o que sugere a dificuldade dos professores em compreender o assunto, que é confirmado no diálogo seguinte.

> AM: E você, me fala, escreve uma fase? Faz três semanas, quatro semanas. Escreve a frase como se você tem todas essas...
>
> LH: Desse jeito parece porcentagem, não é?
>
> OB: Mas sabe por que isso aqui é legal? Que linguagem eu vou usar pra que todos entendam? Por isso que foi interessante. Se nem eu sabia o que era o desvio padrão, né? Eu sabia que era um cálculo só. Aí, depois quando eu fiquei sabendo o que era desvio padrão, como é que eu vou fazer, como é que eu vou escrever com palavras...
>
> AM: Não, mas é, pois é, depois dessa explanação ali da Cláudia de hoje, quer dizer, você para e vê que realmente não dá pra fazer a frase. Tem um monte de coisa por trás disso que como é que eu ia escrever? Quer dizer, escrevi. Aquilo lá eu já vou rasgar. (risos)

Pretendia-se continuar o debate sobre o conceito de desvio padrão como um intervalo em torno da média, mas o tempo cedido para a implementação desta pesquisa já havia se esgotado. Entretanto, foi possível perceber que os professores desconheciam ou não se lembravam do intervalo composto pela média mais ou menos $n$ desvios padrão, tão empregado em muitas áreas do conhecimento.

Isto mostra a necessidade de continuação dessa pesquisa, para explorar o raciocínio sobre variação quando utilizam o intervalo de $n$ desvios padrão da média.

## 14.4 Nível do raciocínio sobre variação na interpretação do desvio padrão

Durante a realização dessa atividade, apenas dois raciocínios foram considerados idiossincráticos. Um deles foi a simples leitura de que o desvio padrão era uma medida de variação em torno da média, que não permite verificar a apreensão desse conceito. A outra interpretação considerada idiossincrática foi a análise do aluno LJ, que calculou a razão da amplitude do intervalo de um desvio padrão da média pelo número total de observações.

O raciocínio mais verbalizado dizia respeito à variação entre as observações e não em torno da média, interpretado como raciocínio verbal. Este fato pode ter sido em decorrência de que essa é a análise predominante nos livros didáticos consultados por esses professores. Mas, deve-se levar em consideração que o mesmo tipo de raciocínio foi apresentado nos estudos analisados no Capítulo 4 deste trabalho.

Outras verbalizações categorizadas como raciocínio verbal estão apresentadas no Quadro 39, que ilustram a complexidade do conceito de desvio padrão.

Muitas verbalizações sobre uma ou duas dimensões do conceito foram categorizados como raciocínio verbal, quando entendeu-se que não permitiriam ao professor efetuar qualquer procedimento com o desvio padrão e nem relacionar com suas outras dimensões.

Vale destacar que o professor AM permaneceu com a interpretação de que o desvio padrão baixo é bom. Certamente, na maioria das situações cotidianas, o fato de o desvio padrão ser baixo favorece uma série de “processos”, inclusive o educacional, mas isto não pode ser generalizado.

Quadro 39 – Raciocínio Verbal sobre variação na interpretação do desvio padrão

| raciocínio sobre variação | Prof. | Verbalização |
|---|---|---|
| Desvio padrão como uma medida da diferença entre os valores | LF | Ele tirou 2 numa prova, 7 na outra e 6 na outra. Então, quando você soma tudo isto e divide por 3. Você vai ter uma média 5, mas você vai ter uma diferença muito grande de uma nota para as outras duas. |
| | AM | desvio padrão vai apontar, vai acusar, vai ser um alarme, vai dizer: opa! Tem uma dispersão de notas aqui. |
| | AM | Ele mostra a irregularidade |
| | CI | quando nós pegamos os valores, eu falei: “Olha, se a gente comparar um com o outro tem uma diferença, mas...” |
| | AM | próximo do zero dos dados mais concentrados |
| | IS | e com variação de uma idade para outra de aproximadamente 9 anos |
| | IS | OB: Ela disse que completou. Ela colocou aqui a variação de idade que é o desvio padrão. Ela tava tentando descobrir uma palavra para desvio padrão. Que ela diz que a variação de idade, ela diz que é desvio padrão. |
| Desvio padrão baixo é bom | AM | em que o desvio padrão está muito próximo do zero, você não precisa se preocupar porque a média é verdadeira. |
| | AM | Calcula o desvio padrão e o valor baixo significa que está tudo bem. Mas se o desvio padrão deu muito longe, tá mostrando que as idades estão muito dispersas, então ai eu já tenho que me preocupar com estas idades aqui |
| Tentativa de encontrar um intervalo | LF | a partir dessa média, os valores positivos a direita, que ele chama de sigma, então uma unidade, duas unidades, então aí é que eu consigo ver a amplitude, aonde é que esse número está se encaixando. Acima da média ou abaixo da média |
| | LF | o desvio padrão em relação à média e ia deslocando de um lado para o outro em relação a esses valores que estão se repetindo. |
| | LH | Desvio padrão são os valores dos dados quantitativos que mostram a variação desses acima ou abaixo da média. |
| | LH<br>RN | Então, eu acho que teria aí de trabalhar com as porcentagens, tantos por cento estaria nessa...<br>Ou tantos professores, mesmo assim de 108, tantos professores |
| | LH | de 39 anos pra baixo a concentração de professores é bem maior do que acima |
| Desvio padrão alto é um alerta | AM | Só que eu tenho o aluno B, aluno 2, esse aluno tá com 9, 5 e 1. Os dois não estão com a mesma média? Ele está com mesmo desvio padrão?<br>LH: Eles estão com mesmo desvio, só que invertido. |

O Quadro 40 apresenta os raciocínios classificados como de procedimento.

Quadro 40 – Raciocínio de procedimento sobre variação na interpretação do desvio padrão

| Raciocínio sobre variação | Prof. | Verbalização |
|---|---|---|
| Intervalo de um desvio da média | AM e CI | sobre a média há uma variação de dez anos para cima e dez anos para baixo, ou seja, a faixa de variação é de 29 anos até 49 anos. |
| | AM | Mas não eu consegui olhar pra média e ver o desvio padrão ali, então tem um indicador ali de 10 pra cima, ou 9,8 pra cima, 9,8 pra baixo. |
| Alguma quantidade de obs no intervalo de um desvio padrão da média | LJ | Boa parte da amostra está dentro do intervalo que a gente achou: de 29 a 49 anos. |

A compreensão de um intervalo de *n* desvios padrão da média é um grande avanço no raciocínio sobre variação, que se mostra aqui não ser uma tarefa tão intuitiva quanto era imaginado pela autora deste trabalho. E a compreensão de que existe uma grande quantidade de observações nesse intervalo só foi observada no aluno LJ, que após a contagem do número de professores no intervalo, concluiu que “boa parte” da amostra estava contida nesse intervalo.

Realmente é muito difícil entender que analisar o desvio padrão significa analisar a média aritmética também. Isto pode ser exemplificado pelos gráficos elaborados pelo professor AM para representar as notas de importância da Estatística. Estes gráficos foram enviados para a PSQ em 10 de Setembro de 2005.

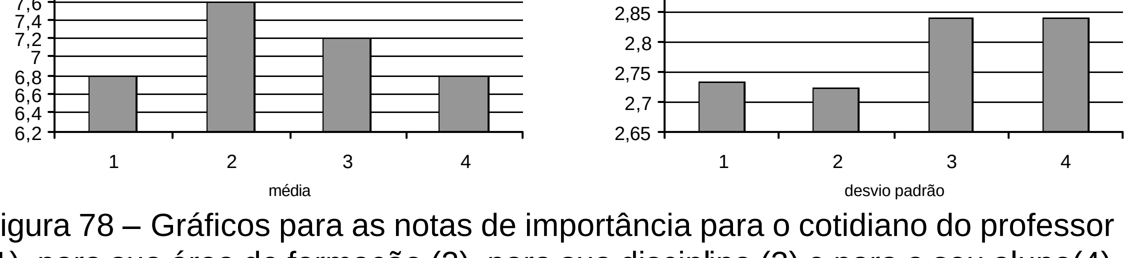



Figura 78 – Gráficos para as notas de importância para o cotidiano do professor (1), para sua área de formação (2), para sua disciplina (3) e para o seu aluno(4).

É possível observar na Figura 78 que, além do fato da média estar separada do desvio padrão, que permite inferir a não compreensão dos dois conceitos conjuntamente, as escalas dos gráficos são diferentes, o que não permite uma análise visual da dimensão do desvio padrão. Isto permite entender que o raciocínio de procedimento apresentado anteriormente, quando este professor encontrou o intervalo de um desvio padrão da média foi induzido pela observadora e não apreendido pelo professor.

Entretanto, este professor desenvolveu seu pensamento estatístico, pois quando enviou estes gráficos explicou que fez a análise da importância de Estatística para todos os professores, que está apresentado na Figura 78, e fez a mesma análise por disciplina ministrada e pôde perceber que os professores de Biologia atribuíam muito mais importância para a Estatística do que os outros professores.

Este fato mostra a prontidão deste professor para investigar as causas da variação, analisá-las mediante o conhecimento estatístico que possuía e elaborar uma conclusão. Para isso, vivenciou o ciclo investigativo, transitou, mesmo que parcialmente pelo pensamento geral e fundamental do pensamento estatístico, estava no ciclo interrogativo e apresentava suas primeiras disposições acerca dos resultados. Em outras palavras, seu pensamento “navegava” em todas as dimensões do pensamento estatístico definido por Wild e Pfannkuch (1999).

## 15 Considerações Finais

A dificuldade apresentada pelos alunos de graduação da autora para a compreensão do desvio padrão e a oportunidade de assistir a apresentação das possibilidades para ensinar o intervalo de confiança com o uso do *Fathom* (no ICOTS-6, em Julho de 2002) foram os principais motivos pelos quais deu-se início a este trabalho.

Os resultados obtidos não permitem uma solução inovadora para o ensino de desvio padrão, mas certamente apontam a complexidade do conceito a partir das diferentes estratégias de raciocínio que surgiram.

Considerando que os participantes da pesquisa eram professores de Matemática da escola básica e, portanto, já tinham aprendido sobre desvio padrão, a fase inicial da pesquisa identificou a ausência de significação desse conceito, exceção feita ao professor AM, que apresentava um raciocínio idiossincrático de variação.

As dimensões do pensamento estatístico (WILD e PFANNKUCH, 1999) serviram de estrutura conceitual e como organização da implementação desta pesquisa-ação, pois os autores declararam ser a variação seu componente fundamental. Portanto, o objetivo deste trabalho foi identificar o raciocínio de variação/variabilidade dos professores em cada etapa do ciclo investigativo, de maneira que permitisse compreender a relação entre pensamento estatístico e raciocínio sobre variação.

Foi observado que os poucos professores que já utilizavam a pesquisa como metodologia de ensino de Estatística, o faziam de maneira restrita, utilizando apenas a distribuição de freqüência e sua respectiva representação gráfica para analisar os resultados, o que indicou a não abordagem do conceito de variação em suas aulas.

Essa não utilização das medidas de tendência central e dispersão como ferramentas de análise dos dados da pesquisa pode ser em decorrência do tipo de variável utilizada, geralmente qualitativa, tais como sabor de sorvete e estilo musical de preferência, ambas verbalizadas por eles.

Outro fato que merece destaque é a utilização de uma única variável pesquisada, que não permite realizar a comparação de grupos e, portanto, dificulta o estímulo para observar a variabilidade.

Embora alguns professores já fizessem “pesquisa” com seus alunos, pode-se considerar que a fase de planejamento do objetivo da pesquisa, da determinação dos sujeitos, da elaboração do questionário e sua aplicação e da construção do banco de dados foram consideradas atividades novas para esses professores e, foi possível notar, que utilizaram quase completamente o ciclo investigativo, utilizaram parcialmente os tipos de pensamento e o ciclo interrrogativo e, apenas o professor AM apresentou disposições sobre os resultados, ou seja, transitou por todas as dimensões do pensamento estatístico.

Nessa fase de planejamento do ciclo investigativo percebeu-se a omissão da PSQ, que deveria ter influenciado diretamente na elaboração do questionário e, verificando a não condição dos professores em apresentar alternativas para as questões, deveria tê-lo feito. A permissão de utilização de muitas questões abertas prejudicou o desenvolvimento desta pesquisa, pois demandou muito tempo para sua análise e não permitiu que os professores concluíssem o relatório para a escola, sem o que não conseguiriam vivenciar completamente as dimensões do pensamento estatístico.

Uma explicação para este fato era a inexperiência da autora em trabalhar com professores em atuação e o seu receio de induzir a solução da situação, haja vista que sua formação foi no ensino tradicional.

Embora alguns professores tenham desenvolvido quase completamente o pensamento estatístico, isto não implicou diretamente no desenvolvimento do raciocínio sobre variação, como pôde ser observado nos resultados. Mesmo o professor AM, que já estava no ciclo interrogativo e apresentando disposições sobre os resultados (dimensões do pensamento estatístico), seu raciocínio de variação ainda não estava no nível mais avançado.

Isso ratifica a afirmação de Gal (2002), que não é a possibilidade de participar de projetos (pesquisa) que permitem o desenvolvimento do raciocínio sobre variação, mas sim as situações desafiadoras que o professor vai proporcionar ao aprendiz.

Durante a análise dos resultados da pesquisa idealizada e desenvolvida pelos professores foi possível observar a naturalidade do surgimento do raciocínio de variabilidade e a dificuldade do desenvolvimento do raciocínio sobre variação, principalmente nos níveis de procedimento e completo, conforme modelo proposto por Garfield (2002).

No que tange à variabilidade, foi possível observar a facilidade em admitir a existência de variação, porém foi verificado a dificuldade de viver e conviver diante dessa variabilidade. Talvez, o fato mais ilustrativo dessa observação tenha sido o desentendimento dos professores acerca das diferentes percepções de mundo que influenciaram a decisão sobre a categorização das respostas das questões abertas.

O raciocínio sobre variação foi observado em quatro atividades propostas: análise de uma distribuição de freqüências simples, análise da representação de uma variável contínua, na interpretação da média aritmética e do próprio desvio padrão.

Essas quatro atividades permitiram observar o raciocínio em quatro dos sete aspectos do modelo epistemológico elaborado por Garfield e Ben-Zvi (2005), apresentado no Capítulo 4 deste trabalho. Os aspectos foram: desenvolvimento de idéias intuitivas de variabilidade, descrição e representação da variabilidade, uso da variabilidade para fazer comparações e a consideração da variabilidade como parte do pensamento estatístico.

No momento em que os professores compararam três distribuição de freqüências simples de uma variável discreta, eles utilizaram a percepção da moda, a observação dos valores máximo e mínimo, a menor frequencia e elaboraram um intervalo de variação composto pelos valores da variável que apresentavam, pelo menos, freqüência um em todas as distribuições. Esse intervalo pode ser entendido como uma quebra da distribuição, estratégia relatada em alguns estudos publicados, o que demonstra a não utilização da média e do desvio padrão para a comparação de distribuições.

Ainda sobre distribuição de freqüências simples, foi possível observar que a utilização do termo variação levou alguns professores a perceber a variação entre as freqüências de uma mesma categoria e escolher a categoria cujas freqüências eram mais uniforme, ou seja, apresentavam menor variação entre elas, o que não faz sentido na análise dos dados e foi categorizado como raciocínio idiossincrático.

A atividade que mais permitiu a identificação do raciocínio sobre variação foi a solicitação de uma análise de uma distribuição, sem a utilização de recurso gráfico ou tabular. As discussões sobre o significado das medidas de variação permitiram identificar raciocínios equivocados, classificados como idiossincráticos e que um deles merece destaque. A compreensão de média como maioria, que se refere ao conceito de moda, foi um fator impeditivo para a percepção da necessidade de uma medida de variação.

Os professores interpretaram a média da idade, 39 anos, como representativa da maioria, excluindo qualquer necessidade de complementação. No momento em que foi discutido que essa análise não estava correta, surgiram outras estratégias para complementar a informação fornecida pela média, tais como os valores máximo e mínimo e a própria moda, mas não surgiu a necessidade do desvio padrão, confirmando os resultados obtidos nos estudos analisados.

A tarefa de interpretar o desvio padrão permitiu identificar a dificuldade em compreender que é uma medida de variação em torno da média. O raciocínio verbal predominante na análise do desvio padrão referia-se à uma medida das diferenças entre as observações e que pode ser reforçado pelos livros didáticos quando apresentam o desvio padrão como uma medida de homogeneidade da amostra ou como a regularidade dos dados, pois ambas afirmações não remetem um *link* mental para a média aritmética.

O raciocínio da professora SB novamente ilustra as diferentes interpretações possíveis quando se pergunta sobre a variação. Na atividade de comparar as notas de duas distribuições, ela verbalizou: “Aqui, a cada aluno varia a nota, né? tem mais variações, são seis variações, seis notas diferentes. E aqui são três notas iguais, três notas iguais, é maior variação em relação à média”, categorizado como raciocínio completo.

Quando se pronuncia maior variação, pode ser interpretado como variação entre as observações e, raramente, será entendido como variação em torno da média ou de qualquer outra medida de tendência central.

Se o raciocínio de variação leva em consideração a média aritmética, a elaboração do intervalo de *n* desvios padrão da média deveria surgir naturalmente, o que não se verificou neste trabalho, o que demonstra que a verbalização que o desvio padrão era a variação em torno da média não foi apreendida.

Somente o professor AM e o aluno LJ conseguiram elaborar o intervalo de um desvio padrão da média, embora supõe-se que este conceito ainda não tenha sido completamente apreendido pelo professor, haja vista os gráficos distintos de média e desvio padrão elaborados.

Possivelmente a compreensão de uma estimativa da proporção de observações no intervalo de *n* desvios padrão da média fosse arrojada, mas esperava-se que eles realizassem a tarefa de contar o número de casos no intervalo, que o professor AM havia feito com a quebra da distribuição na tarefa dos sobrenomes.

A relação entre o conceito de maioria fornecido pela moda e pelo intervalo de *n* desvios padrão da média e a observação do todo (100%) obtida pela amplitude poderia ser elucidativa e proporcionar um avanço no raciocínio sobre variação se as outras dimensões do deste conceito estivem disponíveis pelos professores.

Pôde-se observar nos professores uma apatia diante dessa discussão e uma forte tendência pela escolha da classe modal da distribuição de freqüência com dados agrupados, que indica que o campo conceitual de desvio padrão ainda não estava completamente formado. Ou seja, existia raciocínio sobre variação, com a mobilização de muitos conceitos, alguns deles relacionados, mas não pôde ser considerado como raciocínio completo de variação, no sentido de Garfield (2002).

Uma das questões que se levanta para próximas pesquisas é se esse tipo de raciocínio pode ser estimulado com aplicativos computacionais que trabalhem especificamente com intervalos em torno da média, semelhante ao aplicativo desenvolvido por Delmas e Liu (2005), que trabalhavam a compreensão da magnitude do desvio padrão ou da ferramenta do *Fathom*, que trabalhava com o intervalo de confiança.

Surge, portanto, uma das limitações dessa pesquisa, que foi a não utilização deste tipo de aplicativo, devido ao não conhecimento das técnicas para desenvolvê-lo. Acredita-se que a formação de uma equipe multidisciplinar de pesquisa possa colaborar nesse sentido.

Outra questão que deixou a autora intrigada e que merece atenção das pesquisas em ensino de Estatística foi a dificuldade apresentada pelos professores para trabalhar com variáveis contínuas, principalmente com a distribuição de freqüências com dados agrupados, embora já tivesse sido salientado por Meletiou (2000).

Quando da discussão sobre esse assunto, as verbalizações foram raras e, quando aconteciam, diziam respeito aos aspectos da própria distribuição de freqüência e não sobre a variação, que permitiu identificar apenas a utilização da classe modal para analisar os resultados, que foi categorizado como raciocínio verbal.

Uma frustração da autora foi não ter tido tempo na fase de implementação da pesquisa-ação para desenvolver o trabalho sobre o raciocínio de variação na análise de um intervalo interquartílico, que pretende ser a continuidade dessa pesquisa.

**Referências Bibliográficas**


ARTIGUE, Michèle. Ingenieria didáctica. In: Artigue, Michele; Douady, Regine; Moreno, Luis e Gómez, Pedro. Ingenieria didáctica en educacion Matematica. Bogotá (Colombia): Iberoamérica, 1995.

BAKKER, Arthur. Reasoning about shape as a Pattern in Variability. *Statistics Education Research Journal*, v. 3, n. 2, p. 64-83, 2004. Disponível em http://stat.auckland.ac.nz/serj. Acesso em dez. 2004.

BARBETTA, Pedro Alberto. Estatística aplicada às Ciências Socias. 2. Ed. Florianópolis: Ed. Da UFSC, 1998.

BATANERO, Carmen. Significado y comprensión de las medidas de posicion central. *UNO*, v. 25, p. 41-58, 2000.

BARBIER, René. *A pesquisa-ação*. Tradução de Lucie Didio. Brasília: Liber Livro, 2004.

BEN-ZVI, Dani. Seventh grade students'sense making of data and data representations. In: PHILLIPS, Brian (org.). *Proceedings of the Sixth International Conference on Teaching Statistics*. South Africa, 2002. International Statistical Institute. CD

_____. Reasoning about Variability in Comparing distributions. *Statistics Education Research Journal*, v. 3, n. 2, p. 42-63, 2004. Disponível em http://stat.auckland.ac.nz/serj. Acesso em dez. 2004.

BEN-ZVI, Dani, GARFIELD, Joan. Research on reasoning about variability: A forward. *Statistics Education Research Journal*, v. 3, n. 2, p. 4-6, 2004. Disponível em http://stat.auckland.ac.nz/serj. Acesso em dez. 2004.

BRASIL. Ministério da Educação. Secretaria de Educação Fundamental. *Parâmetros Curriculares Nacionais: matemática.* Brasília: Ministério da Educação/Secretaria de Educação Fundamental, 1998.

BRASIL, Ministério da Educação. Secretaria de Educação Média e Tecnológica. *Parâmetros Curriculares Nacionais (Ensino Médio) – Parte III: Ciências da Natureza, Matemática e suas Tecnologias.* Brasília: Ministério da Educação, 2000.

BRASIL, Ministério da Educação. Secretaria de Educação Média e Tecnológica. *PCN+ Ensino Médio: Orientações educacionais complementares aos Parâmetros Curriculares Nacionais.* Brasília: Ministério da Educação, 2002.

BROUSSEAU, Guy. *Theory of didactical situations in mathematics*. Tradução de Nicolas Balacheff, Martin Cooper, Rosamund Sutherland e Virginia Warfield. Netherlands: Kluwer, 1997, cap 1 p. 21-75.

BUSSAB, Wilton de Oliveira, MORETTIN, Pedro Alberto. *Estatística Básica*. 5ª ed. São Paulo: Saraiva, 2003.

CANADA, Daniel. Elementary Pré-service Teachers´conceptions of variation in a Probability Context. *Statistics Education Research Journal*, v. 5, n. 1, p. 36-63, 2006. Disponível em http://stat.auckland.ac.nz/serj . Acesso em 30 de dez. 2006.

CARVALHO, Carolina. Olhares sobre a Educação Estatística em Portugal. In: *Anais do Simpósio Internacional de Pesquisa em Educação Matemática*, 2006, Recife. Recife: UFPE, 2006. p. 1-16.

CAZORLA, Irene Mauricio, SANTANA, Eurivalda Ribeiro dos Santos. *Tratamento da informação para o ensino fundamental e médio*. Itabuna: Via Literarum, 2006.

CHEVALLARD, Yves. La fonction professorale: esquisse d'un modele didactique**.** In: NOIRFALISE, Robert (org). *Actes de L'Ècole D'Éte***.** Clermont-Ferrand: IREM, 1995.

______. Conceitos Fundamentais da Didática: as perspectivas trazidas por uma abordagem antropológica. In: BRUM, Jean (org.). *Didáctica das Matemáticas*. Lisboa: Editora Piaget, 1996. p. 115-152.

______.L'Analyse des pratiques enseignantes en Théorie Anthropologique Du Didactique. *Recherches en Didactique des Mathématiques*, v. 19, n. 2, p. 221-266, 1999.

COSTA, Carlos Eduardo, CAPOVILLA, Fernando César. Resolução de Problemas e Raciocínio. *Torre Babel*, v. 4, n. 1, p. 105-130, 1997.

COSTA NETO, Pedro Luiz de Oliveira. *Estatística*. 2 ed. São Paulo: Edgar Blucher, 2002.

COUTINHO, Cileda de Queiroz e Silva. *Introduction aux situations aléatoires dès le Collège: de la modelisation à la simulation d'expériences de Bernoulli dans l'environment informatique Cabri-géomètre II.* 2001. Tese (Doutorado em Educação) Université Grenoble I - Joseph Fourier, Grenoble.

DANTE, Luiz Roberto. *Matemática: Contexto e Aplicações – volume único*. São Paulo: Ática, 2000.

DELMAS, Robert, LIU, Yan. Exploring Student's Conceptions of the Standard Deviation. *Statistics Education Research Journal*, v. 4, n. 1, p. 55-82, 2005. Disponível em http://stat.auckland.ac.nz/serj. Acesso em 2005.

EHRENBERG, A.S.C. Symbols or Concepts? *The Statistician*, v.24, n.3, 1976.

EMBSE, Charles Vonder, ENGEBRETSEN, Arne. Visual Representations of Mean and Standard Deviation. *The Mathematics Teacher*, v.89, n. 8, p. 688-692, 1996.

FRANCO, Maria Amélia Santoro. Pedagogia da Pesquisa-Ação. *Educação e Pesquisa*, v. 31, n.3, p. 483-502, 2005.

GAL, Iddo. Adult´s Statistical Literacy: Meanings, Components, Responsabilities. *International Statistical Review*, v. 70, n. 1, p. 1-25, 2002.

GARFIELD, Joan. The Challenge of Developing Statistical Reasoning. *Journal of Statistics Education*, v. 10, n.3, p. 1-11, 2002. Disponível em http://amstat.org/publications/jse/v10n3/garfield.html . Acesso em 24 de set. 2006.

GARFIELD, Joan, BEN-ZVI, Dani. A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, v. 4, n. 1, p. 92-99, 2005. Disponível em http://stat.auckland.ac.nz/serj. Acesso em 2005.

GENTIL, Nelson, SANTOS, Carlos Alberto Marcondes dos, GRECO, Antonio Carlos, BELOTTO FILHO, Antonio e GRECO, Sergio Emilio. *Matemática para o 2º grau. Volume 2*. São Paulo: Ática, 1996.

GIOVANNI, José Ruy, BONJORNO, José Roberto e GIOVANNI JUNIOR, José Ruy. *Matemática Fundamental: uma nova abordagem: ensino médio: volume único.* São Paulo: FTD, 2002(a).

______. *Matemática Fundamental: uma nova abordagem: Guia Pedagógico.* São Paulo: FTD, 2002(b).

GUIMARÃES, Rui Campos, CABRAL, José A. Sarsfield. *Estatística.* Edição Revista. São Paulo: McGraw-Hill de Portugal, 1997.

HAMMERMAN, James K. , RUBIN, Andee. Strategies for managing Statistical Complexity with New Software Tools. *Statistics Education Research Journal*, v. 3, n. 2, p. 17-41, 2004. Disponível em http://stat.auckland.ac.nz/serj. Acesso em dez. 2004.

HART, Anna. How Should We Teach the Standard Deviation? *Teaching Statistics*, v. 6, n. 1, p. 24-25, 1984.

IEZZI, Gelson, DOLCE, Osvaldo, DEGENSZAJN, David, PERIGO, Roberto e ALMEIDA, Nilze de*. Matemática: Ciência e Aplicações, vol. 3.* São Paulo: Atual, 2001

JAMES, Barry R. *Probabilidade: um curso em nível intermediário*. 3 ed. Rio de Janeiro: IMPA, 2004.

KONOLD, Clifford , POLLATSEK, Alexander. Data Analysis as the Search for Signals in Noisy Processes. *Journal for Research in Mathematics Education*, v. 33, n. 4, p. 259-289, 2002.

LEHRER, Rich , SCHAUBLE, Leona. Distribuition: a resource for understanding error and natural variation. In: PHILLIPS, Brian (org.). *Proceedings of the Sixth International Conference on Teaching Statistics*. South Africa, 2002. International Statistical Institute. CD.

LEVIN, Jack. *Estatística para Ciências Humanas*. São Paulo: Harbra, 1977.

LEVIN, Jack e FOX, James Alan. *Estatística para Ciências Humanas*. 9 ed. São Paulo: Prentice Hall, 2004.

LIMA, Rosana Nogueira. *Resolução de Equações Algébricas.* [2007]. Tese (Doutorado em Educação Matemática) - PUC-SP, São Paulo.

LOOSEN, Franz, LIOEN, Monique, LACANTE, Margareta. The Standard Deviation: some Drawbacks of na Intuitive Approach. *Teaching Statistics*, v. 7, n. 1, p. 2-5, 1985.

MAGALHÃES, Marcos Nascimento , LIMA, Antonio Carlos Pedroso. *Noções de Probabilidade e Estatística*. 6 ed. São Paulo: Ed. Da USP, 2004.

MAKAR, Katie, CONFREY, Jere. "Variation-Talk": Articulating meaning in Statistics. *Statistics Education Research Journal*, v. 4, n. 1, p. 27-54, 2005. Disponível em http://stat.auckland.ac.nz/serj. Acesso em 2005.

MARTINS, Gilberto de Andrade. *Estatística Geral e Aplicada*. São Paulo: Atlas, 2001.

MEIRA, Luciano de Lemos, DIAS, Maria da Graça, SPINILLO, Alina Galvão. Raciocínio lógico-matemático: aprendizagem e desenvolvimento. Temas em Psicologia, n.1, p. 113-127, 1993.

MELETIOU, Maria. *Developing Student's Conceptions of Variation: an Untapped well in Statistical Reasoning*. 2000. Tese (Doutorado em Filosofia) - Universidade do Texas, Austin, 2000.

MELETIOU, Maria e LEE, Carl. Student understanding of histograms: a stumbling stone to the development of intuitions about variation. PHILLIPS, Brian (org.). *Proceedings of the Sixth International Conference on Teaching Statistics*. South Africa, 2002. International Statistical Institute. CD

MELETIOU-MAVROTHERIS, Maria e LEE, Carl. Teaching Students the Stochastic Nature of Statistical Concepts in na Introductory Statistics Course. *Statistics Education Research Journal*, v. 1, n. 2, p. 22-37, 2002.

MEYER, Paul. *Probabilidade: Aplicações à Estatística*.Rio de Janeiro: LTC, 2003.

MOOD, Alexander M., GRAYBILL, Franklin A. E BOES, Duane C. *Introduction to the theory of statistics*. Rio de Janeiro: Mc Graw-Hill, 1974.

MOORE, David. New Pedagogy and New Content: The case of Statistics. *International Statistical Review*, v. 65, n. 2, p. 123-137, 1997

MOREIRA, Marco Antonio. A teoria dos Campos Conceituais de Vergnaud, o Ensino de Ciências e a Pesquisa nesta área. *Investigações em Ensino de*

*Ciências*, v.7, n.1, p.1-24, 2002. Disponível em http://www.if.ufrgr.br/public/ensino/vol7/n1/v7_n1_a1.html, acesso em 06/2006.

NIFOCI, Renata Ercília Mendes, SILVA, Cláudia Borim da. Teorema de Pitágoras: uma aprendizagem significativa? In: Sociedade Brasileira de Educação Matemática. *Resumos – VII Encontro Paulista de Educação Matemática*. São Carlos, 2004.

PAIVA, Manoel Rodrigues. *Matemática*. São Paulo: Moderna, 1995. vol.2.

PFANNKUCH, Maxine. Thinking tools and variation. *Statistics Education Research Journal*, v. 4, n. 1, p. 83-91, 2005. Disponível em http://stat.auckland.ac.nz/serj. Acesso em 2005.

PONTE, João Pedro, BROCARDO, Joana, OLIVEIRA, Hélia. *Investigações Matemáticas na sala de aula*. Belo Horizonte: Autêntica, 2003, cap. 5, p. 91-108.

READING, Chris. Student description of Variation while working with Weather Data. *Statistics Education Research Journal*, v. 3, n. 2, p. 84-105, 2004. Disponível em http://stat.auckland.ac.nz/serj. Acesso em dez. 2004.

READING, Chris e SHAUGHNESSY, Michael. Reasoning about Variation. In: BEN-ZVI, Dani e GARFIELD, Joan. *The challenge of developing Statistical Literacy, Reasoning and Thinking*. Netherlands: Kluwer, 2004. p. 201-226.

RODRIGUES, Flavio Wagner. A Estatística e as pesquisas eleitorais. *Revista do Professor de Matemática*, v. 40, p. 23-30, 1999.

ROSSINI, Renata. *Saberes docentes sobre o tema Função: uma investigação das praxeologias*. 2006. Tese (Doutorado em Educação Matemática) - PUC-SP, São Paulo, 2006.

RUMSEY, Deborah J. Statistical Literacy as a goal for Introductory Statistics Courses. *Journal of Statistics Education*, v. 10, n. 3, p. 1-12, 2002. Disponível em: http://www.amstat.org/publications/jse/v10n3/rumsey2.thml . Acesso em 2006.

SILVA, Cláudia Borim da. *Atitudes em relação à Estatística*: um estudo com alunos de graduação. 2000. Dissertação (mestrado em Educação) Faculdade de Educação, UNICAMP, Campinas.

SILVA, Cláudia Borim da, COUTINHO, Cileda de Queiroz e Silva. O nascimento da Estatística e sua relação com o surgimento da Teoria de Probabilidade. *Revista Integração*, n. 41, p.191-196, 2005. disponível em ftp://ftp.usjt.br/pub/revint/191_41.zip

SILVA, Maria José Ferreira da. *Investigando saberes de professores do ensino fundamental com enfoque em números fracionários para a quinta série.* 2005. Tese (Doutorado em Educação Matemática) - PUC-SP, São Paulo, 2005.

SMOLE, Kátia Cristina Stocco e DINIZ, Maria Ignez de Souza Vieira. Matemática – volume 1 – 1a série – ensino médio. 3. ed. Reform. São Paulo: Saraiva, 2003(a).

______. Matemática – volume 2 - ensino médio. 3. ed. Reform. São Paulo: Saraiva, 2003(b)

______. Matemática – volume 3 – ensino médio. 3. ed. Reform. São Paulo: Saraiva, 2003(c).

SNEE, Ronald D. Statistical Thinking and its contribution to Total Quality. *The American Statistician*, v. 44, n. 2, p. 116-121, 1990.

SOARES, José Francisco, SIQUEIRA, Arminda Lucia. *Introdução à Estatística Médica*. Belo Horizonte: Departamento de Estatística – UFMG, 1999.

SOARES, Magda. Letramento e Alfabetização: as muitas facetas. *Revista Brasileira de Educação*, n. 25, p.1-13, 2004. disponível em: http:// www.scielo.br/pdf/rbedu/n25/n25a01.pdf. Acesso em: 29 de dez. 2006.

TOROK, Rob. Chance and Data. *Australian Mathematics Teacher*, v. 56, n. 2, p. 25-31, 2000.

TIBONI, Conceição Gentil Rebelo. *Estatística Básica para o curso de Turismo*. São Paulo: Atlas, 2002.

TRIOLA, Mario F. *Introdução à Estatística*. 7ª ed. Rio de Janeiro: LTC, 1999.

TRIPP, David. Pesquisa-ação: uma introdução metodológica. *Educação e Pesquisa*, v. 31, n. 3, p. 443-466, 2005.

YOSSEF, Antonio Nicolau, SOARES, Elizabeth e FERNANDEZ, Vicente Paz. *Matemática: volume único para o ensino médio*. São Paulo: Scipione, 2004. (Coleção De olho no mundo do trabalho).

WALTON, Douglas N. What is reasoning? What is an argument? *The journal of Philosophy*, v. 87, n. 8, p. 399-419, 1990.

WATSON, Jane. Statistical Literacy at the School Level: What should students know and do? In: *Proceedings of the 54º Bulletin of the International Statistical Institute*, 2003, Berlim. Berlim: ISI, 2003. p. 1-4.

WATSON, Jane, CALLINGHAN, Rosemary. Statistical Literacy: a complex hierarchical construct. *Statistics Education Research Journal*, v. 2, n. 2, p. 3-46, 2003. Disponível em: http://fehps.une.edu.au/serj . Acesso em 2006.

WATSON, Jane, KELLY, Bem A. Can grade 3 students learn about variation? PHILLIPS, Brian (org.). *Proceedings of the Sixth International Conference on Teaching Statistics*. South Africa, 2002. International Statistical Institute. CD

WILD, Chris, PFANNKUCH, Maxine. Statistical Thinking in Empirical Enquiry. *International Statistical Review*, v. 67, n. 3, p. 223-265, 1999.

ZAMPIROLO, Maria José C.V., SCORDAMAGLIO, Maria Terezinha e CÂNDIDO, Suzana Laino. *As estatísticas revelam...* São Paulo: Editora do Brasil, 2000a. (Projeto Escola e Cidadania: Matemática).

ZAMPIROLO, Maria José C.V., SCORDAMAGLIO, Maria Terezinha e CÂNDIDO, Suzana Laino. *As estatísticas não mentem jamais?* São Paulo: Editora do Brasil, 2000b. (Projeto Escola e Cidadania: Matemática).

**Apêndice 1: Independência de variáveis aleatórias**

Antes de explicar o que é independência de variáveis aleatórias, considera-se importante a compreensão de probabilidade condicional, que será apresentada a partir de um exemplo apresentado na Tabela 33.

Considere os eventos A – pessoa escolhida ser mulher e B – pessoa escolhida não fumar. Suponha que tenha sido retirada uma amostra aleatória de alunos de uma universidade, cuja variável aleatória X é gênero e a variável Y é a opção em relação ao fumo de cigarros. Os dados (fictícios) estão apresentados na tabela a seguir.

Tabela 33 – Distribuição da freqüência dos alunos por gênero e opção em relação ao fumo de cigarros

| Gênero | Opção em relação do fumo | | Total |
|---|---|---|---|
| | Fuma | Não Fuma | |
| Homem | 125 | 175 | 300 |
| Mulher | 30 | 70 | 100 |
| Total | 155 | 245 | 400 |

A probabilidade de uma pessoa escolhida ser mulher é $P(A)=\frac{100}{400}=0{,}25$, e a probabilidade de uma pessoa escolhida não fumar é $P(B)=\frac{245}{400}=0{,}6125$.

A partir da observação da tabela é possível dizer que a probabilidade de uma pessoa escolhida ser mulher **e** não fumar, que representa a intersecção dos eventos A e B é dada por $P(A\cap B)=\frac{70}{400}=0{,}175$.

Porém, poderia ser observada a probabilidade de uma pessoa escolhida ser mulher dentre as que não fumam, que restringe o espaço amostral para apenas as pessoas da amostra que não fumam, correspondente a 245. Á esse raciocínio denomina-se probabilidade condicional, pois se refere à probabilidade de acontecer um evento sabendo-se que o outro evento aconteceu.

Para este exemplo, a probabilidade de ser mulher dado que a pessoa sorteada não fuma é $P(A/B)=\frac{70}{245}\cong 0{,}2857$, obtido pela observação da Tabela 33.

Daí, a definição de probabilidade condicional é "dados dois eventos A e B, a probabilidade condicional de A dado que ocorreu B é representada por P(A/B) e dada por $P(A/B) = \frac{P(A \cap B)}{P(B)}$ , com P(B)>0" (MAGALHÃES e LIMA, 2004, P. 42), que permite a obtenção de $P(A \cap B) = P(B) \cdot P(A/B)$

Para o mesmo exemplo anterior, a probabilidade de sortear uma mulher dado que não fuma a partir da utilização da definição de probabilidade condicional é $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{70}{400}}{\frac{245}{400}} = \frac{70}{400} \cdot \frac{400}{245} = \frac{70}{245} \cong 0{,}2857$.

No exemplo fornecido, deve ser observado que a $P(A/B) \neq P(A)$. Daí a definição de independência de eventos: "Dois eventos são independentes, se a informação da ocorrência ou não de B não altera a probabilidade de ocorrência de A, isto é $P(A/B) = P(A)$, P(B)>0" (MAGALHÃES e LIMA, 2004, p. 44).

É possível concluir, então, que o evento A (ser mulher) e o evento B (não fumar) não são independentes pois $P(A/B) \neq P(A)$.

Decorre que, para dois eventos serem independentes $P(A/B) = P(A)$, em que $P(A/B) = \frac{P(A \cap B)}{P(B)}$. Então, $\frac{P(A \cap B)}{P(B)} = P(A) \Rightarrow P(A \cap B) = P(A) \cdot P(B)$, que está tratado neste trabalho como um critério de verificação da independência de dois eventos.

Para dar início à discussão sobre independência de variáveis, a Tabela 33 está reapresentada na Tabela 34 não mais com a freqüência absoluta, mas com a probabilidade.

Tabela 34 – Distribuição de P(X), P(Y) e P(X,Y)

| Gênero (X) | Opção em relação do fumo (Y) | | P(X) |
|---|---|---|---|
| | Fuma (Y=1) | Não Fuma (Y=2) | |
| Homem | $125/400$ | $175/400$ | $300/400$ |
| Mulher | $30/400$ | $70/400$ | $100/400$ |
| P (Y) | $155/400$ | $245/400$ | 1 |

A última coluna da Tabela 34 apresenta a probabilidade de X quando X representa a categoria homem e quando X representa a categoria mulher, que pode ser escrito como $P(X = x_i)$ e a última linha da Tabela 34 apresenta a probabilidade de Y quando Y representa a categoria fumar e não fumar, ou seja, $P(Y = y_j)$. Em outras palavras, a última coluna e a última linha da Tabela 34 representam a distribuição de probabilidade de X e Y, respectivamente, e que são usualmente tratadas por distribuições marginais.

As quatro células centrais da Tabela 34 apresentam a distribuição de probabilidade conjunta, ou seja, $P(X = x_i \cap Y = y_j)$ que também pode ser escrito como $P(X = x_i, Y = y_j)$ que refere-se à probabilidade da variável aleatória bidimensional.

Segundo Meyer ( 2003, p. 121), "seja (X,Y) uma variável aleatória discreta bidimensional. Diremos que X e Y são variáveis aleatórias independentes se, e somente se, $p(x_i, y_j) = p(x_i)q(y_j)$ para quaisquer i e j. Isto é, $P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$, para todo i e j" onde o autor denomina de $q(y_j) = P(Y = y_j)$, ou seja, a existência de um único par ( x, y ) tal que $P(X,Y) \neq P(X) \cdot P(Y)$ é suficiente para concluir que as variáveis X e Y não são independentes. Logo, para o exemplo de gênero e opção pelo fumo, pode-se concluir que as variáveis não são independentes, pois

$P(X = \text{mulher}, Y = \text{não\_fuma}) = P(X = 2, Y = 2) = \frac{70}{400}$ é diferente do produto das marginais, $P(X = 2) \cdot P(Y = 2) = \frac{100}{400} \cdot \frac{245}{400} = \frac{245}{1600}$.

Para exemplificar um caso de independência, é utilizado o Exemplo 3 do Capítulo 2. A Tabela 35 (já enunciada no Capítulo 2) apresenta os valores de duas variáveis X e Y e a Tabela 36 apresenta a distribuição conjunta de X e Y.

Tabela 35 – Distribuição dos valores de X e Y

| Variável | Elemento | | | | | |
|---|---|---|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ | $E_6$ |
| X | 1 | 1 | 2 | 2 | 3 | 3 |
| Y | 5 | 6 | 5 | 6 | 5 | 6 |

Tabela 36 – Distribuição conjunta de X e Y.

| Nota na disciplina Y | Nota na disciplina X | | | Total |
|---|---|---|---|---|
| | 1 | 2 | 3 | |
| 5 | 1/6 | 1/6 | 1/6 | 3/6 |
| 6 | 1/6 | 1/6 | 1/6 | 3/6 |
| Total | 2/6 | 2/6 | 2/6 | 1 |

A partir da distribuição conjunta apresentada na Tabela 36, a Tabela 37 apresenta a verificação da independência de X e Y a partir do critério $P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$ para todo i e j.

Tabela 37 – Verificação de $P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$

| $P(X = x_i, Y = y_j)$ | $P(X = x_i) \cdot P(Y = y_j)$ | Confirma? |
|---|---|---|
| $P(X = 1, Y = 5) = \frac{1}{6}$ | $P(X = 1) \cdot P(Y = 5) = \frac{2}{6} \cdot \frac{3}{6} = \frac{6}{36} = \frac{1}{6}$ | Sim |
| $P(X = 1, Y = 6) = \frac{1}{6}$ | $P(X = 1) \cdot P(Y = 6) = \frac{2}{6} \cdot \frac{3}{6} = \frac{6}{36} = \frac{1}{6}$ | Sim |
| $P(X = 2, Y = 5) = \frac{1}{6}$ | $P(X = 2) \cdot P(Y = 5) = \frac{2}{6} \cdot \frac{3}{6} = \frac{6}{36} = \frac{1}{6}$ | Sim |
| $P(X = 2, Y = 6) = \frac{1}{6}$ | $P(X = 2) \cdot P(Y = 6) = \frac{2}{6} \cdot \frac{3}{6} = \frac{6}{36} = \frac{1}{6}$ | Sim |
| $P(X = 3, Y = 5) = \frac{1}{6}$ | $P(X = 3) \cdot P(Y = 5) = \frac{2}{6} \cdot \frac{3}{6} = \frac{6}{36} = \frac{1}{6}$ | Sim |
| $P(X = 3, Y = 6) = \frac{1}{6}$ | $P(X = 3) \cdot P(Y = 6) = \frac{2}{6} \cdot \frac{3}{6} = \frac{6}{36} = \frac{1}{6}$ | Sim |

Como pode ser observado na Tabela 37, a probabilidade conjunta foi igual ao produto das marginais, para todas as conjuntas,o que permite concluir pela independência de X e Y.

Como foi salientado no Capítulo 2, os outros dois exemplos apresentados (exemplo 1 e exemplo 2 daquele capítulo) não diziam respeito à variáveis independentes e isto pode ser verificado a seguir.

A Tabela 38 já foi apresentada no Capítulo 2 e a Tabela 39 diz respeito à sua distribuição conjunta.

Tabela 38 – Valores de X e Y do exemplo 1 do Capítulo 2.

| Variável | Elemento | | | | |
|---|---|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ |
| X | 0 | 1 | 2 | 5 | 7 |
| Y | 3 | 5 | 7 | 9 | 4 |

Tabela 39 – Distribuição conjunta de X e Y – exemplo 1 do Capítulo 2.

| Nota na disciplina Y | Nota na disciplina X | | | | | Total |
|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 5 | 7 | |
| 3 | 1/5 | 0 | 0 | 0 | 0 | 1/5 |
| 4 | 0 | 0 | 0 | 0 | 1/5 | 1/5 |
| 5 | 0 | 1/5 | 0 | 0 | 0 | 1/5 |
| 7 | 0 | 0 | 1/5 | 0 | 0 | 1/5 |
| 9 | 0 | 0 | 0 | 1/5 | 0 | 1/5 |
| Total | 1/5 | 1/5 | 1/5 | 1/5 | 1/5 | 1 |

Utilizando a verificação da independência, tem-se que $P(X=0, Y=3)=\frac{1}{5}$ é diferente de $P(X=0)\cdot P(Y=3)=\frac{1}{5}\cdot\frac{1}{5}=\frac{1}{25}$, que permite concluir que as variáveis não são independentes.

Utilizando a mesma estratégia para o Exemplo 2 do Capítulo 2, a Tabela 40 apresenta as observações e a Tabela 41 apresenta a distribuição conjunta.

Tabela 40 – Distribuição dos valores de X e Y do exemplo 2 do Capítulo 2.

| Variável | Elemento | | | | |
|---|---|---|---|---|---|
| | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ |
| X | 1 | 2 | 3 | 4 | 5 |
| Y | 2 | 0 | 2 | 0 | 2 |

Tabela 41 – Distribuição conjunta de X e Y – exemplo 2 do Capítulo 2

| Nota na disciplina Y | Nota na disciplina X | | | | | Total |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| 0 | 0 | 1/5 | 0 | 1/5 | 0 | 2/5 |
| 2 | 1/5 | 0 | 1/5 | 0 | 1/5 | 3/5 |
| Total | 1/5 | 1/5 | 1/5 | 1/5 | 1/5 | 1 |

Utilizando a verificação da independência, tem-se que $P(X=1, Y=0)=0$ é diferente de $P(X=1)\cdot P(Y=0)=\frac{1}{5}\cdot\frac{2}{5}=\frac{2}{25}$, que permite concluir que as variáveis X e Y não são independentes.

**Apêndice 2 – Questionário aplicado aos professores participantes**

Este questionário tem o objetivo de conhecer o perfil dos participantes desta pesquisa e, por este motivo, pedimos a gentileza de responder as questões sinceramente.
Por favor, certifique-se que você respondeu todas as questões.

1) Idade :__________anos

2) Você fez um ou mais cursos de graduação?
( ) Um curso de graduação ( ) Mais de um curso de graduação

3) Para cada curso de graduação que você realizou, responda:
Nome do curso: ____________________________________________________________
Universidade onde cursou: _______________________________________________
Ano de conclusão: _________.

4) Você já fez algum curso de pós-graduação? ( ) sim ( ) não

5) Se você já fez algum curso de pós-graduação, responda:
qual curso?_______________________________________________________________
em que Universidade?______________________________________________________

6) Se você já fez algum curso extra-curricular ou já participou de projetos de educação continuada, especifique o curso, o conteúdo abordado e o ano de realização:
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

7) Neste ano de 2005, você está lecionando para que séries?
__________________________________________________________________________
__________________________________________________________________________

8) Nestas séries em que você está lecionando, você estará trabalhando algum conteúdo estatístico? Se sim, especifique o conteúdo para cada série:

______________________________________________________________

______________________________________________________________

______________________________________________________________

9) Se você, em anos anteriores, trabalhou com conteúdos estatísticos, especifique a série e o conteúdo.

______________________________________________________________

______________________________________________________________

10) Você conhece os Parâmetros curriculares nacionais para o Ensino Fundamental?

( ) sim ( ) não

11) Você conhece os Parâmetros curriculares nacionais para o Ensino Médio?

( ) sim ( ) não

12) Você conhece o PCN+?

( ) sim ( ) não

13) Você usa os parâmetros para a construção do seu plano de aula?

( ) sim ( ) não

14) Que tipo de material você costuma usar para preparar sua aula? (consulta de livros, internet, colegas, etc...)

______________________________________________________________

______________________________________________________________

15) Cite metodologias que você usa normalmente em sua aula. Você acredita que essas metodologias seriam também adequadas para tratar temas ligados à estatística? Por quê?

___________________________________________________________________________

___________________________________________________________________________

____________________________________________________________________

16)  Você quer acrescentar alguma informação sobre sua formação?

___________________________________________________________________________

___________________________________________________________________________

17) Você quer acrescentar alguma informação sobre seu trabalho docente?

___________________________________________________________________________

___________________________________________________________________________

18) Você quer acrescentar alguma informação sobre o conteúdo estatístico que você tem a trabalhar ou já trabalhou?

___________________________________________________________________________

__________________________________________________________________________

**Apêndice 3 - Carta de Esclarecimento e o Termo de Consentimento Livre e Esclarecido aplicado aos professores participantes**

**Carta de esclarecimento sobre o Projeto e a Pesquisa**

**PROJETO**: PROCESSO DE ENSINO E APRENDIZAGEM ENVOLVENDO RACIOCÍNIO ESTOCÁSTICO: conceitos estatísticos e probabilísticos na escola básica (PEA)

**Pesquisa:** A NOÇÃO DE VARIABILIDADE: Um estudo com professores de matemática do ensino fundamental e médio.

**Pesquisadora**: Cláudia Borim da Silva
**Orientadora do Projeto e da Pesquisa** : Profa. Dra. Cileda de Queiroz e Silva Coutinho

**Informações sobre o projeto e sobre a pesquisa:**

A pesquisa a ser realizada faz parte da tese de doutorado em desenvolvimento no Programa de Pós Graduação em Educação Matemática da Pontifícia Universidade Católica de São Paulo (PUC-SP).

O objetivo principal da pesquisa é verificar a aplicabilidade de uma seqüência de ensino para conteúdos estatísticos propostos pelos Parâmetros Curriculares Nacionais para o ensino fundamental e médio. Verificar como professores do ensino fundamental e médio apreendem a idéia de variabilidade e analisam a aplicabilidade de uma seqüência didática com seus alunos.

Se você é professor(a) de matemática da Educação Básica, convidamos para participar desta pesquisa.

A pesquisa se realizará com encontros semanais, todas as sextas feiras, no período da manhã, com início previsto para Abril de 2005 e término previsto para Junho de 2005.

Em cada encontro semanal serão discutidos aspectos didáticos e estatísticos propostos para serem trabalhados no ensino fundamental e médio, assim como textos específicos para a reflexão sobre a prática docente (como estamos e onde queremos chegar? Como atingir nossos objetivos? Como melhorar nossas práticas docentes?).

Todos os encontros serão gravados em áudio e eventualmente serão filmados.

Durante todo o desenvolvimento dos trabalhos, teremos também a oportunidade de interagir por um fórum via internet. Com isso pretendemos potencializar nossas discussões e também prestar toda a assistência necessária aos participantes do projeto.

Todos as informações obtidas, em registros escritos, gravados ou filmados, permanecerão em completo sigilo por 5 anos. Assegura-se a não divulgação de nomes dos participantes e nem das instituições a que estão vinculados nos resultados da pesquisa.

**Pesquisa:** A NOÇÃO DE VARIABILIDADE: Um estudo com professores de matemática do ensino fundamental e médio.

**Termo de Consentimento Livre e Esclarecido**

Eu, ________________________________________________________________, com ___ anos de idade, portador (a) do RG ___________________________, residente na _______________________________________________________________, com número de telefone ______________________________ e e-mail ______________________________, abaixo assinado, dou meu consentimento livre e esclarecido para participar como voluntário da pesquisa supra citada, sob a responsabilidade da pesquisadora Cláudia Borim da Silva, aluna do curso de Doutorado em Educação Matemática da PUC-SP e da Professora Dra. Cileda de Queiroz e Silva Coutinho, orientadora da pesquisa e docente do Programa de Doutorado da PUC-SP.

Assinando este Termo de Consentimento, estou ciente de que:

1) O objetivo da pesquisa é desenvolver e aplicar uma seqüência de ensino para conteúdos de Estatística;
2) A realização desta pesquisa é fundamental para a produção de material didático que apoie os professores de matemática no ensino de Estatística na escola básica;
3) Durante o estudo, estarei preenchendo questionários, participando de entrevistas, elaborando material didático sobre o tema Estatística e participando semanalmente da pesquisa supra citada.
4) Assim que for terminada a pesquisa, terei acesso aos resultados globais do estudo;
5) Estou livre para interromper, a qualquer momento, minha participação nesta pesquisa;
6) A participação nesta pesquisa é voluntária, sendo que os professores não receberão qualquer forma de remuneração;
7) Os dados pessoais dos professores serão mantidos em sigilo e os resultados obtidos com a pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, incluindo a publicação na literatura científica especializada;

8) Poderei entrar em contato com os pesquisadores sempre que julgar necessário. Com Cláudia Borim da Silva, no telefone 9970-6498 ou pelo email dasilvm@uol.com.br e com a pesquisadora Dra. Cileda de Queiroz e Silva Coutinho, pelo email cileda@pucsp.br.
9) Obtive todas as informações necessárias para poder decidir conscientemente sobre a minha participação na referida pesquisa;
10) Este Termo de Consentimento é feito em duas vias, de maneira que uma permanecerá em meu poder e a outra com os pesquisadores responsáveis.

São Paulo, ______de ____________________ de 2005.

_______________________________________________

assinatura do participante

_______________________________________________

assinatura da responsável pela pesquisa

_______________________________________________

assinatura da pesquisadora

**Apêndice 4: Síntese do texto Investigações em Estatística ( Ponte, Brocardo e Oliveira, 2003)**

Síntese do Texto: **Investigações em Estatística**

Realizada por **Cláudia Borim da Silva**

**Extraído do livro:** Ponte, João Pedro da, Brocardo, Joana e Oliveira, Hélia. **Investigações Matemáticas na Sala de Aula.** Belo Horizonte: Autêntica: 2003.

"No currículo de Matemática, a Estatística é um tema relativamente recente. As abordagens usuais deste tópico enfatizam os aspectos computacionais e procedimentais: como se calcula a média ou o desvio padrão, como se faz um gráfico de barras, um gráfico circular ou diagrama de caule e folhas. Como conseqüência, a Estatística pode tornar-se um dos temas de Matemática mais aborrecidos de ensinar e de aprender”.

No entanto, este tema matemático desempenha um papel essencial na educação para a cidadania. Na verdade, a Estatística constitui uma importante ferramenta para a realização de projetos e investigações em numerosos domínios, sendo usada no planejamento, na recolha e na análise de dados e na realização de inferências para tomar decisões. A sua linguagem e conceitos são utilizados em cada passo do dia-a-dia para apoiar afirmações em domínios como a saúde, o desporto, a educação, a ciência, a economia e a política. Todo o cidadão precisa saber quando um argumento estatístico está ou não a ser utilizado com propriedade".

Neste capítulo, os autores descrevem uma experiência realizada com alunos de uma turma de 6ª série do ensino fundamental, que foi objeto de estudo de uma dissertação orientada pelo Professor João Pedro da Ponte.

O objetivo da experiência era descobrir o aluno típico. O que foi discutido em cada aula está resumido no Quadro 1:

Quadro 1: Síntese da experiência realizada pela orientanda do Ponte.

| **Aula** | **Objetivo da aula** | **Questionamento das professoras** | **Observações** |
|---|---|---|---|
| 1ª | Preparar as questões de investigação | Suponha que queira comunicar a um aluno de um país distante, ou mesmo, quem sabe, a um extraterrestre, como são os alunos de sua turma.<br>Discussão feita pelos alunos: que dados devem entrar na caracterização do aluno típico? Será necessário traçar um perfil para os moços e outro para as moças? Por quê? | Os alunos perceberam que altura e peso necessitariam de medição |
| 2ª | Orientar os alunos na preparação da coleta de dados | Escrever o que vai pesquisar em forma de pergunta.<br>Como vai obter as respostas (observação, medição, entrevista)? | "Na reflexão feita no final da aula, as professoras concluíram que os alunos são capazes de se organizarem e têm iniciativa quando estão a resolver problemas que lhes interessam" |

| Aula | Objetivo da aula | Questionamento das professoras | Observações |
|---|---|---|---|
| 3ª | Organizar e representar os dados | "1. Qual é o valor mínimo dos seus dados? E o valor máximo? E a distância entre estes dois valores? Acha que os seus dados estão muito concentrados ou estão espalhados?<br>2. Tente descobrir uma forma de organizar os dados de modo que seja fácil ver quantas vezes aparece cada valor.<br>3. Qual é o valor mais frequente? (moda)<br>4. Qual é o valor do meio? (mediana)<br>5. A média de um conjunto de valores obtém-se somando todos os valores e dividindo esta soma pelo número total de dados. Calcule a média dos seus dados. Escreva algumas propriedades da média.<br>6. A moda, a mediana e a média são três medidas estatísticas que pode usar na caracterização de um conjunto de dados. Qual destas medidas, pensa que dá uma melhor idéia acerca do seu conjunto de dados? Por quê?<br>7. Um conjunto de dados pode ser representado de muitas maneiras diferentes: tabelas, diagramas, gráficos, etc. Escolha uma representação para os seus dados que seja diferente da dos seus colegas de grupo. Compare as diferentes representações e escolha aquela que, no teu entender, dá uma melhor visão dos dados. Justifique a sua escolha. | O conceito de moda surgiu mais naturalmente quando os alunos discutiam sobre a cor dos olhos que seria típica da turma e perceberam que era castanho.<br>A dificuldade ficou com a mediana. |
| 4ª | Discutir sobre a aprendizagem | | |

**Apêndice 5: Instrumento de coleta de dados elaborado pelos professores participantes**

**Descrição Metodológica da Pesquisa**

**População:** Professores de todas as disciplinas do Ensino Fundamental e Médio

**Amostra**: Professores de todas as disciplinas do Ensino Fundamental e Médio que lecionem na mesma escola dos professores participantes do Projeto.

**Objetivo Geral:** Identificar a visão dos professores de diferentes áreas de conhecimento sobre a Estatística.

**Objetivos específicos:**
- Qual a opinião dos professores sobre a Estatística?
- O que os professores utilizam de Estatística em suas aulas?
- O que eles esperam do professor de matemática?

**Procedimento:**
Será solicitado ao diretor da escola autorização para a realização da pesquisa no horário de H.T.P.C. Esta autorização será solicitada mediante uma carta, em que também constará o objetivo da pesquisa.

São Paulo, 15 de Abril de 2005.

Caro(a) Diretor(a)

Solicitamos sua autorização para a realização de uma pesquisa com os professores da escola sobre a utilidade da Estatística nas diferentes áreas de conhecimento.

Esta pesquisa tem como objetivo coletar dados que serão ferramenta para a discussão de estratégias de ensino-aprendizagem de conteúdos estatísticos na disciplina Matemática.

O professor de Matemática que estará aplicando os questionários faz parte do Grupo de Pesquisa " O Pensamento Matemático no Ensino Fundamental" coordenado pelo Professor Doutor Saddo Ag Almouloud da Pontifícia Universidade Católica de São Paulo.

Salientamos que os dados permanecerão em sigilo e só serão divulgados os resultados gerais da pesquisa, em que não serão identificados nem o nome dos professores nem o nome da escola.

Após o término da pesquisa, o relatório com os resultados serão enviados para a escola.

Gratos pela atenção

Cláudia Borim da Silva - pesquisadora

Profa. Dra. Cileda de Queiroz e Silva Coutinho - orientadora da pesquisa

Prof. Dr. Saddo Ag Almouloud - coordenador do Projeto

**Questionário**

Caro professor,

Este questionário faz parte de um trabalho que está sendo realizado no Grupo de Pesquisa da PUC-SP intitulado " O Pensamento Matemático no Ensino Fundamental" e tem como objetivo identificar a visão dos professores de diferentes áreas de conhecimento sobre a Estatística e sua utilidade em suas aulas.

Salientamos que todos os dados permanecerão em sigilo e só serão divulgados os resultados gerais da pesquisa. Pedimos a gentileza de responder todas as questões sinceramente.

Muito Obrigado!

1) Sexo: (   ) Feminino (   ) Masculino.
2) Idade: ____anos
3) Quanto tempo você tem de magistério? ____anos
4) Qual sua habilitação específica?

______________________________________________________________

______________________________________________________________

5) Assinale com um "X" em que tipo de escola você leciona:

(   ) escola privada

(   ) escola pública

(   ) escola privada e pública

6) Qual (is) a(s) disciplina(s) que você está lecionando este ano?

______________________________________________________________

______________________________________________________________

7) Em que série(s) você está lecionando este ano?

______________________________________________________________

______________________________________________________________

8) Para você, o que lembra a palavra Estatística? ________________________

9) Atribua uma nota <u>entre</u> zero (nada importante) e dez (extremamente importante) para a importância de Estatística em:
   a) seu cotidiano: _______pontos de importância
   b) sua área de formação: _____pontos de importância
   c) a sua disciplina: _______ pontos de importância
   d) seu aluno: _______ pontos de importância

10) Que conteúdo(s) de Estatística você aborda em suas aulas?

_______________________________________________________________

11) De que forma você usa a Estatística para desenvolver o conteúdo da disciplina que você leciona?

_______________________________________________________________

12) Você incentiva a leitura e discussão de textos que envolvam dados estatísticos? (   ) sim(   ) não

a) Se você respondeu **SIM**, que tipo de texto você costuma usar? ____________
b) Se você respondeu **NÃO**, explique porque você não usa. _________________

13) Qual é a aceitação do seu aluno quando você usa Estatística em suas aulas?

_______________________________________________________________

14) Como você avalia a compreensão (aprendizagem) de seus alunos a respeito do uso da Estatística em suas aulas?

_______________________________________________________________

15) Qual (is) a(s) dificuldade(s) que você encontra para utilizar a Estatística em suas aulas?

_______________________________________________________________

16) Que sugestões você apresenta para enfrentar as dificuldades encontradas com o uso da Estatística em suas aulas?

_______________________________________________________________

## Apêndice 6: Parte do Banco de Dados da pesquisa “Teorema de Pitágoras: uma aprendizagem significativa?”

| Alunos | sexo | idade | escola | disc+gosta | disc-gosta | ... | gostaexer | tipodific | decora | facilidade | nota mat | nota teste |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 0 | 15 | 1 | Português | Matemática | | 1 | 4 | 2 | 0 | 8 | 9,5 |
| **2** | 0 | 14 | 0 | Biologia | Geografia | | 0 | 1 | 2 | 0 | | 8,5 |
| **3** | 0 | 15 | 1 | Matemática | História | | 0 | 1 | 2 | 0 | | 10,0 |
| **4** | 0 | 15 | 1 | Biologia | Física | | 0 | 3 | 2 | 1 | 7,5 | 5,5 |
| **5** | 1 | 15 | 1 | Geografia | Matemática | | 0 | 1 | 0 | 1 | 6 | 4,0 |
| **6** | 0 | 14 | 1 | História | Matemática | | 1 | 1 | 0 | 1 | 7 | 6,0 |
| **7** | 0 | 14 | 1 | História | Matemática | | 1 | 3 | 2 | 1 | 8 | 9,0 |
| **8** | 0 | 16 | 1 | Português | Matemática | | 1 | 0 | 0 | 1 | | 2,0 |
| **9** | 0 | 15 | 1 | Biologia | Física | | 1 | 4 | 2 | 0 | 7 | 6,5 |
| **10** | 0 | 14 | 1 | Português | Matemática | | 1 | 0 | 2 | 1 | 6 | 5,0 |
| **...** | | | | | | | | | | | | |
| **...** | | | | | | | | | | | | |
| **55** | 1 | 15 | 1 | Física | Português | | 0 | 3 | 2 | 0 | 7,5 | 5,0 |
| **56** | 1 | 14 | 1 | História | Física | | 0 | 3 | 2 | 1 | 8,5 | 8,5 |
| **57** | 0 | 14 | 1 | Biologia | Matemática | | 0 | 0 | 2 | 0 | 7 | 8,5 |

| sexo | escola | gostaexerc | tipodific | decora | facilidade |
|---|---|---|---|---|---|
| 0 = feminino | 0 = escola pública | 0 = sim | 0 = visualizar figuras | 0 = sim | 0 = sim |
| 1= masculino | 1 = escola particular | 1 = não | 1 = associar nomes e figuras | 1 = não | 1 = não |
| | | | 2 = utilizar materiais | 2 = às vezes | |
| | | | 3 = fórmulas | | |
| | | | 4 = não tem dificuldades | | |

**Apêndice 7: Banco de dados da pesquisa realizada pelos professores participantes**

| Questionario | escola | sexo | idade | tempo de magistério | habilitação_1 | habilitação 2 | habilitação 3 | tipo escola | leciona_1 | leciona_2 | serie | palavra Estatistica | importancia cotidiano | importancia area | importancia disciplina | importancia aluno | conteudo | forma | incentiva | sim, qual texto | não, porque | aceitação do aluno | compreensão | dificuldades | sugestoes |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 35 | 14 | 1 | | | 2 | 1 | | 5 | 1 | 8 | 10 | 10 | 8 | 5 | 1 | 1 | 1 | | 1 | 2 | 10 | 6 e 7 |
| 2 | 1 | 2 | 39 | 5 | 1 | | | 2 | 1 | | 2 | 1 | 10 | 10 | 10 | 10 | 3 e 5 | 1 | 2 | | 2 | 1 | 9 | 6 | 4 |
| 3 | 1 | 1 | 26 | 5 | 2 | | | 3 | 2 | 3 | 5 e 7 | 2 | 5 | 5 | 5 | 8 | 3 | 1 | 1 | 5 | | 1 | 2 | 4 | 2 e 6 |
| 4 | 1 | 1 | 28 | 7 | 2 | | | 3 | 2 | 3 | 5 | 3 | 3 | 3 | 3 | 3 | 5 | 1 e 2 | 1 | 1 | | 5 | 6 | 4 e 5 | |
| 5 | 1 | 1 | 44 | 10 | 2 | | | 2 | 3 | | 5 | 5 | 6 | 6 | 6 | 10 | 4 | 2 | 1 | 3 | | 2 | 9 | | |
| 6 | 1 | 1 | 47 | 17 | 3 | | | 2 | 4 | | 2 | 1 e 3 | 7 | 8 | 7 | 8 | 3 e 4 | 1 | 1 | 3 | | 1 | 1 e 4 e 6 | 2 | 2 e 6 e 7 |
| 7 | 1 | 1 | 25 | 8 | 4 | | | 2 | 4 | 5 | 5 | 4 | 9 | 10 | 10 | 8 | 5 | 1 e 2 | 1 | 1 | | 6 | 4 e 7 e 9 | 4 | |
| 8 | 1 | 2 | 24 | 1 | 5 | | | 2 | 6 | | 2 | 1 | 8 | 8 | 8 | 8 | 2 e 5 | 1 | 1 | 4 | | 5e6e7 | 2 | 4 | 4 e 6 |
| 9 | 1 | 1 | 51 | 28 | 5 | 6 | | 2 | 6 | | 5 | 2 | 6 | 6 | 6 | 10 | 4 | 5 | 2 | 2 | 1 | 2 | 9 | | |
| 10 | 1 | 2 | 33 | 8 | 7 | | | 2 | 7 | 8 | 2 | 1 e 3 | 7 | 7 | 7 | 6 | 1 | 1 | 1 | | | 3 | 4 | 6 | 5 |
| 11 | 1 | 1 | 38 | 18 | 8 | | | 2 | 8 | | 5 | 2 | 5 | 8 | 5 | 5 | 3 | 1 | 1 | 3 | | 5 | 6 | 4 e 7 | 3 e 4 e 6 |
| 12 | 1 | 1 | 41 | 22 | 8 | | | 2 | 7 | 8 | 5 | 1 | 10 | 10 | 10 | 10 | 2 e 5 | 1 | 1 | 1 e 2 | | 1 | 1 | 8 | 3 e 5 e 7 |
| 13 | 2 | 2 | 29 | 10 | 4 | | | 3 | 4 | 9 | 5 | 1 | 8 | 10 | 5 | 5 | 3 e 4 | 1 | 1 | 1 | | 5 | 6 | 2 e 6 | 2 |
| 14 | 2 | 1 | 25 | 4 | 9 | | | 2 | 10 | | 2 | 4 | 2 | 5 | 5 | 5 | 2 e 4 | 5 | 2 | | 4 | 5 | 1 | 2 | 7 |
| 15 | 2 | 2 | 34 | 12 | 7 | | | 3 | 7 | | 5 e 8 | 1 | 10 | 10 | 8 | 8 | 2, 3 e 4 | 4 | 1 | 1 | | 1 | 1 e 4 e 9 | 6 e 7 | 4 |
| 16 | 2 | 2 | 31 | 2 | 1 | | | 2 | 1 | | 5 | 1 | 7 | 10 | 10 | 10 | 5 | 4 | 1 | 1 | | 3 | 5 | 7 | 7 |
| 17 | 2 | 1 | 36 | 14 | 8 | | | 2 | 8 | | 5 | 3 | 10 | 10 | 10 | 10 | 2 e 5 | 1 | 1 | 1 | | 3 | 6 | 7 | 2 |
| 18 | 2 | 2 | 23 | 3 | 1 | | | 2 | 1 | 11 | 5 | 1 | 10 | 10 | 10 | 10 | 5 | 5 | 1 | 5 | | 1e4e6 | 6 | 2 e 6 | 9 |
| 19 | 2 | 1 | 24 | 5 | 2 | 10 | | 2 | 3 | 13 | 2 | 2 | 5 | 5 | 5 | 7 | | 1 | 1 | 3 | | 5 | 6 | 6 | |
| 20 | 2 | 1 | 25 | 5 | 1 | 11 | | 2 | 12 | | | 2 | 6 | 8 | 10 | 5 | 5 | 5 | 1 | 2 | | 5 | 1 | 2 e 11 | 5 e 9 |
| 21 | 2 | 1 | 29 | 10 | 3 | 4 | | 3 | 4 | 9 | 5 | 4 | 10 | 9 | 10 | 5 | 2 e 5 | 4 | 2 | | 2 | 4 | 6 | 9 | 2 |
| 22 | 2 | 1 | 37 | 17 | 2 | | | 2 | 2 | | 5 | 3 | | | | | 1 | 3 | 2 | | | | | | |
| 23 | 2 | 1 | 29 | 10 | 2 | | | 2 | 3 | | 5 e 7 | 1 e 3 | 7 | 5 | 7 | 7 | 2 e 3 | 2 | 1 | 1 | | 2 | 2 | 4 e 6 | 2 |
| 24 | 2 | 2 | 41 | 15 | 5 | | | 2 | 6 | | 5 | 1 | 0 | 0 | 0 | 0 | 1 | 3 | 2 | | 4 | 3 | 7 | 10 | 8 |
| 25 | 2 | 1 | 40 | 13 | 2 | | | 2 | 2 | 3 | 5 | 1 | 10 | 10 | 9 | 10 | | 2 | 1 | 5 | | 4 | 9 | 7 | 7 |
| 26 | 2 | 2 | 60 | 14 | 12 | | | 2 | | | 5 | 1 | 4 | 10 | 6 | 6 | 5 | 2 | 2 | | 2 | 7 | 4 e 7 e 9 | 12 | 9 |
| 27 | 2 | 1 | 23 | 4 | 1 | | | 2 | 1 | 13 | 2 | 3 | 8 | 8 | 10 | 8 | 3 e 4 | 5 | 1 | 5 | | 1 | 2 | 1 | |
| 28 | 2 | 2 | 34 | 10 | 7 | | | 2 | 7 | 8 | 5 | 2 | 9 | 8 | 10 | 10 | 2 e 4 | 1 | 1 | 5 | | 1 | 1 | 3 e 6 | 1 |
| 29 | 3 | 1 | 21 | 2 | 1 | 13 | | 1 | 14 | | 6 | 5 | 10 | 10 | 10 | 10 | 5 | 6 | 1 | 1 | | 1 | | | |
| 30 | 3 | 2 | 41 | 5 | 1 | | | 3 | 11 | 15 | 3 | 1 e 5 | 10 | 10 | 5 | 5 | 1 | 5 | 2 | | 3 | 1 | 1 | 6 | 7 |
| 31 | 3 | 1 | 42 | 20 | 10 | 13 | | 1 | 13 | | 1 | 2 | 10 | 10 | 10 | 10 | 5 | 5 | 1 | 2 | | 1 | | 1 | 6 e 7 |
| 32 | 3 | 1 | 39 | 11 | 8 | | | 3 | 7 | 8 | 5 | 1 | 5 | 8 | 8 | 8 | 2 e 5 | 1 | 1 | 3 | | 1 | 2 | 1 | |
| 33 | 3 | 1 | 36 | 17 | 4 | | | 1 | 4 | | 2 | 1 | 7 | 8 | 7 | 8 | 2 e 5 | 1 | 1 | 3 | | 1 | 1 | 1 | |
| 34 | 3 | 1 | 34 | 14 | 1 | | | 3 | 1 | | 2 | 1 | 10 | 9 | 10 | 10 | 2 e 5 | 2 | 2 | | | 1 | 9 | 7 | 7 |
| 35 | 4 | 1 | 26 | 4 | 2 | | | 2 | 3 | | 2 e 7 | 1 | 10 | 10 | 10 | 10 | 5 | 2 | 1 | 1 | 1 | 1 | | 4 e 5 | 4 e 9 |
| 36 | 4 | 2 | 50 | 18 | 12 | | | 2 | 8 | | 2 e 7 | 1 | 6 | 8 | 8 | 8 | 2 e 5 | 1 e 2 | 1 | 3 | | 1 | 2 | 4 | 4 |
| 37 | 4 | 1 | 39 | 15 | 3 | 4 | 11 | 3 | 1 | 4 | 2 | 2 | 10 | 10 | 10 | 10 | 3 | 1 | 1 | 1 | | 3 | 2 | 6 | 7 |
| 38 | 4 | 1 | 42 | 15 | 7 | | | 2 | 7 | 8 | 5 | 1 | 4 | 5 | 5 | 0 | 3 | 4 | 2 | | 4 | 3 | 3 | 3 e 11 | 8 |
| 39 | 4 | 2 | 42 | 15 | 7 | | | 2 | 7 | | 5 | 1 | 7 | 5 | 5 | 3 | 2 e 4 | 2 | | 5 | | 3 e 7 | 1 | 1 | 9 |
| 40 | 4 | 1 | 19 | 15 | 1 | | | 2 | 1 | | 3 | 1 e 5 | 8 | 10 | 8 | | 5 | 4 | 1 | 5 | | 2 | 2 e 3 e 7 | 5 | 9 |
| 41 | 4 | 1 | 41 | 10 | 8 | | | 2 | 8 | | 5 e 7 | 1 e 5 | 8 | 10 | 10 | 8 | 2 e 5 | 1 | 1 | 1 | | 1 | 6 | 1 | 8 |
| 42 | 4 | 1 | 36 | 13 | 1 | | | 2 | 1 | | 2 | 1 | 8 | 10 | 10 | 8 | 3 | 1 e 4 | 1 | 1 e 3 | | 1 | 1 | 3 e 6 | 4 e 6 e 9 |
| 43 | 4 | 1 | 28 | 10 | 9 | | | 2 | 10 | | 5 | 4 | 6 | 8 | 5 | 1 | 1 | 1 | 2 | | | 3 | 3 | 3 e 7 | 8 |
| 44 | 4 | 2 | 34 | 8 | 1 | | | 2 | 1 | | 5 | 1 | 7 | 10 | 10 | 7 | 5 | 5 | 2 | | 4 e 5 | 5 | 9 | 4 e 5 | 3 e 9 |
| 45 | 4 | 1 | | 19 | 1 | | | 2 | 1 | | 1 | | 6 | 10 | 10 | 6 | 1 | 3 | 1 | 1e3e5 | | 7 | 7 | 3 e 11 | 6 e 9 |
| 46 | 4 | 1 | 41 | 17 | 4 | | | 2 | 4 | 9 | 5 | 1 | 5 | 6 | 6 | 5 | 4 | | 1 | 1 e 3 | | 2 e 6 e 7 | 9 | 1 | |
| 47 | 4 | 1 | 39 | 19 | 6 | | | 2 | 6 | | 5 | 2 | 0 | 10 | 10 | 10 | 1 | | 2 | | | | | | |
| 48 | 4 | 1 | 42 | 20 | 2 | | | 2 | 3 | 16 | 5 | 2 e 4 | | | | | | | 2 | | | | | | |
| 49 | 4 | 1 | 50 | 6 | 8 | | | 2 | 8 | | 5 | 1 | 5 | 9 | 9 | 5 | 2 e 5 | 1 | 1 | 2 | | 6 | 1 e 2 e 7 | | |
| 50 | 4 | 1 | 53 | 27 | 2 | | | 2 | 2 | 3 | 5 | 1 | 8 | 7 | 8 | 8 | 3 | 2 | 1 | 5 | | 6 | 4 | 3 e 6 | 8 e 9 |
| 51 | 4 | 1 | 42 | 15 | 7 | | | 2 | 7 | | 5 e 7 | 1 | 8 | 8 | 10 | 10 | 2 e 5 | 1 | 1 | 3 | | 4 | 9 | 3 | 6 |
| 52 | 4 | 2 | 51 | 16 | 14 | | | 2 | 7 | 17 | 5 | 1 e 2 | 8 | 6 | 6 | 10 | 5 | 2 | 2 | | 1e4e5 | 1 | 1 e 2 e 9 | 1 | |
| 53 | 5 | 1 | 35 | 15 | 4 | 15 | | 3 | 9 | | 3 | 1 e 3 | 8 | 10 | 8 | 5 | 2 e 5 | | 1 | 1 | | 3 | 2 | 4 | 9 |
| 54 | 5 | 1 | 58 | 21 | 2 | | | 2 | 3 | 16 | 3 | 1 | 4 | 10 | 10 | 10 | 2 e 5 | 6 | 2 | | 1 | | | | 1 e 2 |
| 55 | 5 | 1 | 44 | 3 | 2 | 12 | | 2 | 3 | | 5 | 1 | 7 | 7 | 8 | 10 | 5 | 6 | 1 | 3 | | 1 | 1 | | 1 |
| 56 | 5 | 1 | 34 | 8 | 2 | | | 2 | 3 | | 2 | 1 | 9,5 | 9 | 9 | 8 | 2 e 5 | 1 | 2 | | 5 | 1 | 7 | 2 e 3 | 7 |
| 57 | 5 | 1 | 51 | 20 | 2 | 12 | | 2 | 3 | | 5 | 4 | 10 | 10 | 10 | 10 | 5 | 4 | 1 | 5 | | 1 | 7 | 1 | |
| 58 | 5 | 1 | 34 | 10 | 11 | | | 3 | 1 | 11 | 5 | 1 | 5 | 5 | 3 | 3 | 1 | 2 | 1 | 5 | | 1 | 1 e 7 | 3 | 5 |
| 59 | 5 | 2 | 39 | 15 | 1 | | | 2 | 1 | | 3 e 7 | 3 | 9 | 9 | 7 | 7 | 3 | 1 | 1 | 1 | | 1 | 1 | 1 e 3 | 4 e 6 |
| 60 | 5 | 1 | 36 | 12 | 1 | 3 | | 2 | 1 | | 2 | 1 | 7 | 8 | 8 | 10 | 5 | 4 | 1 | 1 | | 1 | 1 | 1 e 3 | 7 |
| 61 | 5 | 1 | 49 | 25 | 7 | | | 2 | 7 | | 3 | 1 | 8 | 8 | 8 | 8 | 3 | 1 | 2 | 3 | | 3 e 5 e 7 | 9 | 2 e 3 | 2 |
| 62 | 5 | 1 | 44 | 10 | 7 | 8 | | 2 | 7 | | 2 | 3 | 7 | 6 | 5 | 7 | 2 e 5 | 6 | 1 | 5 | | 1 e 2 e 7 | 9 | 1 | 9 |
| 63 | 5 | 2 | 27 | 4 | 14 | | | 2 | 17 | | 3 | 1 e 2 | 5 | 5 | 5 | 5 | 1 | 3 | 2 | | 5 | | | | 9 |
| 64 | 5 | 2 | 46 | 5 | 15 | | | 2 | 5 | 18 | 6 | 1 | 10 | 10 | 5 | 8 | 1 | 3 | 2 | | 4 | 7 | 1 | 3 | 8 |
| 65 | 6 | 1 | 47 | | 2 | 12 | | 2 | 3 | | 5 | 2 | 8 | 7 | 7 | 6 | 1 | 2 | 1 | 2 e 5 | | | | 3 e 11 | |
| 66 | 6 | 2 | 21 | 2 | 2 | 12 | | 2 | 3 | | 5 | 2 e 3 | 7 | 7 | 7 | 7 | 5 | | 1 | 1 | | 1 | 1 | 3 | 5 |
| 67 | 6 | 1 | 25 | 5 | 2 | 12 | | 2 | 3 | | 5 | 2 | 7 | 7 | 7 | 7 | | | 1 | 1 | | 1 | 1 | 3 | |
| 68 | 6 | 1 | 44 | 9 | 3 | 2 | | 2 | 2 | | 3 | 1 | 5 | 8 | 8 | 8 | 1 | | 2 | | 3 | | | | |
| 69 | 6 | 2 | 32 | 6 | 1 | | | 2 | 1 | 11 | 5 | 1 | 8 | 8 | 8 | 7 | 3 e 4 | 1 | 1 | 5 | | 1 | 1 | 5 e 13 | 4 e 7 |
| 70 | 6 | 1 | 51 | 21 | 7 | | | 2 | 7 | | 2 | 1, 2 e 3 | | | | | 5 | | | | | | | | |
| 71 | 6 | 1 | 48 | 6 | 3 | 4 | | 2 | 4 | 9 | 5 | | 9 | 9 | 9 | 8 | | 1 | 1 | 5 | | 1 | 9 | 4 | 3 e 4 |
| 72 | 6 | 1 | 36 | 5 | 1 | | | 2 | 1 | 4 | 5 | 1 | 6 | 8 | 10 | 7 | 3 e 4 | 1 | 1 | 3 | | 4 | 2 | 10 | 9 |
| 73 | 6 | 1 | 38 | 8 | 1 | | | 2 | 1 | 4 | 5 | 1 | 10 | 10 | 10 | 10 | 3 e 4 | 4 | 1 | 5 | | 1 | 1 | 6 e 10 | 2 e 9 |
| 74 | 6 | 2 | 22 | 2 | 2 | | | 2 | 3 | | 3 | 1 | 8 | 8 | 7 | 7 | 5 | 6 | 1 | 1 | | 1 | 8 | 4 | 2 e 9 |
| 75 | 6 | 2 | 26 | 6 | 8 | | | 3 | 8 | | 3 | 1 e 2 | 4 | 5 | 5 | 5 | 5 | 4 | 1 | 3 | | 4 e 6 e 7 | 3 e 7 | 7 e 14 | 9 |
| 76 | 6 | 1 | 42 | 10 | 1 | | | 2 | 1 | | 3 | 1 | 8 | 10 | 10 | 6 | 3 e 4 | 1 | 1 | 3 | | 4 | 1 | 4 | 3 |
| 77 | 6 | 2 | 27 | 5 | 1 | | | 2 | 1 | 11 | 5 | 1 | 10 | 10 | 10 | 10 | 4 | 1 | 1 | 1 | | 1 | 1 | 4 | 7 |
| 78 | 6 | 1 | 45 | 17 | 2 | | | 3 | 3 | | 5 | 3 | 5 | 5 | | | 1 | 3 | 2 | | | | | | |
| 79 | 7 | 1 | 45 | 24 | 7 | 8 | 10 e 16 | 2 | 8 | 19 | 5 | 3 | 6 | 7 | 7 | 5 | 2 e 5 | 1 | 1 | 1 | | 5 | 6 | 7 | 7 e 9 |
| 80 | 7 | 1 | 52 | 19 | 2 | 12 | | 2 | 2 | 3 | 5 | 2 | 8 | 8 | 8 | 8 | 3 | 2 | 1 | 5 | | 1 | 9 | 7 | 7 e 9 |
| 81 | 7 | 1 | 41 | 19 | 4 | | | 2 | 9 | | 3 | 1 | 8 | 10 | 10 | 10 | 1 | 3 | 2 | | 5 | 5 | 1 e 7 | 2 e 6 e 10 | 9 |
| 82 | 7 | 1 | 33 | 8 | 8 | 13 | | 2 | 8 | | 5 | 1 e 2 | 5 | 8 | 8 | 8 | 2 e 3 | 1 | 1 | 1 | | 6 e 7 | | 7 | 8 e 9 |
| 83 | 7 | 1 | 40 | 16 | 7 | | | 2 | 7 | | 5 | 1 | 10 | 10 | 10 | 10 | 2 e 5 | 4 | 1 | 1e2e5 | | 3 | 7 | 6 e 7 | 5 |
| 84 | 7 | 1 | 31 | 10 | 7 | | | 2 | 7 | 19 | 5 | 3 | 7 | 7 | 7 | 6 | 1 | 3 | 2 | 5 | | 3 | 7 | 3 | 1 |
| 85 | 7 | 2 | 54 | 27 | 1 | | | 3 | 1 | | 2 | 1 | 8 | 9 | 7,5 | 6 | 3 | 4 | 1 | 3 | | 1 | 7 | 8 | 9 |
| 86 | 7 | 1 | | 2 | 15 | | | 2 | 5 | | 3 | 1 | 5 | 5 | 5 | 5 | 5 | 1 | 1 | 1 | | 2 | 7 | 3 | 8 |
| 87 | 7 | 1 | 38 | 9 | 7 | 8 | | 3 | 8 | | 5 | | | | | | | | | | | | | | |
| 88 | 7 | 1 | 60 | 14 | 1 | 4 | | 2 | 1 | 9 | 5 e 7 | 3 | 10 | 10 | 10 | 9 | 5 | 1 | 1 | 2 | | 6 | 2 e 6 | 10 | 9 |
| 89 | 7 | 1 | 56 | 23 | 4 | | | 2 | 4 | | 2 | 2 | 5 | 5 | 8 | 8 | 4 | 1 e 2 | 1 | 1 | | 1 | 3 | 6 e 10 | 1 e 9 |
| 90 | 7 | 2 | 45 | 20 | 1 | | | 3 | 1 | 11 | 5 | 1 | 10 | 10 | 10 | 5 | 3 e 4 | 2 | 1 | 1 | | 1 | 1 | 3 e 10 | 1 e 2 e 9 |
| 91 | 7 | 1 | 38 | 15 | 2 | 12 | | 2 | 2 | 3 | 5 | 1 | 8 | 10 | 10 | 10 | 2 e 5 | 2 | 1 | 1 | | 1 | 1 | 5 e 10 | 7 |
| 92 | 7 | 1 | 44 | 10 | 2 | | | 2 | 3 | 16 | 5 | 1 | 8 | 7 | 7 | 5 | 5 | 2 | 1 | 1e3e5 | | 1 | 1 | 4 | 1 |
| 93 | 7 | 1 | 53 | 10 | 12 | | | 2 | 2 | | 5 e 7 | 1, 4 e 5 | 5 | 4 | 5 | 5 | 1 | 1 e 2 | 1 | 1e2e5 | | 6 e 7 | 1 e 7 | 1 e 3 | |
| 94 | 7 | 1 | 51 | | 2 | 12 | | 2 | 3 | | 3 | 1 e 2 | 5 | 5 | 5 | 10 | | | 1 | 5 | | 1 | | | |
| 95 | 8 | 1 | 32 | 14 | 4 | | | 2 | 9 | 4 | 5 | | 8 | 8 | 8 | 8 | | | 1 | | | | | | |
| 96 | 8 | 2 | 24 | 1 | 1 | | | 2 | 13 | | 6 | | 8 | 10 | 10 | 7 | | | 1 | | | | | | |
| 97 | 8 | 1 | 27 | 5 | 14 | | | 2 | 17 | | 3 | | 2 | 5 | 3 | 8 | | | 2 | | | | | | |
| 98 | 8 | 2 | 39 | 8 | 1 | | | 2 | 1 | 11 | 5 | | 7 | 6 | 7 | 8 | | | 1 | | | | | | |
| 99 | 8 | 2 | 56 | 30 | 7 | 8 | | 2 | 8 | | 5 | | 8 | 10 | 10 | 6 | | | 1 | | | | | | |
| 100 | 8 | 1 | 47 | 21 | 2 | | | 2 | 2 | 3 | 5 | | | | | | | | 2 | | | | | | |
| 101 | 8 | 2 | 37 | 14 | 7 | | | 3 | 7 | | 5 | | 5 | 5 | 5 | 10 | | | 2 | | | | | | |
| 102 | 8 | 2 | 47 | 15 | 8 | 16 | | 2 | 8 | | 2 | | 10 | 10 | 10 | 5 | | | 1 | | | | | | |
| 103 | 8 | 2 | 39 | 17 | 7 | | | 2 | 7 | | 2 | | | | | | | | | | | | | | |
| 104 | 9 | 1 | 41 | 20 | 10 | | | 1 | 13 | | 1 | | 10 | 10 | 10 | 10 | | 6 | 1 | 1 e 5 | | 1 | 9 | 1 | 4 |
| 105 | 9 | 1 | 36 | 18 | 7 | 8 | | 3 | 7 | 8 | 5 | | 7 | 9 | 8 | 6 | | 1 e 2 | 1 | 1 | | 1 | 1 | 3 e 14 | 7 e 9 |
| 106 | 9 | 1 | 34 | 15 | 2 | 13 | | 1 | 13 | 2 e 3 | 1 e 7 | | 5 | 10 | 9 | 10 | | 3 | 2 | | 5 | 7 | 7 | 10 | 8 |
| 107 | 9 | 2 | 55 | 11 | 4 | | | 3 | 1 | 4 e 9 | 5 | | 6 | 10 | 8 | 6 | | 1 | 1 | | | 3 | 3 | 3 e 10 | 7 |
| 108 | 9 | 2 | 30 | 10 | 7 | | | 1 | 7 | 8 | 5 | | 9 | 9 | 9 | 9 | | 4 e 6 | 1 | 1 | | 7 | 7 e 9 | 2 e 7 | 7 |
| 109 | 9 | 1 | 37 | 5 | 2 | | | 3 | 2 | 3 | 5 | | 7 | 8 | 8 | 9 | | 2 e 6 | 1 | 1 | | 1 | 9 | 1 | |
| 110 | 9 | 1 | 54 | 30 | 2 | 13 | | 1 | 16 | 20 | 3 | | 4 | 5 | 4 | 8 | | 6 | 1 | 5 | | 7 | | | |