SERGIO ALVES PEREIRA

UM ESTUDO A RESPEITO DO PROFESSOR DE MATEMÁTICA E A IMPLEMENTAÇÃO DE UMA SEQÜÊNCIA DIDÁTICA PARA A ABORDAGEM DA ESTATÍSTICA NO ENSINO MÉDIO

MESTRADO PROFISSIONAL EM ENSINO DE MATEMÁTICA

PUC
SÃO PAULO
2007

SERGIO ALVES PEREIRA

UM ESTUDO A RESPEITO DO PROFESSOR DE MATEMÁTICA E A IMPLEMENTAÇÃO DE UMA SEQÜÊNCIA DIDÁTICA PARA A ABORDAGEM DA ESTATÍSTICA NO ENSINO MÉDIO

Dissertação apresentada à banca Examinadora da Pontifícia Universidade Católica de São Paulo, como Exigência para QUALIFICAÇÃO NO MESTRADO (PROFISSIONAL) EM ENSINO DE MATEMÁTICA. sob orientação da Profª Drª Cileda de Queiroz e Silva Coutinho

PUC/SP

São Paulo

2007

BANCA EXAMINADORA

AGRADECIMENTOS

A todos que de alguma maneira contribuíram para a realização deste trabalho.

<u>Em especial:</u>

- ✓ À Professora Doutora Cileda, orientadora e amiga pelo incentivo, paciência e dedicação;
- ✓ À minha Noiva Cintia, pela força e compreensão nos momentos em que estive ausente;
- ✓ Aos amigos que fiz na pós-graduação pelas discussões e contribuições ao trabalho;
- ✓ Aos amigos Ricardo Sergio B. Vasques ; Ricardo Cardoso, Pedro Bigattão e Jonas Borsetti, pelo companheirismo e amizade verdadeira que me presentearam.
- ✓ Ao professor Doutor Saddo Ag Almouloud pelas valorosas contribuições ao trabalho. e também como professor na formação para o mestrado;
- ✓ Ao amigo Clemente Ramos, por todo o incentivo que me ofereceu e principalmente pela amizade sincera.
- ✓ À minha família, por sempre estar me apoiando e incentivando-me a seguir a diante.
- ✓ À professora Doutora Maria José Ferreira da Silva, por ter me incentivado para fazer o Mestrado, e pelas grandes contribuições que trouxe ao trabalho.
- ✓ Á professora Doutora Irene Mauricio Cazorla, por sua grande contribuição a este trabalho.

RESUMO

Os Parâmetros Curriculares Nacionais (1997) incluíram conteúdos de estatística no ensino Fundamental e Médio como parte do programa disciplinar de Matemática, que foi muito importante para que houvesse mudanças no ensino da Estatística, o fato levou as editoras a uma preocupação maior quanto ao ensino deste tópico, assim, passaram incluir a Estatística de uma forma um tanto mais elaborada em seus manuais e o professor de matemática vem refletindo sobre sua prática. Mas as pesquisas mostram que as universidades nos dias de hoje vêm trabalhando de forma muito limitada no ensino da estatística, fazendo que saiam para as salas de aulas profissionais pouco preparados a respeito desse tópico e como trabalhar, pois a maioria dos professores de matemática cristalizaram em sua prática conhecem muito pouco sobre estatística e as propostas dos PCN para o ensino deste tópico. A proposta deste trabalho aplica-se a uma investigação que envolverá professor e alunos do ensino médio. Verificar-se este educador ensinaria seus alunos a resolver uma lista com problemas elaborados para compreender média; mediana; moda; desvio padrão; quartis e gráficos como procederia depois de alguns ajustes proposto quanto ao Técnico, Mobilizável e Disponível. Após a preparação do professor e os resultados deste preparo ter sido implementado em sua sala de aula, como os alunos responderiam a lista proposta ao teste, e em qual nível de conceitualização proposto por Robert (1998), estes alunos aferiram.

Palavras-chave: Estatística; Professor; Aluno, Ensino-aprendizagem.

ABSTRACT

The National Curricular Parameters (1997) had included contents of statistics in Basic and Average education as part of the program of Mathematics, that was very important so that it had changes in the education of the Statistics, the fact took the publishing companies to a bigger concern about the education of this topic, thus, they had passed to include the Statistics of a form in such a way elaborated in its manuals more and the mathematics teachers comes reflecting on its practical. But the researches show that the universities nowadays come working of a very limited form in the education of the statistics, making that they leave classrooms as professionals with low preparation about this topic and how to work, therefore the majority of the mathematics professors had crystallized in their practical knows little on statistics and the proposals of the PCN for the education of this topic. The proposal of this work is applied in an inquiry that will involve professor and pupils of average education. Verify if this educator would teach its pupils to solve a list with elaborated problems to understand average; medium; mode; standard deviation; quartiles and graphs as it would proceed after some adjustments considered about the Technician, Mobilized and Available. After the preparation of the professor and the results of this preparation have been implemented in its classroom, as the pupils would answer the list proposal to the test, and in which level of conceitualization considered for Robert (1998), these pupils had surveyed.

KEYWORDS - Statistics; Teacher; Pupil, Teach-learning.

SUMÁRIO

LISTA DE GRÁFICOS

LISTA DE FIGURAS

LISTA DE TABELAS

INTRODUÇÃO

Dentre os vários saberes da docência, os saberes da experiência são para os professores o saber de referência. Nessa medida, pergunta-se: por que não oferecer ao professor espaço para a troca entre pares, submetendo os saberes da sua experiência vivida ao reconhecimento por parte dos outros grupos produtores de saberes, impondo-se, desse modo, como produtor de um saber originado em sua prática e sobre o qual poderia reivindicar controle legítimo?

Penim (1995) assumindo que o processo de constituição do conhecimento pelo professor acontece no interior do espaço de representação em que vive e que tal espaço é constituído pelas concepções que vai acumulando sobre o ensino e pela vivência, afirma que os movimentos entre conhecimento sistematizado, o saber cotidiano e a vivência na construção do conhecimento ocorrem no movimento lógico-dialético entre representação/conhecimento, assim, constituindo o campo de possibilidades de processo de formação continuada.

Observamos que os avanços tecnológicos vêm modificando as relações de trabalho e os perfis de competência profissional. O ponto de partida desta pesquisa fundamenta-se na importância que a Estatística vem sendo tratada, por ser um dos ramos da matemática cada vez mais vem conquistando um lugar de destaque na educação.

Os Parâmetros Curriculares Nacionais (PCN) para o Ensino Médio propõem o bloco de análise de dados. Por sua vez, os educadores matemáticos também têm se dedicado à investigação do desenvolvimento de um pensamento estatístico e probabilístico, pelo fato de que fazem parte do mundo atual em que a todo momento os indivíduos precisam entender as informações transmitidas pelos meios de comunicação, no qual, centrados na Estatística e nas probabilidades (SANTOS, 2005)

A Estatística exerce um papel fundamental na formação do sujeito, pois possibilita que lide com a aleatoridade e a variabilidade, permitindo uma análise dos fatos complexos que, sob uma visão determinista, tornam-se impossíveis de ser tratadas, conforme assinala Lopes (1998).

A respeito da aprendizagem dos conceitos estatísticos, Batanero (2001) cita que muitos professores necessitam incrementar seu conhecimento não só quanto ao conteúdo especifico da disciplina, mas também quanto aos aspectos didáticos ao tema. Esta preparação deveria incluir o conhecimento das dificuldades e erros que os alunos encontram na aprendizagem de Estatística .

Segundo a autora citada, as pessoas que se interessam pela educação estatística têm como preocupação fundamental identificar os pontos difíceis e os equívocos que continuam após a aprendizagem da disciplina.

Assim, buscaremos nesta pesquisa diagnosticar o grau de conceitualização estatística do professor, para que se possa ajuda-lo na incorporação destes conceitos e sua implementação em sua sala de aula.

Após estas aulas aplicaremos uma seqüência de atividades elaborada por Carlos BIFI (2006), atividades esta que serviram de base para elaboração de mais três trabalhos que se desenvolvem de forma conjunta, em um mesmo projeto, são eles: PEREIRA, Sergio Alves: Um estudo a respeito do professor de matemática e a implementação de uma seqüência didática para a abordagem da estatística no ensino médio. VASQUES, Ricardo Sergio Braga: Mobilização dos Conceitos Estatísticos - um estudo diagnóstico desses conceitos, envolvendo variabilidade, com os alunos do ensino médio. CARDOSO, Ricardo: O professor de matemática e a análise exploratória de dados no ensino médio.

A partir destes resultados atingidos com estas pesquisas buscaremos uma generalização, para assim, organizarmos um plano de trabalho eficaz para auxiliar o professor do Ensino Médio no trabalho com os conceitos estatísticos.

1 OBJETIVO

Identificar como o professor de Matemática do Ensino Médio ensina estatística a seus alunos.

Verificar como este professor passa a ensinar a pós passar por uma formação Estatística.

Conhecer como os PCN apresentam a proposta do ensino de Estatística para o Ensino Médio.

1.1 JUSTIFICATIVA

A Estatística faz parte da Educação geral desejável a toda pessoa culta, além de capacitá-la para interpretar e avaliar de modo crítico a informação quantitativa nos meios de comunicação e no trabalho, e fornecer capacidade para discutir e comunicar sua opinião sobre informações estatísticas.

Os conhecimentos de Estatística sempre desenvolvem habilidades no sujeito e a formação estatística adequada segundo os preceitos da Educação, permite ao aluno construir seus conhecimentos de forma autônoma pela mediação do professor.

Desse modo, o aluno pode construir os conceitos de base da Estatística para poder mobilizá-los de modo adequado em situações concretas de sua vida.

A justificativa do presente estudo é reforçar a idéia de que o ensino da Estatística no ensino Médio deve ser direcionado a contextos reais da vida prática do aluno para que ele possa vivenciar e transferir seus conhecimentos para futuras tomadas de decisões.

Consideramos que este trabalho possa ser útil na pesquisa em Educação Matemática e poderá contribuir para uma Matemática com uma didática coerente com a proposta educacional.

Nossa pesquisa será caracterizada como um estudo de caso, com descrições pormenorizadas de tudo o que ocorreu em sala de aula na concretização das tarefas, adaptando tal abordagem qualitativa para as tarefas propostas.

Desse modo, o presente estudo será estruturado em quatro partes:

A primeira parte refere-se à introdução do estudo, objetivos e justificativa.

A segunda trata dos estudos preliminares a respeito do Ensino da Estatística no Ensino Médio.

A terceira parte apresenta o problema de pesquisa e os procedimentos metodológicos.

A quarta parte refere-se à fase experimental com apresentação das atividades e a análise a priori dos mesmos.

2 ESTUDOS PRELIMINARES

Neste capítulo, abordaremos o ensino da estatística no Ensino Médio. Apresentaremos uma análise da importância da Matemática no ensino da disciplina de estatística, abordando os objetos matemáticos envolvidos nos cálculos estatísticos e os obstáculos pedagógicos que a Matemática traz ao ensino da disciplina e quais as expectativas dos alunos frente à disciplina da Matemática e da estatística.

Levantamentos históricos nos mostram que desde os tempos remotos os governantes interessavam-se por conhecer informações sobre sua população e riquezas, que poderiam ser utilizadas com fins tributáveis e militares. Levantamentos feitos na China mostram que há mais de 2.000 anos antecedendo à Era Cristã, os conhecimentos estatísticos já existiam, no Egito os faraós já faziam uso deste conhecimento, conforme mostram os dados arqueológicos. Estes estudos mostram que os Astecas, Incas e Maias, também, já faziam uso sistemático desse conhecimento.

Em Roma, a estatística era usada no recenseamento dos judeus, ordenado pelo Imperador Augusto, para fazer balancetes, controlar suas posses, terras e riquezas. Estes são exemplos anteriores à estatística descritiva que surgiu no século XVI, na Itália. Este estudo era parecido com o que hoje faz o Instituto Brasileiro de Geografia e Estatística (IBGE), órgão responsável por nossa estatística oficial.

2.1 A ESTATÍSTICA COMO OBJETO DE ENSINO

O termo estatística surge da expressão em Latim statisticum collegium palestra sobre os assuntos do Estado, de onde surgiu a palavra em língua italiana statista, que significa "homem de estado", ou político, e a palavra alemã Statistik, designando a análise de dados sobre o Estado. A palavra foi proposta pela primeira vez no século XVII, em latim, por Schmeitzel na Universidade de Lena e adotada pelo acadêmico alemão Godofredo Achenwall. Aparece como vocabulário na Enciclopédia Britânica em 1797, e adquiriu um significado de coleta e classificação de dados, no início do século 19.

Desde então a estatística vem sendo aprimorada e cada vez mais é aplicado aos mais variados campos do conhecimento. Assim, as tabelas tornaram-se mais complexas e surgiram representações gráficas cada vez modernas. A teoria da probabilidade consolidou a fundamentação teórica da inferência estatística e a Estatística deixou de ser uma simples catalogação de dados numéricos coletivos para se tornar o estudo de como chegar a conclusões sobre o todo (população), partindo da observação do todo (amostra)

Os resultados obtidos com a aplicação dos métodos estatísticos na resolução de problemas dos diversos domínios do conhecimento, aliados à evolução tecnológica dos últimos anos, fizeram com que a estatística se tornasse indispensável nas grades curriculares, tanto no Ensino Fundamental e Médio, conforme orientações fornecidas pelos PCN como no Ensino Superior.

Nos dias de hoje, a estatística apresenta inúmeras definições porém a que mais se aproxima da forma de conduzirmos este trabalho é de Barnett (1973, apud Cordani, 2001, P.1): “ a estatística é o estudo de como a informação deveria ser empregada como reflexão e ação em uma situação prática envolvendo incerteza”.

Mas, para podermos entender de fato a definição de Estatística, precisaremos compreender os conceitos fundamentais, é o que veremos no capítulo seguinte.

2.2 A ESTATÍSTICA E SEUS CONCEITOS DE BASE

O texto apresentado a seguir foi elaborado por um grupo de alunos do Mestrado Profissional em Educação Matemática da PUC-SP, participantes do projeto no qual este trabalho está inserido (PEA-ESTAT: processos de ensino e aprendizagem em Estatística), orientados pela Profa Dra. Cileda de Queiroz e Silva Coutinho. Desta forma, este texto será comum a todos os trabalhos do projeto e, a partir dos subprojetos, será feita a pesquisa que é foco em cada um dos trabalhos.

Consideramos, no projeto maior, como conceitos estocásticos elementares, aqueles que estão na base do desenvolvimento do raciocínio estocástico e, por conseqüência, da alfabetização estocástica.

Adotaremos assim como conceitos elementares de Estatística àqueles enunciados por Gal (2002, p.10) como os cinco blocos do conhecimento estatístico de base: o reconhecimento da necessidade dos dados e de como estes podem ser produzidos; a familiaridade com termos básicos e idéias relacionadas com a estatística descritiva, familiaridade com termos básicos e idéias relacionadas com representações gráficas e tabulares, a compreensão de noções básicas de probabilidade e, finalmente, o conhecimento do alcance das conclusões estatísticas ou inferências.

Como o nosso trabalho visa o nível de escolaridade relativo ao Ensino Médio, vamos tomar o significado de número e os conhecimentos básicos da Estatística Descritiva como ponto de partida para a construção do raciocínio estatístico: organização e representação de um conjunto de dados, cálculo ou determinação de medidas-resumo e percepção da variabilidade.

Lembramos que os conceitos apresentados a seguir devem ser abordados a partir da resolução de problemas (PCN, 1997, p. 40). Por isso é importante que sejam apresentados de forma contextualizada para que permita a experimentação e a criação de modelos que levem ao desenvolvimento do raciocínio estatístico.

Apresentaremos na seqüência as idéias básicas relacionadas a alguns destes conceitos.

Ao final do Ensino Médio em relação à Estatística, o aluno deve ser capaz de: organizar questionários, como instrumento de coleta de dados, resumir os dados em tabelas com freqüência absoluta e relativa, construção de gráficos, calcular medidas resumo (média, mediana e moda, desvio-padrão e quartis, estas duas não constantes dos PCN, mas que julgamos importantes), articular com outras áreas da Matemática, como álgebra (frações, porcentagem, números decimais), tópicos de geometria (circunferência, ângulos, retângulo e área), entre outras.

Em probabilidade, adotaremos os conceitos de base enunciados em Coutinho (2001): a percepção do acaso, a idéia de experimento aleatório e a noção de probabilidade. Ou seja, é necessário que o sujeito que está pronto para resolver um problema no campo da probabilidade perceba que a situação a ser analisada não é determinística, que envolve o desenvolvimento de uma experiência aleatória (reprodutível nas mesmas condições, na qual o resultado final não se pode calcular ou determinar, mas se pode identificar todas as possibilidades de resultados), e que o evento observado, resultado dessa experiência aleatória, pode ser avaliado em termos probabilísticos por uma razão entre o número de sucessos possíveis e o número total de casos (nos termos enunciados por Laplace (1814, apud Coutinho, 2001, p.37) em seu segundo princípio: « A probabilidade de um evento é igual à razão entre o número de casos favoráveis e o número de casos possíveis. »)

Logo, ao final do Ensino Fundamental, o aluno deverá ser capaz de observar e descrever um experimento através de diagrama de árvores, calcular a probabilidade de um evento, relacionar a Estatística e a Teoria das Probabilidades, assim como resolver problema de contagem.

Portanto essa análise tem por objetivo verificar se ao final do Ensino Médio, estes conceitos tenham sido construídos pelos alunos, para que possa fazer observações dos aspectos quantitativos e qualitativos em situações da vida cotidiana, e também estabelecer o maior número possível de relações entre eles, fazendo assim, com que percebam o que pode acontecer ou qual sua chance de acontecer. Com isso o aluno pode tomar consciência da utilização da Matemática no seu dia-a-dia.

A primeira delas diz respeito à diferenciação entre população e amostra. Segundo Barbetta (2003), uma população é o conjunto de elementos que formam o universo do nosso estudo e que são passíveis de serem observados. Uma amostra é qualquer subconjunto finito da população.

Uma vez definida a população (amostra), precisamos definir o que será observado, logo outra idéia básica é a variável estatística.

Variável estatística é a característica da população que se quer observar. Esta característica pode ser qualitativa (nominal ou ordinal) ou quantitativa (discreta ou contínua). Uma variável qualitativa nominal descreve uma qualidade sem, no entanto, estabelecer níveis de hierarquia. Citamos, entre outras, música predileta, esporte preferido, cor dos olhos. Uma variável qualitativa ordinal descreve uma qualidade, mas identificando níveis hierárquicos. Citemos por exemplo nível de escolaridade, classe sócio-econômica, ou qualquer tipo de opinião expressa em tipo de escalas. As variáveis quantitativas são expressas por números. As quantitativas discretas são aquelas cujo conjunto admite uma relação biunívoca com o conjunto dos Números Naturais (ou seja, é um conjunto enumerável). Para Barbetta (2003), são as variáveis que "só assumem valores que podem ser listados". As variáveis quantitativas contínuas são aquelas que, contrariamente às variáveis discretas, podem assumir qualquer valor em um intervalo real.

Uma vez coletados os dados[1] passa-se à sua organização e representação, seja em forma de tabelas de distribuição de freqüências ou em forma gráfica. Uma distribuição de freqüências pode ser considerada uma função empírica, na qual a cada valor da variável observada é associado a sua freqüência.

Cada tipo de representação da distribuição de freqüência construída nos dá uma visão sobre diferentes aspectos e, portanto a soma dessas representações faz com que possamos desenvolver uma análise exploratória de dados satisfatória.

Segundo Batanero:

> Uma idéia fundamental da análise exploratória de dados é o uso de representações múltiplas de dados e se converte em um meio de desenvolver novos conhecimentos e perspectivas. Por exemplo, passar de listagem de números a uma representação do tipo ramo e folha, pode facilitar a exploração da estrutura total, assim como construindo gráficos, como o de Box-Plot que

[1] De um ponto de vista didático, os resultados de pesquisas indicam que esta coleta deve, preferencialmente, ser feita por meio de pesquisa realizada com os próprios alunos de um assunto que seja de interesse comum para os alunos e professores.

> possibilite a comparação de várias amostras, Batanero (2001, p.29).

Quanto às representações gráficas para variáveis unidimensionais, podemos citar o diagrama de setores, o diagrama de colunas ou de barras, o histograma, os gráficos de linha e o box-plot. Todos estes gráficos são de pouca complexibilidade cognitiva, na medida em que exigem poucos pré-requisitos matemáticos, podendo, portanto ser explorados desde o Ensino Fundamental.

Neste texto, abordaremos alguns deles, que podem ser usados pelos sujeitos de nossa pesquisa durante a resolução das atividades propostas.

O Histograma, construído no sistema de eixos ortogonais, é uma representação gráfica em forma de colunas justapostas. Para sua construção deve ser observado que a área total do histograma deve ser proporcional à freqüência total, e as áreas parciais (área de cada retângulo), proporcionais às freqüências das classes. As bases dos retângulos são proporcionais a amplitude do intervalo de classe. Assim, amplitudes iguais levam à bases de mesma medida, enquanto que amplitudes distintas devem ser representadas por medidas distintas. Caso as amplitudes dos intervalos de classes sejam unitárias, a altura de cada retângulo terá como correspondente suas respectiva freqüência.

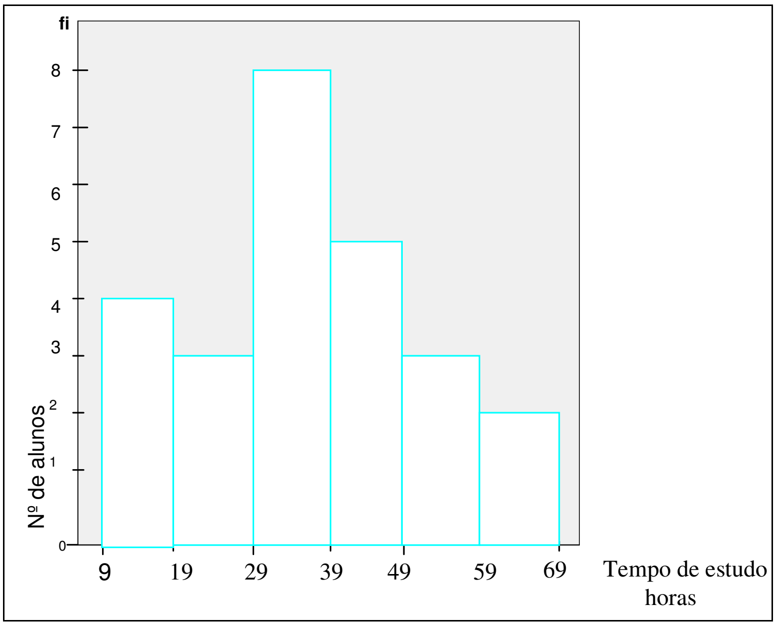


Gráfico 1: Tempo mensal de estudo de 25 alunos de um curso

Segundo Barbeta (2003), uma maneira de apresentar aspectos relevantes de uma distribuição de freqüência é através do chamado diagrama em caixas ou box-plot ou desenho esquemático. Sua construção se inicia pela determinação do primeiro quartil, mediana e o terceiro quartil e é preciso também observar o limite inferior e o superior. Vejamos construção: traçamos uma semi-reta com as medidas acima indicadas proporcionalmente. Sobem-se três retas verticais de altura qualquer a partir das medidas primeiro quartil, mediana e terceiro quartil, traçam-se duas retas paralelas não coincidentes a semi-reta formando na maioria das vezes dois retângulos como podemos observar no gráfico abaixo (observação nem sempre aparecerão dois retângulos, pois existe a possibilidade de coincidência entre as medidas), observando a semi-reta interpreta-se que quanto maior a distância menor a concentração e quanto menor a distância maior a concentração.

Caso o primeiro quartil coincida com a mediana a variabilidade é nula existe uma concentração total entre o primeiro quartil e a mediana, se a mediana coincidir com o terceiro quartil pode-se analisar analogamente.

Casos existam valores discrepantes – além de 1,5 (dQ) -,a linha é traçada até o ultimo valor não discrepante; e os valores discrepantes são indicados por pontos

No exemplo da distribuição de freqüência representada pelo histograma, (Gráfico 1) teremos:

Q1 = 26,5h

Q2 = Md = 35,875h

Q3 = 46,5h

Min. = 9h

Max. = 69h

Passando estes dados para um eixo orientado:

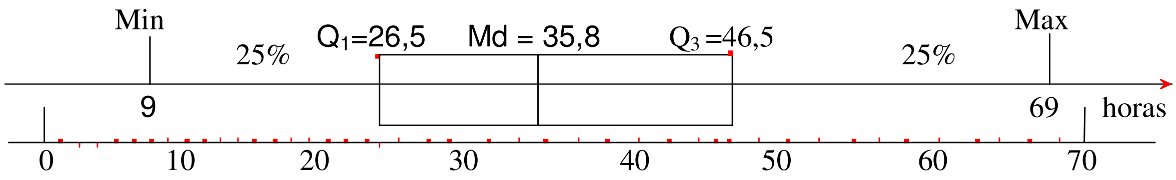


Gráfico 2: Tempo de estudo de 25 alunos

Este tipo de gráfico permite visualizar o conjunto de dados organizado em 4 grupos, com 25% dos elementos em cada um destes grupos, podemos observar por este gráfico que 50% dos alunos estudam entre 26,5h à 46,5h e que entre o primeiro quartil e a mediana existe uma concentração um pouco maior que a mediana e o terceiro quartil.

Vejamos um outro exemplo para a complementação possível na interpretação do histograma, da tabela e do box-plot.

Seja a idade de quarenta participantes da associação de pais e mestres da “escola da criança do futuro.”

Tabela 1: Idades dados hipotéticos

Idade (Anos)	Fi
[18 ; 22[	2
[22 ; 26[	4
[26 ; 30[	6
[30 ; 34[	12
[34 ; 38[	9
[38 ; 42[	7
Σ	40

Fonte: Arquivo do pesquisador

De acordo com a tabela acima, podemos observar as variação entre as idades, a classe modal, o número de participantes por cada classe e o de participantes.

O histograma nos permite apreender os aspectos relativos a localização dos valores centrais, a densidade em torno da média, a classe modal, formada da distribuição, as idades superior e inferior

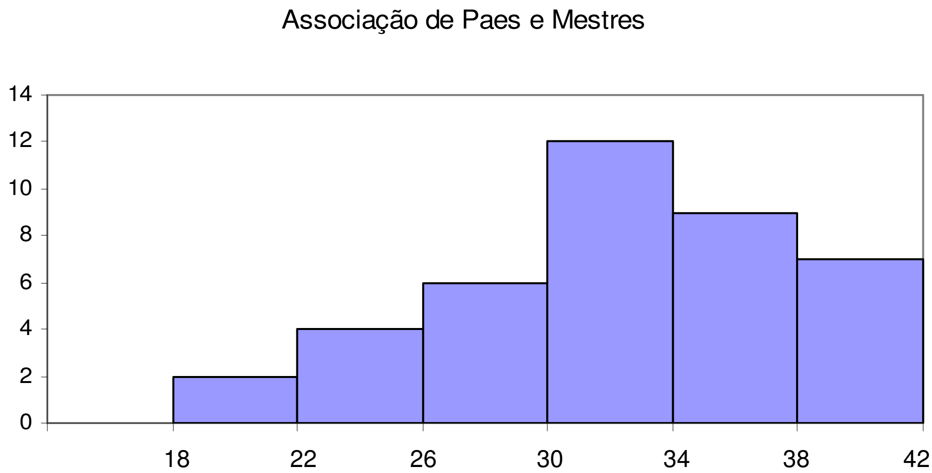


Gráfico 3: Histograma Associação de Paes e mestres

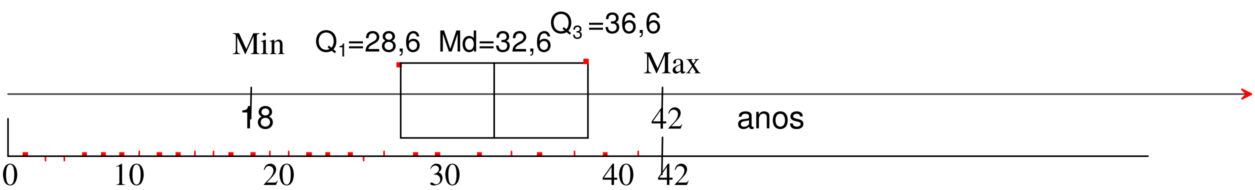


Gráfico 4: Box-plot Associação de Paes e mestres

Podemos notar pelo box plot, que entre o primeiro e o terceiro quartil existe uma amplitude de 8, isso significa dizer que aproximadamente 33% das idades da tabela estão localizadas neste intervalo. E que de 18 à 32,66 anos representam 50% das idades da tabela e 32,66 até 42 anos representam os outros 50% dos dados da tabela. Quando comparado com a média que é de 32,3, e o desvio-padrão que é 5,61, temos uma amplitude de 11,22, para uma amplitude total de 24, isso, significa dizer que 46,75% das idades ficam representadas neste intervalo, portanto a média representa melhor está distribuição.

Para a interpretação das medidas-resumo de um conjunto de dados

Uma distribuição de freqüências (ou mesmo um conjunto de dados apresentado simplesmente em uma série) pode ser interpretada com auxílio de medidas de tendência central em associação com medidas de dispersão (é o que permite a apreensão da variabilidade). Um exemplo usual no ensino é um conjunto formado por valores assumidos por uma variável quantitativa, a amplitude total exprime um dos tipos de variação que pode ser observado neste conjunto. É definido como sendo a diferença entre o maior e o menor valor observado. No exemplo**,** Gráfico 1: Tempo mensal de estudo de 25 alunos de um curso, que representa o tempo de estudo dos alunos, teríamos 69 – 9 = 60. Ou seja, a amplitude total é de 60 minutos.

A Média pode ser considerada como o ponto de equilíbrio de uma distribuição. É um valor típico ou representativo dos dados. Podemos dizer que a média de uma distribuição é o valor em torno do qual os demais valores oscilam. Calculada a partir de um algoritmo de custo cognitivo bastante baixo.

$$u = \frac{\sum_{i=1}^{N} X_i}{N} \quad \text{ou } \overline{\text{X}}\text{, no caso de dados amostrais}$$

No entanto, perde muito do seu significado se não pudermos identificar o grau dessa concentração dos valores ao redor da média.

O Desvio-padrão é à medida que melhor indica essa concentração. Seu valor é determinado pelo cálculo da raiz quadrada da variância, que por sua vez, é determinado pelo cálculo da média dos quadrados da diferença entre cada valor observado e a média. Ou seja:

$$\sigma^2 = \frac{\Sigma xi^2 - \frac{(\Sigma xi)^2}{N}}{N}$$

Tratamos aqui a variância populacional. Para o caso da variância amostral:

$$s^2 = \frac{\sum_{i=1}^{k} (xi - \bar{x})^2}{N - 1}$$, ou

$$s^2 = \frac{\sum_{i=1}^{k} xi^2 - \frac{(\sum_{i=1}^{k} xi)^2}{N}}{N - 1}$$

Não faremos, neste bloco, a discussão sobre a correção N – 1 que aprece para a variância amostral, mas isso pode ser pesquisado em livros específicos da área de estatística.

Segundo Pereira e Tanaka (1990, p. 97), mediana é o valor central de um rol, ou seja, a mediana de um conjunto de dados ordenados (crescentes ou decrescentes) é a medida que divide este conjunto em duas partes iguais.

Quartis são números reais que dividem a série de Rol em quatro partes com o mesmo número de elementos. Assim, devem-se ter três quartis: Q1 representa o primeiro quartil é precedido por 25% dos valores da distribuição; o segundo quartil Q2 é precedido por 50% e o terceiro quartil Q3, é representado por 75%. Observe que o segundo quartil Q2 coincide exatamente com a mediana. Para o cálculo da posição dos quartis, usamos processo análogo ao da mediana, ou seja: $PQ_1 = \frac{N+1}{4}$ para $\mathrm{PQ}_3 = \frac{3.(N+1)}{4}$

Outra ferramenta fundamental: a média e sua associação com o desvio-padrão e a mediana com sua associação com os quartis. Sem essas associações fica difícil a percepção da variabilidade dos dados. Por exemplo podemos ter dois ou mais conjuntos de dados com o mesmo valor de média, ou de mediana porém esse conjunto pode diferir em sua distribuição.

Vejamos o exemplo das quatro notas de uma determinada avaliação e sua respectiva média.

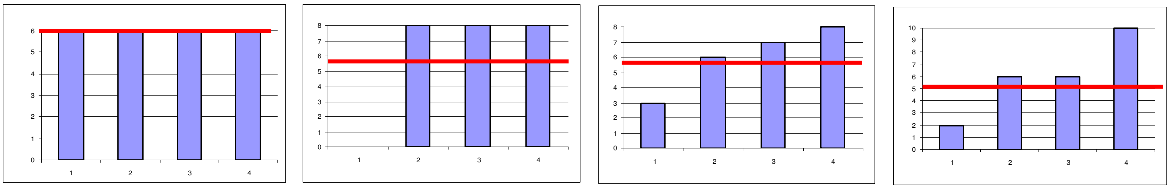
Tabela 2. Notas de avaliação e sua respectivas médias

Aluno 1	6 – 6 – 6 – 6	$\mu_1 = 6$
Aluno 2	0 – 8 – 8 – 8	$\mu_2 = 6$
Aluno 3	3 – 6 – 7 – 8	$\mu_3 = 6$
Aluno 4	10 – 2 – 2 – 10	$\mu_4 = 6$

No exemplo acima podemos observar que a média, nos quatro casos, é mesma, porém a variação em cada um dos conjuntos não. Logo, a simples aplicação do algoritmo para o cálculo da média não garante a percepção da variabilidade.

Associação entre média e desvio-padrão é garantida pela própria definição e seu algoritmo e a utilização de representações podem nos ajudar na compreensão e são representadas a seguir:

Aluno 1 Aluno 2 Aluno 3 Aluno 4



As representações gráficas acima, nos permitem visualizar a evolução de cada aluno a cada avaliação aplicada, ou seja, visualizar a variação das quatro notas.

Esperamos que os conceitos apresentados neste capitulo ajude na compreensão deste trabalho, e do capítulo que vêm a seguir pois estaremos falando das dificuldades e erros que os alunos enfrentam.

2.3 IDENTIFICAÇÃO DAS DIFICULDADES NA ÁREA DO CONHECIMENTO

Em toda pesquisa sobre ensino e aprendizagem em uma área do conhecimento têm-se como objetivo principal identificar os pontos difíceis e os erros que surgem durante a aprendizagem. Na Matemática, existem vários estudos sobre os erros conceituais ou algorítmicos na disciplina. Na Educação Estatística, não é diferente. Nosso trabalho também buscará este enfoque, no qual tentaremos identificar quais as dificuldades e erros que os alunos enfrentam.

Batanero (2001, apud Bifi 2006) afirma que as dificuldades não acontecem de forma aleatória, encontramos erro que se repetem regularmente e produzem associações com variáveis próprias das tarefas propostas, dos sujeitos, de circunstâncias presentes ou passadas e, até mesmo, das concepções dos alunos.

Assim, verificamos que esses conceitos são apoiados por alguns outros, como ferramenta para construção e mobilização dos conceitos estatísticos aqui relacionados.

Ainda segundo a autora, a concretização da tarefa de investigação permite criar condições para que os alunos pensem estatisticamente, definindo objetos e traçando seus próprios caminhos. Formular questões e conjecturas e pôr em prática estratégias de validação dessas conjecturas, criticar e comunicar os resultados obtidos são algumas das competências que podem ser desenvolvidas com este tipo de tarefa. Uma investigação diz-se estatística se, em sua concretização, o aluno utilizar metodologias quantitativas, integrando a linguagem e os métodos estatísticos a um processo mais global de investigação.

Em nosso trabalho investigativo, esperamos que os alunos formulem questões e planejem estudos que lhes permitam responder as questões norteadoras desta nossa pesquisa. Estes estudos englobam a tomada de decisões, quanto ao tipo de dados que necessitam e ao modo de coletar a interpretação desses dados. A análise e interpretação dos dados pressupõe sua manipulação de várias formas. As conclusões obtidas poderão responder ou não às questões de investigação e, neste último caso, será necessário recolher novos dados e/ou reformular as questões de investigação. Uma vez terminado o estudo, os alunos comunicam os resultados de sua investigação, tendo o cuidado de preparar argumentos para defender as opções que tomaram e as interpretações que fizeram ao longo do processo de investigação.

> Pelas investigações estatísticas, o contato com as técnicas e instrumentos de coleta de dados e seus diferentes modos de representá-los e sintetizar surgem da necessidade efetiva de usá-los. Os alunos podem ser envolvidos em uma aprendizagem autêntica dos processos e conteúdos estatísticos (HEATON; MICKELSON, 2002, p. 39 apud BIFI, 2006).

Além disso, conforme os autores é muito conhecida a facilidade com que a estatística pode ser integrada e aplicada ao estudo de situações que envolvem conteúdos matemáticos e não-matemáticos. Apoiados nesses argumentos, os autores referidos citam que parece mais razoável e, possivelmente, mais praticável que a estatística seja integrada ao ensino elementar, promovendo a consistência e as conexões entre os assuntos do que tentar acrescentá-la a um currículo já demasiado extenso.

A realização de investigações estatísticas pode contribuir para que a prática de procedimentos deixe de se constituir uma atividade preparatória, repetitiva, isolada e sem significado e transforme-se em uma prática compreensiva capaz de promover nos alunos a aquisição de habilidades utilizáveis com segurança e autonomia (PCN, 1997)

A partir do PCN, quando foram incluídos os conteúdos de estatística no Ensino fundamental e Médio como parte da disciplina de matemática, as editoras e os autores dos livros didáticos deveriam buscar não ficar deslocados do contexto político e cultural, já que o livro didático tem uma função de legitimação de sistemas de ensino muitos profissionais da educação o utilizam como fonte principal para planejar suas atividades e sua prática pedagógica, veremos algumas considerações no próximo capítulo.

2.4 O LIVRO DIDÁTICO: CONSIDERAÇÕES

No ensino da Matemática, o livro didático é um dos recursos de ensino que influência a organização, o desenvolvimento e a avaliação do trabalho pedagógico.

Segundo Fonseca (1999)

> O livro didático e a educação formal não estão deslocados do contexto político e cultural e das relações de dominação, sendo, muitas vezes, instrumentos utilizados na legitimação de sistemas de poder, além de representativos de universos culturais específicos. (...) Atuam, na verdade, como mediadores entre concepções e práticas políticas e culturais, tornando-se parte importante da engrenagem de manutenção de determinadas visões de mundo. (p. 204, apud Corrêa), O livro escolar como fonte de pesquisa em História da Educação.

Entretanto, para (Lajolo e Zilberman, 1999, apud Corrêa) apesar de ilustre, o livro didático é o primo pobre da literatura, texto para ler e botar fora, descartável porque anacrônico: ou ele fica superado dados os progressos da ciência a que se refere ou o estudante o abandona, por avançar em sua educação. Sua história é das mais esquecidas e minimizadas, talvez porque os livros didáticos não são conservados, suplantado seu "prazo de validade".

O professor utiliza o livro didático de forma ostensiva, para preparar suas aulas, selecionar os conteúdos, definir a seqüência lógica dos conteúdos, selecionar questões para as provas, Este também é uma fonte de atividades e guia para escolha de pesquisas.

3. PROBLEMA DE PESQUISA

O ponto de partida desta pesquisa vem do fato que a estatística cada vez mais vem conquistando um lugar de destaque na educação.

Os Parâmetros Curriculares Nacionais (1997) incluíram conteúdos de estatística no Ensino Fundamental e Médio como parte do programa da disciplina de matemática, este fato sem dúvida foi muito importante pois, o mundo moderno requer do cidadão uma consciência estatística para poder compreender as informações que são trazidas, podemos considerar que os conhecimentos estatísticos sempre desenvolve no cidadão habilidades para compreender a abstração lógica que se faz possível no estudo quantitativo dos fenômenos coletivos.

A comunidade cientifica passou a pesquisar com maior afinco o processo de ensino e aprendizagem por concordar que a estatística cada vez mais tem conquista um lugar de destaque. Pesquisadores do mundo todo assinalam para a necessidade desta cultura estatística.

> Em uma sociedade mutante e imprevisível como a que estamos vivendo, nós nos sentimos inseguros sobre qual é a melhor forma de preparar nossos jovens, quais são os conteúdos que devemos ensinar e que não se tornarão obsoletos em pouco tempo. Quais valores permanecem inalteráveis ou preparam para a busca de atualizações para se continuar ativo no mercado? Nesse contexto, a Estatística se apresenta como parte da educação geral desejável a todo cidadão culto, pois, além de capacitá-los para interpretar e avaliar criticamente a informação quantitativa nos meios de comunicação e no trabalho, fornece capacidades para discutir e comunicar sua opinião a respeito de informações estatísticas, como assinala Ottaviani, 1998, apud Batanero, 2002, p.1).

Os parâmetros Curriculares Nacionais (PCN) para o Ensino Médio propõem o bloco de Análise de Dados.

Assim todo cidadão deveria receber uma formação adequada a respeito da estatística. Chamo de formação adequada aquela que, permite ao aluno construir seus conhecimentos de forma autônoma, por meio da mediação do professor, que é de fundamental importância para este aprimoramento, tudo isto nos leva a pensar em uma questão que se subdividem-se em três partes são elas:

- Como o professor de Ensino Médio lida com os conteúdos de estatística? - Qual a opinião que esse professor têm sobre a proposta de trabalhar com tais conteúdos no Ensino Médio? - Como os alunos vão resolver as questões?

Desta forma, nesta pesquisa, pretendemos verificar o nível de conceitualização segundo Robert (1998), são eles (nível técnico, mobilizável e disponível) dos conhecimentos desse componente curricular nos alunos do Ensino Médio.

3.1 PROCEDIMENTOS METODOLÓGICOS

Este trabalho trata de um estudo de caso sobre professor e alunos, sendo que o professor receberá uma formação sobre o que se pretende com a pesquisa por parte do pesquisador, também o pesquisador manterá alguns encontros com o professor para que se possam ser discutidos assuntos ligados a estatística e aos conceitos de base que serão utilizados nesta pesquisa. O professor em contra partida utilizará dos resultados destes encontros para trabalhar com seus alunos que nos ajudará a responder nossa questão de pesquisa que se subdivide em três partes.

Desse modo, surgem as seguintes questões norteadoras:

- Como o professor de Ensino Médio lida com o conteúdo de Estatística?

- Qual a opinião que esse professor tem sobre a proposta de trabalho com tais conteúdos no Ensino Médio?

- Como os alunos vão resolver as questões?

A nossa pesquisa apoiada nestas questões não se limitará somente a uma descrição quantitativa, mas a um estudo especifico que envolve o desempenho profissional e pedagógico do professor.

Fizemos opção pelo método de natureza quanti-qualitativa para direcionar esta pesquisa, que poderá responder ou não nosso objetivo.

Para a realização do nosso trabalho primeiramente submetemos o professor colaborador a um questionário, que foi gravado em áudio e depois transcrito. Este questionário tem como finalidade definir a partir de onde começaremos nosso trabalho com o professor.

As dificuldades deste professor sobre os conceitos estatísticos serão trabalhadas em cinco encontros de aproximadamente 2h15 cada e mais um encontro na própria sala de aula juntamente com seus alunos, sendo que este último encontro preparatório será apenas para discutirmos sobre análise de resultados estatísticos.

Durante a preparação do professor, que será feita pelo próprio pesquisador, também se trabalhará os conceitos em paralelo com seus alunos, que são alunos de uma terceira serie do ensino médio, de modo que em cada encontro, este possa fazer comentários de como tudo estará ocorrendo e sanar também as possíveis dificuldades didáticas que poderão surgir.

Ao final do ultimo encontro, marcaremos os testes que nos ajudarão a responder nossas questões de pesquisa.

Nesta investigação, esperamos que os alunos formulem questões e planejem estudos que lhes permitam responder tais questões. Estes estudos englobam a tomada de decisões, quanto ao tipo de dados que necessitam e ao modo de coletar a interpretar esses dados. A análise e interpretação dos dados pressupõem sua manipulação de várias formas. As conclusões obtidas poderão responder ou não às questões de investigação e, neste último caso, será necessário recolher novos dados e/ou reformular as questões de investigação.

Uma vez terminado o estudo, os alunos comunicarão os resultados de sua investigação, tendo o cuidado de preparar argumentos para defender as opções que tomaram e as interpretações que fizeram ao longo do processo de investigação.

Para responder a nossa questão de pesquisa, apresentada no capítulo (3 Problema de Pesquisa) será proposta uma situação-problema na forma de uma atividade diagnostica qualitativa, dividida em três etapas. Esperamos que esta atividade permita diagnosticar qual o nível de funcionamento dos conceitos, segundo os preceitos de Robert (1998) e, especificamente, aqueles ligados ao estudo da variabilidade, por parte dos alunos e, também, permitir identificar o(s) possível(is) erro(s) cometido(s) por estes alunos.

Acreditamos que as situações-problema, quando apresentadas em várias formas de representação, abrem um leque maior de possibilidades de investigar as possíveis dificuldades que os alunos poderão encontrar na resolução de qualquer atividade diagnostica. Sendo assim, as etapas da atividade se apresentarão em diferentes níveis de complexidade, para que, então, seja possível verificar o nível de mobilização dos conceitos de (Robert, 1998) em que estes alunos se encontram. A situação-problema será aplicada para grupos de alunos que poderão dispor do meio que lhe for mais conveniente para tal resolução, sem que influenciemos nas possíveis interações sociais (diálogo entre os participantes) que acontecerão durante a aplicação da atividade, pois é fundamental para o pesquisador respeitar a forma de resolução dos participantes, deixando-os livres para utilizarem seus próprios meios. Farão uso de calculadoras, ainda teremos a gravação dos diálogos de cada grupo em áudio, será mantida durante a apresentação das atividades a presença do professor e do pesquisador e ainda buscaremos a não contaminação de um grupo para com o outro.

O teste a ser proposto ao final das atividades com os alunos está dividida em três partes com diferentes formas de apresentação. A primeira parte será composta de um banco de dados fictício, no qual constarão idade e renda mensal de quarenta pessoas entrevistadas por uma Empresa de cartões de crédito. Na segunda parte, são apresentadas duas distribuições na forma de tabelas, sendo a primeira sem intervalo de classes e a segunda com intervalo de classes. Na primeira distribuição, relata-se o comportamento de quantidade de carros por número de pessoas, e a segunda distribuição relata o comportamento do tempo no trânsito por número de pessoas.

E, por fim, a terceira parte relata as tabelas da atividade da segunda parte, em forma de gráficos. Nossa idéia em apresentar a atividade em diferentes formas de representação nos ajudará, além do que nos propomos a investigar, é a de identificar se os níveis de dificuldades também se apresentam em diferentes formas. O que queremos dizer é que se um conjunto de dados apresentados ao aluno, estiver na forma de dados brutos (sem nenhum tipo de organização crescente ou decrescente), as dificuldades serão maiores do que se estes mesmos dados forem apresentados com algum tipo de forma organizacional, por exemplo, na forma de gráficos ou tabelas.

As respostas obtidas nesta atividade serão organizadas, tabuladas por categorias e, posteriormente, analisadas segundo os preceitos de Robert (1998).

Esperamos que, por meio desta atividade, possamos atingir nossos objetivos e assim colaborarmos para que o ensino da disciplina de Estatística seja vista com olhar atento, e que as possíveis dificuldades apresentadas pelos alunos não sejam mais encaradas como “rotineiras”, bem como as pesquisas futuras apresentem quadros satisfatórios de aprendizados.

3.2 QUADRO TEÓRICO

A produção dos alunos será analisada segundo os níveis propostos por Robert (1998): mobilizável, técnico e disponível.

O nível técnico, este nível corresponde para nós a dos focos em funcionamento indicados, isolados colocando em jogo as aplicações imediatas de teoremas, propriedades, definições fórmulas, etc. Ele contextualiza de maneira simples, locais sem adaptações.

Exemplo:

Um instituto de pesquisa lançou um balão. Esse balão sobe a uma velocidade constante, percorrendo 25 m a cada minuto.

Qual a altitude que o balão atinge após 18min do lançamento?

Tempo gasto (em minutos)	Altitude do balão (em metros)
1	25
2	2.25 = 50
3	3.25 = 75
4	4.25 = 100
10	10.25 = 250
T	t.25
18	18.25 = 450

Resp: atinge após 18min do lançamento, 450m.

O nível mobilizável corresponde aos focos em funcionamento mais amplos: ainda indicados, mas passando da aplicação simples de uma propriedade. Isso pode ser exemplo pois é preciso adaptar os conhecimentos para aplicar o teorema adequado, ou mudar de ponto de vista ou de quadro (com indicação), isso pode acontecer, pois é necessário aplicar diversas vezes seguidas a mesma coisa ou utilizar coisas diferentes em etapas sucessivas, ou pois é necessário articular duas informações de natureza diferentes. Em todos os casos, este nível testa um foco em funcionamento ou existe um início de justaposição de saberes dentro de um domínio já dado, observar a organização, não existe somente uma aplicação simples, os caractéres ferramentas e objetos podem ser concernidos. Mas o que está em jogo é explícito. Repetimos ao dizer, um saber é dito mobilizável se ele for bem identificado, ele é bem usado pelos alunos mesmo se ele foi adaptado a um contexto.

Um aluno consegue mobilizar os conhecimentos de desigualdade. No exemplo dado ele consegue fazer um relacionamento entre duas formas de representação de conjuntos.

Ex.: A= {x ∃ N / 1≤ X < 4}, faça sua representação onde aparecem todos os elementos desse conjunto.

Resposta: A= {1, 2, 3}.

Para a autora é necessária a adaptação ao conhecimento, no qual esse pode ser sugerido pelo professor ou pelo próprio enunciado do problema, ou seja, o que diferencia conhecimento mobilizável de conhecimento técnico é a possibilidade de estabelecer relações complexas entre o objeto visado a outros e a utilização desse objeto como ferramenta para resolução de problemas complexos. Por exemplo, podemos citar uma atividade que fornece os dados coletados de uma determinada variável e, logo em seguida, pede-se ao aluno uma análise estatística destes dados. Dependendo de como está transcrito o enunciado da atividade, ele pode sugerir um caminho a ser percorrido, ou seja, implicitamente o aluno é conduzido a utilizar cálculos algébricos ou até mesmo raciocínio lógico, para que se alcancem resultados favoráveis na resolução do problema.

O nível disponível corresponde em saber responder o que é proposto corretamente o que é proposto sem indicações, de poder, por exemplo, dar contra-exemplo (encontrar ou criar), fazer relações, aplicar métodos não previstos.

Este nível de conhecimento está associado ao conhecimento de referência variadas que o estudante conhece, servem de questionamentos e de organização. Podendo funcionar para um único problema ou possibilitando fazer resumos.

A teoria de Robert é comparada por Skemp (1978 apud Carvalho e César 2001, p. 7), na qual as autoras abordam o conhecimento como sendo instrumental e relacional. O primeiro é quando se denomina uma coleção isolada de regras e algoritmos apreendidos pela repetição e pela rotina. Quando o conhecimento que o sujeito possui é desse tipo, só conseguimos resolver um conjunto limitado de situações, em contextos semelhantes. O segundo é aquele, no qual o aluno construiu um esquema do conceito que pode ser atualizado sempre que novas situações exijam, ou seja, um conhecimento que consegue mobilizar em face de novas situações.

> “Professores e pesquisadores queixam-se freqüentemente de que a compreensão que os alunos têm da álgebra é meramente instrumental: as crianças são capazes de “avançar nos passos

necessários”, mas não são capazes de explicar aquilo que estão por fazer”.(SFARD;LINCHEVSKI,1994 a, p 203 apud BIFI,2006).

Para observar se os alunos conseguem mobilizar os conceitos de Estatística desenvolvidos durante as aulas ministradas pelo professor colaborador, utilizaremos os conceitos propostos por Robert (1998).

Algumas pesquisas voltadas para as dificuldades de aprendizagem da Estatística que contribuíram e continuam contribuindo para o aprimoramento deste tema e da divulgação de sua importância para a sociedade moderna e para a matemática. Entre elas, podemos citar (Silva, 2000; Vendramine 2000), Novais (2004) e Batanero (2001).

Vendramine (2000) pesquisou aproximadamente 73% dos alunos matriculados em sete cursos da área de ciências Humanas, ciências Exatas e ciências da Saúde, selecionados aleatoriamente entre 29 cursos de uma universidade particular. Apresentou estudos indicando que os alunos apresentavam atitudes negativas em relação à Estatística e desenvolvem ansiedade em relação a disciplina. A autora mostrou ainda que existe correlação positiva e significativa entre as atitudes dos alunos em relação à Matemática e à nota final da disciplina Estatística. Isto é, as atitudes negativas com relação à Matemática são transferidas para a Estatística. A sugestão que a autora nós traz é que se precisa desenvolver junto aos alunos atitudes, (Atitude é a prontidão de uma pessoa para responder a determinado objeto de maneira favorável ou desfavorável) positivas frente a Matemática e o tópico de Estatística.

> Uma variável que pode influenciar a aprendizagem de Estatística, levando o aluno a ter interesse, querer aprender mais e estudar quando apresenta atitudes positivas em relação à disciplina mas também pode tornar o aluno nervoso, ansioso, com medo e sem interesse de aprende-la quando esse aluno apresenta atitudes negativas em relação a ela. (Silva, 2000, p. 15).

Essa afirmação, levanta a hipótese da necessidade de uma abordagem de situações didáticas adequadas para a matemática e em especial para a disciplina de estatística a qual estamos tratando, tem que ser uma abordagem, voltada para a construção, por parte dos alunos, dos conceitos básicos da estatística, deveria se apresentar problemas envolvendo dados reais, concentrando-se em aspectos que não necessitem de memorização, mais sim de interpretação, estratégias para exploração dos dados, com um diagnóstico básico preliminar para a inferência.

É importante alertar os professores de Estatística sobre a necessidade de elaborar programas visando o desenvolvimento de atitudes positivas em relação à Estatística e o desenvolvimento das habilidades matemáticas necessárias para a aprendizagem dessa disciplina, como, por exemplo, a leitura acurada e atenta da proposição do problema visando a obtenção da informação matemática (Vendramini, 2000).

Novaes (2004), em seu trabalho, investigou a mobilização dos conceitos estatísticos de base em alunos do curso superior de Turismo. A autora concluiu que os erros cometidos por esses alunos foram os que envolviam processos algébricos; erros nos conceitos de média, moda e mediana; análise inadequada da variabilidade dos dados e dificuldades vindas de obstáculos epistemológicos e didáticos na resolução de situações-problema.

Ela também pode observar que os erros se repetem e produzem regularidades, sendo que os sujeitos investigados apresentavam, em sua maioria, as mesmas dificuldades. Assim, a autora deixou proposto em seu trabalho o ensino dos conteúdos estatísticos com pelo menos uma ordenação, na forma de situações-problema do que se pretende ensinar, de forma a facilitar a construção dos conceitos por parte dos alunos, bem como a sua adequada utilização.

Desta forma o trabalho de Novaes (2004) vem a colaborar muito em nossa pesquisa. E nos ajudará, pois a manteremos como referencial para futuras análises, em nossa pesquisa que pretende reforçar a necessidade de uma atenção especial à metodologia do ensino da Estatística, mas, também, proporcionar recursos teóricos para futuras pesquisas na área, que sofre de enorme carência.

De acordo com os trabalhos citados, os alunos, em todos os níveis de formação, associam as dificuldades que tiveram com a Matemática às dificuldades atuais e também com a Estatística. Quando reconhecem a importância dessa área do saber para a resolução de situações-problema da vida, ficam ansiosos ao sentirem dificuldades, o que os leva a atitudes negativas.

Podemos verificar também, nesses trabalhos, a preocupação de se elaborar situações-problema de uma forma organizada, no processo de aprendizagem da Estatística, auxiliando na identificação de possíveis erros por parte dos alunos e compreendendo a origem de tais erros, bem como fazendo um estudo de novas propostas para uma forma correta de abordagem do conteúdo da disciplina.

Uma outra pesquisadora muito importante na área da Didática da Estatística é Carmem Batanero, da Universidade de Granada, Espanha. Investigações realizadas por diversos autores sobre os erros e as dificuldades na compreensão dos conceitos estatísticos elementares são analisados por esta pesquisadora. Batanero(2001a) faz a afirmação em seu trabalho que grande parte da investigação teórica e experimental, realizada atualmente em Didática da Matemática, mostra que os alunos produzem respostas erradas ou simplesmente não são capazes de produzir nenhuma resposta, quando são submetidos a certas tarefas. No caso em que não se trata de distração, os professores acreditam que a tarefa é muito difícil para o aluno, porém tais erros não acontecem aleatoriamente, imprevisíveis. Com freqüência, é possível encontrar regularidades, associações com variáveis próprias das atividades propostas, dos sujeitos ou de circunstâncias presentes ou passadas. Pode-se afirmar que estas regularidades são provocadas pela mobilização de forma estável de conhecimentos em ação.

No caso da probabilidade e da estatística, é importante analisar o raciocínio dos alunos, visto que tratamos com idéias abstratas e não tão ligadas à experiência escolar, como foram os conceitos matemáticos.

Podemos perceber que a natureza da estatística é diferente da cultura determinista da matemática. Os indicadores disso são as controvérsias filosóficas sobre a interpretação e a aplicação de conceitos básicos como os de probabilidade, aleatoriedade, independência ou contraste de hipótese, já que estas controvérsias não são comuns no campo da álgebra ou geometria. Por isso, faz-se necessário o trabalho da Estatística desde as séries iniciais como sugerem os Parâmetros Curriculares Nacionais (PCN) e, quanto mais tarde trabalharmos a formação dos conceitos estatísticos, mais dificuldades os alunos terão para entender a aleatoriedade e a variabilidade contida nos fenômenos do cotidiano, que permita uma leitura correta do mundo.

A formação de professores nesse âmbito especifico é quase inexistente. Só recentemente, segundo Batanero (2001a), o ensino de Estatística foi inserido em alguns cursos de licenciatura das Universidades na Espanha e muitos ainda não o contemplam. No Brasil, quando contemplam, nem sempre é no enfoque crítico. Por outro lado, embora tenhamos excelentes livros textos, a investigação didática está começando a mostrar como alguns erros conceituais e pedagogia inadequada estão presentes com freqüência nesses livros, afirma a autora.

3.3 PERFIL DO PROFESSOR ENTREVISTADO

O professor entrevistado que nos ajudará na pesquisa tem vinte e seis anos de idade, é casado, não tem filho, e foi formado no ano de 2005 em uma universidade particular na zona leste na cidade de São Paulo. Leciona a três anos em uma escola Estadual também da zona leste. Pessoa com boa desenvoltura e vontade de se aprimorar na sua área de conhecimento.

Concordou em nos ajudar em nossa pesquisa não se eximindo e buscando tirar proveito dos resultados que pudessem auxiliá-lo no processo de ensino e aprendizagem no tópico estatística.

3.4 QUESTIONÁRIO DIAGNÓSTICO

Este questionário tem o objetivo de detectar qual o conhecimento que o professor possui quanto ao assunto estatística e poder estimar quais os conceitos que deveríamos trabalhar com este para que obtivéssemos sucesso na aplicação dos testes, pois todo o êxito desta pesquisa depende de como este professor compreende ou venha a compreender os conceitos de base da estatística.

Quando falamos em estatística, qual relação vem à sua mente?

"Vem uma relação de pesquisa para saber, para saber certa quantidade de números, para podermos fazer um trabalho de relacionamento entre a pesquisa e o número previamente dito, entre os solucionados no caso".

Você sente-se seguro o suficiente para trabalhar com este conteúdo?.

"**Não, não tenho uma base forte para trabalhar com este conteúdo, Não estou preparado para dar uma boa aula sobre isto, eu precisaria de um conhecimento maior".**

Entrevistador ainda pergunta: O senhor é formado em uma Faculdade da zona leste recentemente, no ano de 2005 correto?.

"Entrevistado confirma que sim".

Quando chegou à Universidade você já tinha algum contato com esse conteúdo?

"Não, só tive a base da base mesmo, um conhecimento meio fraco mesmo da realidade do que é estatística".

A Universidade que o senhor lhe fez um profissional capaz de desenvolver capacidades de transferência de conhecimento da vida e da experiência do cotidiano para ambiente da sala de aula, em diferentes modelos organizacionais?

"É ela assim, (ela, estava referindo-se à universidade) ela começou a tratar do assunto, relacionado ao cotidiano nosso, tendo exemplo de nós como podemos usar no cotidiano no caso, exemplo de coisas que estão acontecendo na sociedade no momento, foi esta base que ela deu, como podemos dizer, passar o conhecimento para o aluno no caso, usando o conhecimento na base do cotidiano".

O que você pode dizer-me sobre o desenvolvimento de competências para utilização de novas tecnologias a serviço da aprendizagem da Estatística?

"Nunca usei software, não conheço, nunca trabalhei com sistema de informação previamente dito mesmo, nunca trabalhei".

Como você vem lidando com o livro didático quando o assunto é a Estatística, você acredita que só este já basta ou utiliza outros materiais?. Quais seriam os outros materiais?.

"Pela base que eu tenho, só este livro(Novo Ensino Médio Volume Único . Curso Completo. Jorge Danirl Silva , Valter Dos Santos Fernandes, Orlando Donizete Mabelini. IBEP), me dá uma linha de raciocínio para seguir, agora que ele é propriamente dito suficiente para o aprendizado isto não, temos que trabalhar com materiais relacionados ao dia-dia do aluno. Eu não acho que proibindo o livro a gente vai acabar ajudando o aluno, Acho que deixando o aluno usar o livro didático ele poderá aprender melhor o conceito do que queremos passar para ele, Acho que não deveríamos proibir o livro didático para o aluno".

Concluído este questionamento pudemos estimar a partir de onde teríamos que trabalhar com o colaborador, como na questão número dois ele responde que não têm segurança para trabalhar com estatística e que precisaria compreender melhor os conceitos, fizemos algumas intervenções sobre o que este já havia ensinado, e que para o êxito de nossa pesquisa teríamos que ajudar o professor a compreender os conceitos desde os mais simples, que seriam trabalhados em alguns encontros.

A seguir mostraremos o trabalho realizado com o professor durante alguns encontros.

3.5 OS ENCONTROS COM O PROFESSOR

Todo professor que leciona uma disciplina que contempla conteúdos de Estatística, seja no Ensino Fundamental, Médio ou Superior, devera ter clara visão da Estatística como ciência. Isso significa ter conhecimento de como ela se estrutura, as áreas de atuação, a relação com a Matemática e suas aplicações. Será que o professor licenciado em Matemática está preparado para esta tarefa?

Buscaremos responder a esta pergunta e discutir aspectos do ensino de Estatística na escola.

A Estatística é a ciência que se preocupa com a coleta, organização e interpretação de dados experimentais, conforme refere costa Neto (1977). Desta forma alguns conteúdos de Estatística devem ser trabalhados tendo como objetivo desenvolver no aluno esta competência estatística.

De acordo com os PCN´s, os seguintes objetivos e conteúdos para o Ensino Fundamental e Médio relacionados a conteúdos de Estatística devem ser trabalhados nas aulas de Matemática, como nosso foco de estudo principal é o Ensino Médio, teremos:

> Ler e interpretar dados ou informações apresentados em diferentes linguagens e representações, como tabelas, gráficos, esquemas, diagramas, árvores de possibilidades; selecionar diferentes formas para representar um dado ou conjunto de dados e informações, reconhecendo as vantagens e limites de cada uma delas; ler e interpretar distintos tipos de textos com informações apresentadas em linguagem matemática como por exemplo artigos de jornais e revistas; acompanhar e analisar os noticiários e artigos relativos à ciência em diferentes meios de comunicação, interpretando com objetividade seus significados e implicações; expressar-se com clareza utilizando a linguagem matemática elaborando textos, gráficos e tabelas; compreender e emitir juízos próprios sobre informações relativas à ciência e tecnologia, de forma analítica e crítica e posicionando-se de forma analítica e crítica; identificar os dados relevantes em uma dada situação-problema para buscar possíveis resoluções; utilizar as ferramentas matemáticas para analisar situações de seu entorno real e propor soluções; reconhecer o caráter aleatório de fenômenos e eventos naturais, científico-tecnológicos ou sociais; quantificar e fazer previsões em situações aplicadas a diferentes áreas do conhecimento e da vida cotidiana que envolvam o pensamento probabilístico.

Conteúdos: Construção e interpretação de tabelas e gráficos medidas de tendência central (média, mediana e moda); medidas de variabilidade (Variância, desvio-padrão), probabilidade.

Os objetivos propostos pelos PCN´s não envolvem apenas a aprendizagem dos conteúdos limitando-se a como realizar os cálculos, mas em seus propósitos uma grande ênfase na interpretação do significado desses cálculos, bem como na valorização da utilização da estatística, como um meio na análise de dados. Esta idéia vem de encontro à importância de desenvolver nos alunos as competências básicas em estatística. Para Rumsey apud Bayer et al (2002), as competências básicas em Estatística envolveriam cinco aspectos:

Conhecimento do dados;

Entendimento sobre a terminologia e conceitos básicos de Estatística;

Compreensão do básico a sobre coleta de dados e geração de estatísticas descritivas;

Habilidades básicas de interpretação (habilidade para descrever o que os resultados significam no contexto do problema);

Habilidades básicas de comunicação (explanar os resultados a outras pessoas).

A falta de preparação do professor de Matemática para o desenvolvimento dos conteúdos relacionados à Estatística faz com que este, muitas vezes, prefira não trabalhar com estes conteúdos em suas aulas.

Conforme cita (Batanero, 2001 apud Bayer; et al, 2000), devem ser considerados fundamentais para a capacitação do professor, aspectos estes como o conhecimento didático, os conhecimentos dos quais o professor de Matemática deverá apresentar para lecionar estatística.

Vejamos o professor que se propôs a nos ajudar na pesquisa:

No começo de trabalho com a formação desse professor colaborador, foram discutidos representação e interpretação de dados em tabelas, gráficos de barra, linhas e setor. Foi uma discussão informal, mas, buscando fazer com que o professor pudesse compreender de fato o trabalho com estes elementos estatísticos. Sabíamos, segundo o questionário respondido por esse professor, que ele não havia tido uma boa formação sobre a estatística e quase tudo era novo.

A forma utilizada para a apresentação dos conceitos que pretendíamos, foi a mesma utilizada no capítulo “A estatística como objeto de ensino”.

Pretendíamos buscar nosso objetivo com esse professor, fazer com que ele pudesse coletar dados, organizar, analisar e interpretar informações representadas em diversas formas matemáticas, como foi exposto no começo deste capitulo, e assim pudesse fazer com que os seus alunos também compartilhassem do novo aprendizado.

Para cada um de nossos encontros preparatórios foram utilizados em média 2h15.

1º encontro

Neste primeiro encontro, falamos sobre População, amostra e variável estatística, procuramos discutir os temas a serem trabalhados por meio de situações-problema. Seguem abaixo estas situações e o diálogo com o professor durante o processo de resolução.

1ª situação: Uma fabricante de calçados fez uma pesquisa em uma escola primária para saber sobre o desenvolvimento do tamanho dos pés das crianças em uma determinada faixa etária, no começo do ano letivo ele mediu o tamanho dos pés das crianças para poder comparar o crescimento em um ano, a anotações foram feitas assim:

Kátia, 10cm, Márcio 13cm, Ricardo 13cm, Maria 11cm, Renato 14cm, Selma 11cm, Diego 13cm, Sandra 12cm, Sergio 13cm, Fabio 15cm, Roberto 15cm, Osvaldo 13cm, Silvana 12cm, Cintia 11cm, Rosa 12cm. Mais tarde, ele pensou em uma maneira para simplificar essas anotações. Veja como ele fez. Ordenou os dados em ordem crescente e criou uma tabela para poder anotar os resultados.

Tabela 3: Estudo sobre o pé das crianças

Nome	Tamanho do pé inicial (cm)	Tamanho do pé final (cm)
1. Kátia	10	
2. Maria	11	
3. Selma	11	
4. Cintia	11	
5. Sandra	12	
6. Silvana	12	
7. Rosa	12	
8. Márcio	13	
9. Ricardo	13	
10. Diego	13	
11. Sergio	13	
12. Osvaldo	13	
13. Renato	14	
14. Fabio	15	
15. Roberto	15	

Fonte: Arquivo do pesquisador dados Hipotéticos

Após verificar que este professor não sabia como fazer o tratamento dos dados, optou-se por aulas “tradicionais” mas ressaltando que tal método não deveria ser usado com os alunos.

PSQ: É fácil notar que o tamanho do pé de uma criança tem 10cm, de três crianças tem 11cm, três crianças 12cm, cinco crianças 13cm, uma criança 14cm e

duas crianças 15cm. Agora ficou fácil calcular a média do tamanho dos pés das crianças veja:

$$\text{Media aritmética} = \frac{1.(10)+3.(11)+3.(12)+5.(13)+1.(14)+2.(15)}{15}$$

Média aritmética = 12,53.

PROF: é mesmo muito fácil de calcular. Mas acaba assim?

PSQ: o cálculo sim mais precisamos compreender o seu significado.

PROF: como assim?

PSQ: Média é um valor típico ou representativo dos dados. Podemos dizer que a média de uma distribuição é o valor em torno do qual os demais valores oscilam podemos interpretar a média como sendo o ponto de equilíbrio dos desvios da distribuição.

PROF: poderia me explicar melhor?

PSQ: a média aritmética, tem por objetivo dar uma idéia razoável a respeito de um conjunto de valores através de um único valor que possa representa-lo bem.

PROF: você pode dar um exemplo para que eu possa compreender melhor?

PSQ: Digamos que um padeiro venda 2100 pães por mês, isso não significa que todos os dias deste mês este padeiro consiga vender a mesma quantidade de pães, provavelmente, em alguns dias, ele vendeu mais de 70 pães, mais em outros ele vendeu menos, assim, o que importa é que nos dias que vende mais cobre os dias que menos se vende, então podemos dizer que um dia pelo outro, ele vendeu em média 70 pães por dia, é bem mais fácil usar está medida a ter que enumerar as quantidades exatas vendidas em cada dia.

PROF: OK, assim não importa se ele vendeu menos em alguns dias desde que consiga recuperar a diferença em outros dias.

PSQ; isso mesmo. Portanto, para calcular a média aritmética de um conjunto de valores somam-se todos e divide-se o resultado pela quantidade de valores do conjunto.

PROF: gostei.

PSQ: Professor, a intenção aqui não é de lhe ensinar a fazer contas, mais é muito importante que você converse com seus alunos a respeito do uso da calculadora. Neste exercício, que

estamos trabalhando por exemplo, os alunos precisam saber que eles têm algumas opções, vamos falar de algumas delas:
Primeiro opção:
Você pode fazer cada uma das multiplicações, utilizando a calculadora e anotar os resultados em uma folha de rascunho, depois da soma de todas estas anotações obtidas, divida o resultado encontrado pelo total das crianças, que, para nosso exemplo, é o 15.
Segunda opção:
Você poderia falar da tecla de memória da calculadora, explicar que eles fariam as multiplicações e armazenariam os resultados na tecla M+ para a soma e se quisessem subtrair, usariam a tecla M-.
Explique, também, sobre a função de outras teclas muito importantes e que serão muito utilizadas, por exemplo, a tecla ON/CE, aparecerá no visor 0, ou seja, o visor ficou limpo.
A tecla MRC mostra o que está armazenado na memória.
Vejamos para nosso exercício, 1.10(M+) 3.11(M+) 3.12(M+) 5.15(M+) 14(M+) 2.15(M+) (MRC).
No visor da calculadora, aparecerá M 180, isso ocorre em razão do fato de que, quando você apertou a tecla M+, a calculadora passou a fazer cada multiplicação e soma, guardando o valor da soma total.
Agora para poder trabalhar com este valor direto na calculadora, aperte a tecla MRC e aparecerá no visor 188, que é só dividir por 15 e teremos a média.
Pratique com seus alunos estas formas de utilização, elas serão muito utilizadas.
PROF: eu confesso que não sabia trabalhar com a tecla da memória da calculadora.

Foi pedido ao professor que praticasse tudo, o que foi aprendido em sua casa e também que tudo o que ele estivesse vendo, fosse passado paralelamente com seus alunos, já que nossos encontros ocorriam apenas uma vez semanal.

2º encontro

O professor iniciou esse encontro comentando que foi muito interessante ter trabalhado noções de média aritmética com os alunos, embora tivesse de providenciar algumas calculadoras, pois, nem todos os alunos tinham uma e a escola não tinha o suficiente para o empréstimo. Havia colocado os alunos sentados em pequenos grupos onde todos pudessem treinar um pouco o uso das principais teclas, porém, de uma maneira geral, foi muito gratificante e todos demonstraram grande interesse.

Continuamos a ver o mesmo exercício da aula passada.

PSQ: o valor da média 12,53, ainda não diz muita coisa, porque já vimos que existem crianças bem abaixo da média e, também, outras muito acima da média, como por exemplo, Kátia e Roberto, respectivamente, isso mostra uma distorção da informação que a média nos traz.
Temos uma medida que nos faz ter uma visão melhor dos fatos, ela é o Desvio-padrão, é uma medida que mostra o quanto estamos próximos ou longe da média encontrada.
PROF: já ouvi falar mais não sei dizer o que é.
PSQ: podemos observar se o dados são considerados homogêneos, quando todos valores estão próximos da média ou heterogêneo, quando muitos valores estão bem abaixo ou muito acima da média.
PROF: mas como calcular o desvio-padrão?
PSQ: primeiro, iremos criar uma tabela ordenada de forma crescente ou decrescente com o nome das crianças, tamanho do pé e tamanho do pé elevado ao quadrado, faça a soma dos resultados de cada coluna, ficará assim:

Tabela 4: Tamanho do pé (cm^2)

Nome	Tamanho do pé inicial cm	Tamanho do pé cm^2
1. Kátia	10	100
2. Maria	11	121
3. Selma	11	121
4. Cintia	11	121
5. Sandra	12	144
6. Silvana	12	144
7. Rosa	12	144
8. Márcio	13	169
9. Ricardo	13	169
10. Diego	13	169
11. Sergio	13	169
12. Osvaldo	13	169
13. Renato	14	196
14. Fabio	15	225
15. Roberto	15	225
Total	188	2386

Fonte Arquivo do Pesquisador dados hipotéticos.

Como já temos os dados individuais podemos calcular a variância;
S^2=[2386-(1882 / 15)] /14 = (2386-2356,26) / 14 = 2,12cm^2.
O desvio-padrão é uma medida que nos mostra quão próximo ou distante da média estão os valores do conjunto.
PROF: como assim?
PSQ: Desvio-padrão indica o grau de concentração dos valores observados ao redor da média. Para encontrarmos o desvio-padrão basta tirarmos a raiz quadrada do valor encontrado para a variância, assim, o valor do desvio padrão será 1,4574.
PROF: o que se faz agora?
PSQ: como a Média é 12,53 + 1,46 = 13,99 e 12,53 –1,46 = 11,07

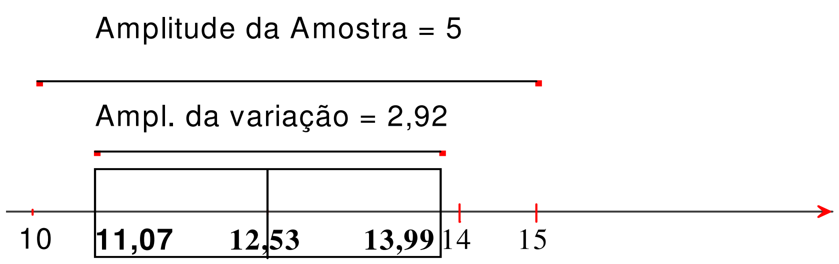


Gráfico 5: Amplitude de variação.

PSQ:Como o desvio somado e subtraído a média apresentam, respectivamente, 13,99 e 11,07, podemos observar que mais de 83% dos valores de nossa tabela encontram-se dentro deste intervalo, então, conclui-se que para nosso exercício a média é uma boa medida de representação, pois os valores são considerados homogêneo.
PROF: nunca tinha visto este tipo de representação, da uma ótima visão deste desvio.

PSQ: um outro tipo de medida conhecida em estatística é a moda é o valor mais típico de uma distribuição de freqüência.
PROF: não entendi.
PSQ: é o valor que mais se repete em um conjunto de dados. No exercício que estamos trabalhando a moda será 13, pois é o valor mais freqüente.
PROF: está medida eu lembro, se não houvesse valores repetindo-se, então, não teríamos moda.
PSQ: também, em outra situação poderíamos ter dois valores com a mesma freqüência, assim, teríamos duas modas, se acontecesse o mesmo com três valores teríamos três modas e, assim, por diante.
PSQ: outra medida muito conhecida é a mediana, é o valor que divide o conjunto de dados ao meio; para nosso caso a mediana será o 13, pois estamos trabalhando com 15 valores e o valor central é 13, deixa sete valores abaixo e mais sete valores acima.
Vejamos um exemplo ilustrativo.

Figura 1: Fonte Cazorla e Santana (2006)

Bia	Ana	Caio	Luiz	João
145 cm	148 cm	150 cm	152 cm	155 cm

Abaixo da mediana: dois dados ←——Mediana——→ Acima da mediana: dois dados

PROF: com esta ilustração ficou fácil para se entender a mediana.

PSQ: note que se trocarmos o João, por uma outra pessoa bem mais alta, a mediana não mudaria, mais se estivéssemos trabalhando com a média, bastaria mudar qualquer das pessoas, que a média também mudaria.
PROF: não vou mais esquecer.
PSQ: os quartis são outro tipo de mediana que é outra medida separatriz.
Assim, os quartis são valores de uma série que a dividem em quatro partes iguais.
Precisamos, portanto, de três quartis (Q1 , Q2 e Q3) para dividir a série em quatro partes iguais.
Q1= limita 25% dos dados da série.
Q2= limita 50% dos dados da série.
Q3= limita 75% dos dados da série.
Obs: O quartil 2 (Q2) sempre será igual à mediana da série.
Para dados não agrupados

Basta calcular mediana aos três quartis. Na realidade, serão calculadas " três medianas " em uma mesma série.
Vejamos para nosso exercício :
{10, 11, 11, 11, 12, 12, 12, **13**, 13, 13, 13, 13, 14, 15,15 }
Neste caso, os dados já estão ordenados, caso contrário, primeiro faríamos a ordenação dos dados (crescente ou decrescente) dos valores:
O valor que divide a série acima em duas partes iguais é **13**, logo a Mediana = 13 que será = segundo quartil, ou seja Q_2.
PROF: mediana também pode ser chamada de quartil 2?.
PSQ: sim, porque é o valor que divide a serie ao meio.
PROF: entendi.
PSQ: temos agora {10, 11, 11, 11, 12, 12, 12 } {13, 13, 13, 13, 14, 15, 15 } como sendo os dois grupos de valores iguais que ficaram abaixo da mediana e acima da mediana (quartil 2). Para o cálculo do quartil 1 e 3 basta calcular da mesma forma as medianas das partes iguais provenientes da verdadeira Mediana da série (quartil 2).
Assim em { 10, 11, 11, **11**, 12, 12, 12 } a mediana é **11** , o quartil será 1º, Q1.
Em {13, 13, 13,**13**, 14, 15, 15 } a mediana é **13** , o quartil será 3º, Q3.
No final, o professor afirmou que iria continuar trabalhando todo esse conteúdo com seus alunos.

3º encontro

Iniciamos o nosso encontro com o relato do professor de que ele estava gostando muito de trabalhar estes conceitos com seus alunos pois eles demostravam bastante interesse em participar das aulas, e quanto a ele professor pela primeira vez estaria sabendo como conduzir as discussões a respeito da estatística, alguns alunos tiveram dúvidas sobre o arredondamento da média, mais eu respondi que, para o caso da média, não se deveria arredondar pois isto não é prudente, dei o exemplo, de que se estivéssemos tratando com dinheiro e a média fosse 35,55, se arredondássemos para 36, então a média deixaria de ser uma medida resumo, pois ela seria desvirtuada e deixaria de ser o pondo equilíbrio desta distribuição.

PSQ: muito bom!
PSQ: vejamos agora outra atividade
Em um supermercado, a distribuição dos salários dos funcionários é feita como nos mostram os dados da tabela abaixo:

Tabela 5: Relação funcionário e Salário

Nº funcionários	Salários em R$
90	330,00
20	420,00
11	590,00
8	710,00
5	1.100,00
2	3.600,00

Calcule a média aritmética dos salários.
Calcule a mediana.
Determine a moda.
Qual medida você acredita que representa melhor o salários dos funcionários desse supermercado,(média, moda ou mediana)? Por quê?
Se uma comissão dos funcionários resolvesse reivindicar uma renegociação de seus salários, qual das três medidas seria melhor empregada? Por quê?
O professor resolveu a questão proposta mostrando que havia entendido os conceitos e, além disso, estaria fazendo um estudo paralelo às nossas aulas.

4º encontro

PSQ: partiremos, então, para a utilização das fórmulas que irão nos ajudar, pois nos darão agilidade na resolução das atividades, é importante que se guarde que os conceitos são os mesmos, o que modifica são as ferramentas de resolução.
Vejamos uma questão proposta por Coutinho (2006):

Um professor deseja estudar a relação entre tempo de estudo de seus alunos e o resultado da avaliação. Para isso, antes do início da prova ele anotará quanto tempo seus alunos declaram ter se preparado para a mesma. Os resultados estão na tabela abaixo. Faça uma análise dos dados, interpretando a informação neles contida.

Tabela 6: Tempo de estudo, em horas, de 25 alunos.

Tempo de estudo, em horas, de 25 alunos do professor xx, para a primeira prova.				
42,3	11,4	26,3	30,1	43,4
60,3	37,7	16,7	38,5	21,4
43,3	49,3	38,8	39,9	55,4
31,9	44,1	53,0	33,8	25,2
34,9	9,0	33,8	38,3	17,7

Fonte Coutinho

Para uma maior visualização dos dados, começaremos colocando os dados em ordem crescente ou decrescente, assim:

Tabela 7 organizada: Tempo de estudo, em horas, de 25 alunos

Dados em ordem crescente				
9,0	25,2	33,8	38,8	44,1
11,4	26,3	34,9	39,9	49,3
16,7	30,1	37,7	42,3	53,0
17,7	31,9	38,3	43,3	55,4
21,4	33,8	38,5	43,4	60,3

Fonte Coutinho

PSQ:Para resolver utilizando esta nova ferramenta, precisamos entender algumas nomenclaturas, vejamos:
n = total de elementos da amostra
fi = freqüência simples
xi = ponto médio das classes
fa = Freqüência acumulada
(At)= Amplitude total (maior valor observado – menor valor observado) At = 60,3 – 9,0 = 51,3
$\bar{x}$ = média
K = número de classes
K = √n, para n = 25
K = √25 = 5, assim, teremos cinco classes
hc = Amplitude do intervalo da classe
hc = At/K
hc= (60,3 – 9) / 5 = 10,2
Muitas vezes, ao efetuar a divisão acima, podemos chegar a um resultado não muito conveniente devido à formatação dos dados observados. (valores quebrados). Neste caso seria conveniente arredondar o valor para adequá-lo ao cálculo, no caso de nossa divisão, arredondaremos 10,2 para 10. hc= 10.
Assim, passaremos à montagem de uma tabela de distribuição de freqüência.
PROF: será que consigo lembrar de tudo isso?
PSQ: todas estas nomenclaturas é só usar que elas serão memorizadas automaticamente.

Tabela 8. Distribuição de freqüência: variável tempo de estudo.

Classes	int. de classe	fi	fa	xi	xi.fi	$(xi - \bar{x})^2.fi$
1	9 ⊢ 19	4	4	14	56	2.007,04
2	19 ⊢ 29	3	7	24	72	461,28
3	29 ⊢ 39	8	15	34	272	46,08
4	39 ⊢ 49	5	20	44	220	288,80
5	49 ⊢ 59	3	23	54	162	929,28
6	59 ⊢ 69	2	25	64	128	1.523,52
∑		25			910	5.256

$\bar{x} = \sum(xi.fi) / n$

$\bar{x} = 910 / 25 = 36,4$

Média = 36,4

Mediana = (Md)

$$Md = li + \left[\frac{\frac{i\ .\ n}{100} - faa}{fi} \right] .\ hc$$

Onde :

li = limite inferior

faa = freqüência acumulada anterior

$Md =$ para P50, pois a mediana é a medida que deixa 50% dos dados, então teremos:

$\left[\frac{i.n}{100}\right] \Rightarrow \left[\frac{50.25}{100}\right] = \frac{1250}{100} = 12,5$ assim observamos no fa para achar a qual classe pertence, no nosso caso 3ª classe.

Assim, para nos a mediana será encontrada, fazendo as seguintes substituições:

$$\mathrm{Md} = 29 + \left[\frac{12,5 - 7}{8}\right].10$$

$$\mathrm{Md} = 35,875 \cong 35,8 \text{ horas}$$

Moda (Mo) = valor de maior freqüência

$$Mo = \mathrm{li} + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right).\mathrm{hc} \Rightarrow \mathrm{Mo} = 29 + \left(\frac{5}{8}\right).10 \Rightarrow \mathrm{Mo} = 35,25$$

Onde: $\Delta_1 = 8 - 3 = 5$ e $\Delta_2 = 8 - 5 = 3$

Cálculo das medidas separatrizes

1º quartil (Q1): deixa 25% dos dados

3º quartil (Q3): deixa 75% dos dados

$$Q_1 = \mathrm{li} + \left[\frac{\frac{i.n}{100} - \mathrm{faa}}{\mathrm{fi}}\right].\mathrm{hc} \Rightarrow \mathrm{Q}_1 = 19 + \left[\frac{6,25 - 4}{3}\right].10$$

$$Q_1 = 26,5 \text{ horas}$$

$$Q_3 = \mathrm{li} + \left[\frac{\frac{\mathrm{i\,.\,n}}{100} - \mathrm{faa}}{\mathrm{fi}} \right] . \mathrm{hc} \Rightarrow Q_3 = 39 + \left[\frac{18{,}75 - 15}{5} \right] .10$$

$$Q_3 = 46{,}5 horas$$

Cálculo das medidas de dispersão

Variância (S2) e Desvio-padrão (S): mostra a dispersão dos dados em torno da média.

$$\text{Variância } S^2 = \sum \left(\frac{x_i - \overline{x}}{n-1} \right)^2 . \mathrm{fi} \Rightarrow S^2 = \frac{5256}{25-1} \Rightarrow S^2 = \frac{5256}{24} = 219$$

$$Desvio - padrão\ S = \sqrt{219} \cong 14{,}79 horas$$

Interpretação dos resultados

PROF: agora é que eu quero ver?
PSQ: os valores se concentram em torno da média, como a média é igual 36,4, então o tempo médio de estudo é de 36,4 horas.
Como a mediana é aproximadamente 35,8, então 50% dos alunos estudam aproximadamente até 35,8 horas.
Como o 1º quartil é igual a 26,5, então 25% dos alunos têm tempo de estudo menor ou igual a 26,5 horas.
Como o 3º quartil è igual a 46,5, então 75% dos alunos têm tempo de estudo menor ou igual a 46,5 horas.
Pela moda, verifica-se que o tempo mais freqüente de estudo é de aproximadamente 35 horas.
Vejamos agora a interpretação no Box-plot
PROF: Box-plot?, este nós ainda não estudamos, o quê é o box-plot?
PSQ: é um tipo de gráfico, vejamos: Traça-se dois retângulos, um representando os espaço entre o quartil inferior e a mediana e outro entre a mediana e o quartil superior. Esses dois retângulos, em conjunto, representam a faixa dos cinqüenta por cento dos valores mais típicos da distribuição. Entre os quartis e os extremos traça-se uma linha. Casos existam valores discrepantes – além de 1,5 (dQ) - ,a linha é traçada até o ultimo valor não discrepante; e os valores discrepantes são indicados por pontos.

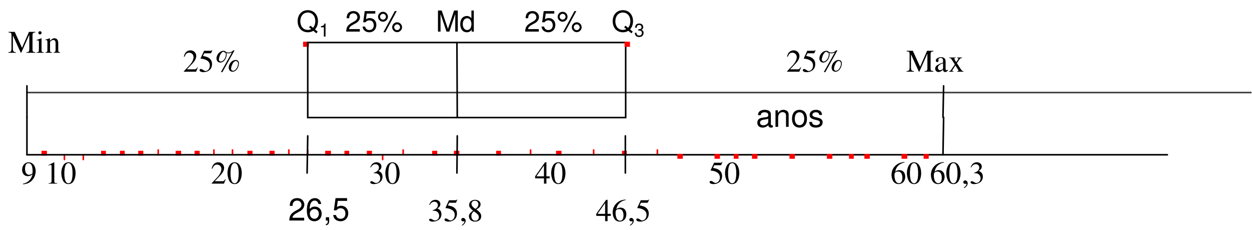


Gráfico 6: Box-plot Variável tempo de estudo

Min: Menor valor do conjunto de dados (mínimo);

Q1: Quartil Inferior (Primeiro quartil);
Md: Mediana;
Q3: Quartil superior (Terceiro Quartil);
Máx: Maior valor do conjunto de dados (máximo)
dQ: Diferença entre o quartil superior e o quartil inferior.
Pelo gráfico box-plot observamos que os tempos de estudo estão mais concentrados entre Q1 e Q3.
PROF: nunca tinha visto este tipo de gráfico.
PSQ: ele nos da uma boa visão da variabilidade.
PROF: tem outro?
PSQ: Histograma de freqüência da variável tempo de estudo. Vejamos.

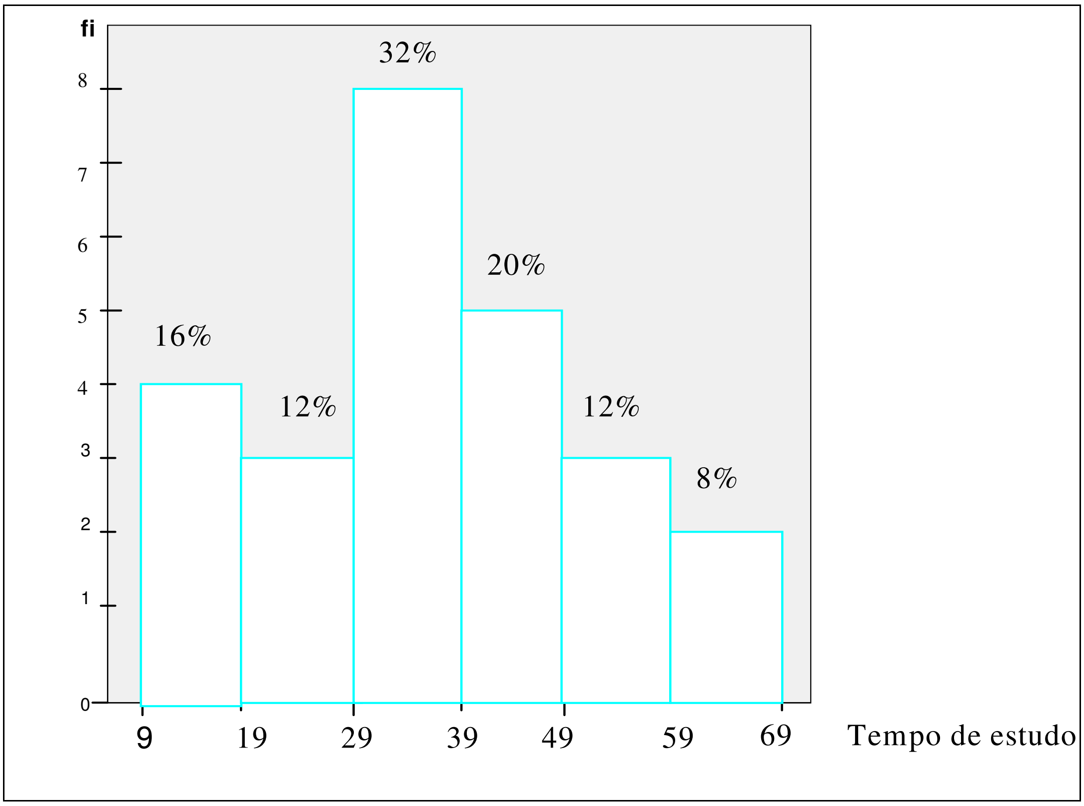


Gráfico 7: Histograma de freqüência da variável tempo de estudo.

PSQ:O intervalo de 9 à 19 mostrado na primeira coluna, têm 10 unidades se multiplicarmos pela a altura fi, então teremos 10 . 4 = 40.

Faremos o mesmo cálculo para todas as outras colunas.

Somando todos os resultados obtidos teremos então: 250.

Podemos então montar uma simples regra de três e encontrarmos qual o percentual de estudo indicado por cada coluna, vejamos na primeira coluna:

250 100 logo teremos: 250 X = 40 . 100

40 x 250 X = 4000

X = 4000/250

X = 16, logo a primeira coluna corresponde a 16%.

PROF: o procedimento para as outras colunas também é o mesmo.
PSQ: sim, assim teremos:
1ª coluna 16%
2ª coluna 12%
3ª coluna 32%
4ª coluna 20%
5ª coluna 12%
6ª coluna 8%, totalizando assim 100% dos dados.

Foi dito ao professor que resolvesse em sua casa mais algumas atividades sugeridas que seria o foco das discussões da nossa próxima aula.

5º encontro

Fizemos as discussões prometidas e principalmente sobre as análises dos resultados.

Entre as questões que o professor havia respondido em casa, tínhamos, por exemplo:

Explique de forma clara e sistemática como podemos relacionar os conceitos estatísticos de base para analisar um conjunto de dados quantitativos.

Pela resposta dada pelo professor, notamos que havia sim dificuldades, pois ele descreveu os passos e não as relações de variabilidade em torno da média e da mediana.

Veja a resposta do professor a esta questão:

> PROF: Conceitos estatísticos de base:
> Conceito de amostra e população
> Conceito de variável
> Organização dos dados numa tabela de distribuição de freqüência
> Saber conceituar e calcular as medidas de posição: média, moda e mediana
> Medidas de dispersão: variância e desvio-padrão
> Estudo da variabilidade em torno da média
> Saber conceituar e calcular as medidas separatrizes: 1º quartil, 2º quartil e 3º quartil
> Estudo da variabilidade em torno da mediana
> Representação gráfica
> Num segundo momento devemos fazer a interpretação dos resultados.

Quanto ao conhecimento técnico, o professor mostrou que já estava muito bem, mas necessitava ainda de aprofundar-se mais em como relacionar tais conteúdos e foi o foco principal desta nossa aula.

Combinamos um dia para que eu fosse a sala de aula do professor para conhecer seus alunos e ver como as aulas estavam sendo ministradas, o professor achou melhor que fosse especificamente à aula em que ele fosse tratar das análises dos resultados, assim foi feito.

No dia marcado pude observar que havia sim grande empenho dos alunos em resolver as atividades. Os alunos sentavam em pequenos grupos e não tinham dificuldades para os cálculos, mais que ainda não sabiam como analisa-los.

O professor pegou um problema especifico para começar este estudo. Precisei fazer algumas intervenções e pude observar que havia ocorrido um grande desenvolvimento daquele professor e que agora poderíamos começar nossos testes.

O teste foi marcado para as aulas seguintes. Mas no dia o professor faltou e com autorização da coordenação da escola, o teste prosseguiu mesmo assim. Precisei levar algumas calculadoras, e o teste foi todo gravado em áudio.

Nas aulas seguintes o professor continuou faltando, mas o teste prosseguiu. Fui informado pela coordenação da escola que o professor havia sofrido um acidente com o seu carro mas que passava bem, e que eu poderia informar aos seus alunos que as atividades seriam utilizadas como um dos instrumentos de avaliação bimestral.

4 O TESTE OFERECIDO AOS ALUNOS

Na disciplina Estatística, para alunos da Escola de Ensino Médio, levanta-se a hipótese da Estatística ser bastante complexa, agravada, pela multiplicidade de fatores que a envolvem por exemplo professores mal preparados para lidar com os conceitos, e ainda segundo Vendramine (2000), os alunos apresentam atitudes negativas em relação a Estatística por relaciona-la a Matemática.

Na atividade proposta, os alunos irão analisar uma situação-problema fictícia, adaptada do trabalho de Bifi (2006), e também utilizadas nos dois outros trabalhos que compõem este projeto: VASQUES, Ricardo Sergio Braga; CARDOSO, Ricardo, já citados anteriormente. Nosso objetivo o de verificar se, uma vez que o professor recebe uma "assessoria" para o trabalho com temas ligados à Estatística, ele pode trabalhar adequadamente com seus alunos, favorecendo estes em sua aprendizagem.

4.1 ATIVIDADES PROPOSTAS

Primeira parte

1º) Uma Empresa de cartões de crédito solicitou uma análise do banco de dados abaixo, construído baseado nas respostas a um questionário que buscava levantar a idade e a renda mensal de 40 pessoas.

Tabela 9: Idade e Renda Mensal 40pessoas

Idade	Renda Mensal	Idade	Renda Mensal	Idade	Renda Mensal	Idade
30	1.180	28	1.420	37	387	40
28	490	46	630	29	1.600	25
28	1.200	30	1.000	43	1.770	30
40	540	31	760	43	1.770	45
29	860	23	1.000	31	1.200	31
31	850	29	700	30	1.200	65
30	500	27	400	30	400	53
32	1.600	48	380	30	1.400	25
41	700	30	1.800	30	1.400	34
39	1.420	40	554	28	800	25

Fonte Bifi (2006)

Questões:

Nas variáveis idade e renda mensal, encontre a média e o desvio-padrão. Como você analisaria esses resultados?

Nas variáveis, idade e renda mensal, encontre a mediana, o 1º quartil e o 3º quartil. Como você analisaria estes resultados?

Se você precisasse explicar o comportamento da variável idade para um cliente, você usaria o item (1) ou o (2)? Explique.

Segunda parte

De acordo com as tabelas abaixo, responda às questões:

Tabela 10: Quantidade de carros/pessoa (dados fictícios)

Quantidade de carros	***Número de pessoas***
1	10
2	25
3	15
4	5
Total	***55***

Fonte Bifi(2006).

Tabela 11: N. de horas no trânsito/pessoa (dados fictício)

Tempo no trânsito	***Número de pessoas***
0 ├ 2	10
2 ├ 4	20
4 ├ 6	30
6 ├ 8	15
Total	***75***

Fonte Bifi(2006).

Questões

Nas tabelas acima, determine a média e o desvio-padrão. Como você analisaria estes resultados?

Nas tabelas acima, determine a mediana, o 1º quartil e o 3º quartil. Como você analisaria estes resultados?

Se você precisasse descrever os dados Quantidade de carros e Tempo no trânsito para um cliente, você usaria o item (1) ou o (2)? Explique por quê.

Terceira parte

Observe os gráficos abaixo. Responda à pergunta: “Se você precisasse descrever esses dados para um cliente, como você os analisaria?”.

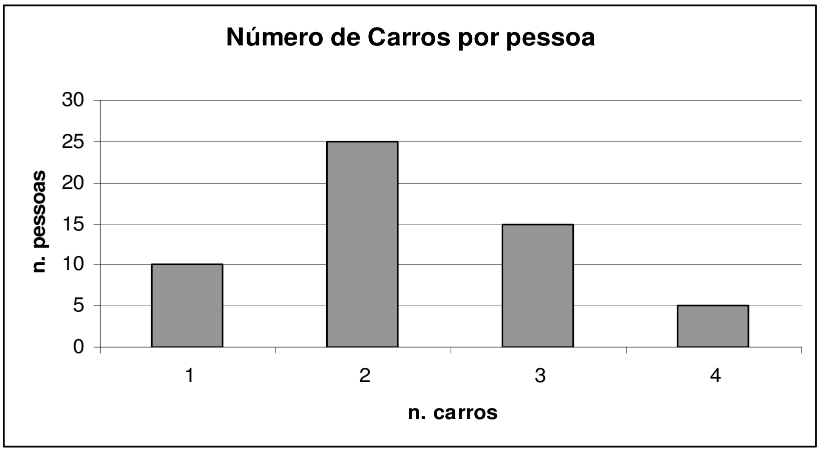


Gráfico 8: Número de carros/pessoa.
Fonte Bifi(2006).

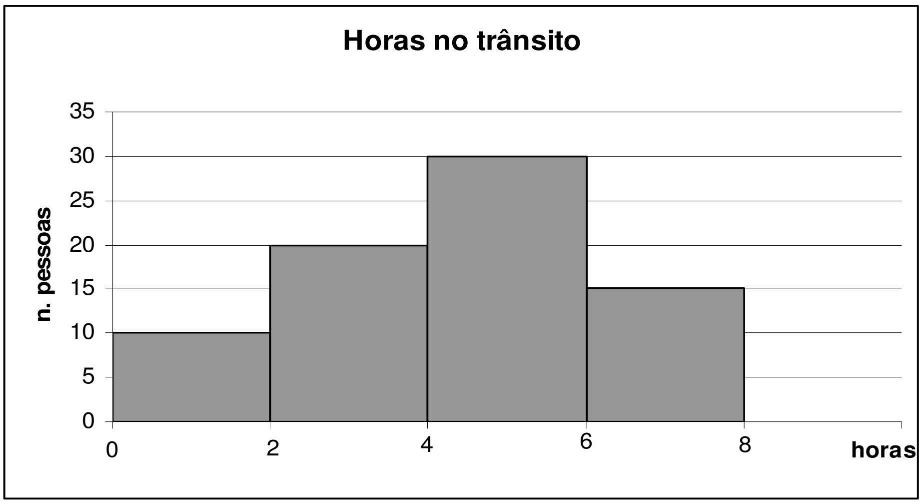


Gráfico 9: Horas no trânsito/pessoa.

Fonte Bifi(2006).

4.2 ALGUMAS CONSIDERAÇÕES

Para estas tarefas:

• O aluno deverá saber que para uma boa análise de dados o primeiro passo deverá ser conhecer a População ou amostra e observar se será necessário trabalhar com todos os elementos desta população ou apenas com parte dela, (amostra), deverá perceber o que esta escolha influirá no resultado final.

• O aluno deverá conhecer a amplitude da amostra, ela poderá sinalizar que tipo de tratamento deverão receber as variáveis quantitativas. Isto se resume a saber dar um tratamento as primeiras informações para otimizar o trabalho de análise desse conjunto de dados. Amplitude da amostra (diferença entre maior e menor valor dos elementos da amostra).

• Podemos observar que a forma de apresentação dos dados de nossas atividades sugere um tratamento unidimensional das variáveis envolvidas nos problemas. Embora a forma como se apresenta a primeira atividade é bidimensional. Fica a livre escolha do aluno a preferência da ordem da variável a tratar.

• Na observação de dados coletados, há sempre um objetivo à vista. Exploramos estes dados e observamos o que eles nos revelam. Os alunos que serão submetidos ao teste, aprenderam com o professor de matemática algumas técnicas de resolução. Não será aqui observado a técnica utilizada pelo aluno e sim a forma como ele conseguirá relacionar (1) natureza ou forma de distribuição; (2) um valor representativo; (3) uma medida de variação.

• É de fundamental importância levar em consideração a distribuição dos dados, estes poderão afetar não apenas a metodologia adotada mais também as conclusões finais.

• Os alunos envolvidos na pesquisa precisam adquirir certa familiaridade com conceitos básicos e a apresentação dos dados que são usados para divulgar resultados ao público-alvo. Dois tipos-chave de conceitos, cuja centralização é notável por muitos são “por cento” (Parker; Leinhardt, 1995, p. 435 apud Bifi 2006) e medidas de tendência central, sobretudo, a média e a mediana.

• Os alunos precisam estar atentos à possibilidade de diferentes erros ou preconceitos (em amostragem, medida, conclusão) e manterem uma preocupação saudável quanto á estabilidade e generalidade dos resultados.

• Finalmente, é importante estar atento às diferenças encontradas ou as tendências que podem existir, mas não, necessariamente, podem ser grandes ou estáveis o bastante para serem consideradas importantes ou podem ser causadas por processos de casualidade.

4.3 RESOLUÇÃO DA ATIVIDADE

Primeira parte

De acordo com o enunciado:

Uma Empresa de cartões de crédito solicitou uma análise do banco de dados abaixo, construído com base nas respostas um questionário que buscava levantar a idade e a renda mensal de 40 pessoas.

Solicitaremos aos alunos para realizar uma análise exploratória dos dados apresentados na Tabela 6, com o objetivo, conforme já especificado, de diagnosticar em que nível de conceitualização, segundo Robert (1998), os alunos se encontram em relação aos conceitos estatísticos de base, já vistos no curso.

A resolução da atividade será feita por meio de tabela unidimensional, ou seja, trataremos cada variável separadamente.

Variável idade (ROL):

Tabela 12: Variável idade.

23	25	25	25	27
28	28	28	28	29
29	29	30	30	30
30	30	30	30	30
30	31	31	31	31
32	34	37	39	40
40	40	41	43	43
45	46	48	53	65

Fonte Bifi(2006)

1ª Estratégia de resolução: Cálculo das medidas com base no ROL

Para o cálculo da média, usaremos a fórmula: $\bar{x} = \sum_{i}^{n} \frac{x_i}{n}.fi$, neste caso, o valor da média será: $\bar{x} = \frac{1364}{40} \cong 35{,}075\, anos$;

A variância será denotada pela fórmula: $s^2 = \frac{\sum (x - \bar{x})^2 .fi}{n-1}$; logo, o valor da variância será $s^2 = \frac{2694}{39} = 69{,}076 \cong 69$;

O desvio-padrão será a raiz quadrada da variância: $s = \sqrt{69} \cong 8{,}31$;

O valor mínimo da amostra: 23 anos;

O valor máximo da amostra: 65 anos;

Amplitude da amostra: 42 anos;

Nas medidas separatrizes, temos:

O primeiro quartil (Q1): 29 anos;

A mediana (md): 30 anos;

O terceiro quartil (Q3): 40 anos;

A moda, o elemento que aparece com mais freqüência dentro desta amostra será: 30 anos.

A análise esperada pelos alunos:

Na amostra, a média é de 35,075 anos aproximadamente e o desvio–padrão de 8,31 anos, tendo uma amplitude amostral de 42 anos. Percebemos que o coeficiente de variação em torno da média é de aproximadamente 25%, considerada alta. A amplitude em torno da média é de 16,62 anos, pode ser representada como mostra o gráfico 10:

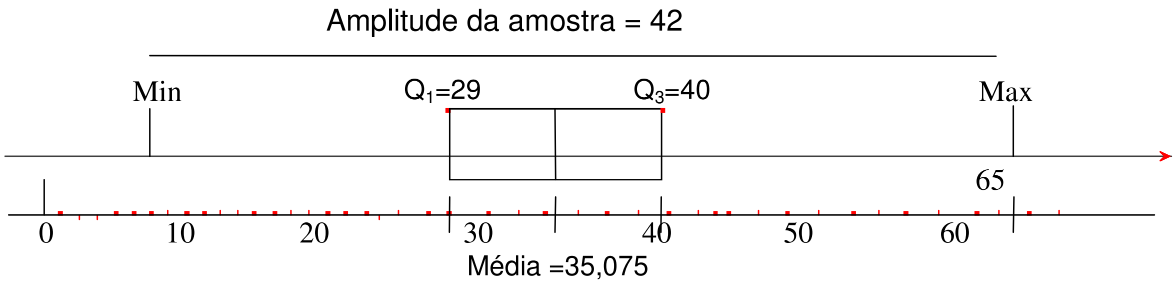


Gráfico 10: Amplitude em torno da média

Pelo gráfico 10, Percebemos em torno da média, uma concentração que varia entre 26,69 anos a 43,31 anos em torno de uma média de aproximadamente 35 anos, representando 39,57% do total da amostra. Na amostra, de cada 100 pessoas entrevistadas, 39,57% estão entre 26,69 e 43,31 anos de idade. Uma análise geométrica dos dados coletados apresenta uma melhor noção do comportamento da amostra. As concepções erradas sobre média apresentam-se de várias formas pelos alunos do Ensino Médio. Entre elas, está a concepção de que média é atrelada ao ponto central, ou seja, não há por parte deles a preocupação de uma análise, a priori, do tipo de distribuição que estão manipulando, se são simétricas ou assimétricas.

> Batanero (2001, p. 87) observa que a média tende a situar o centro dos dados da distribuição, propriedade que é certa para distribuições simétricas. Quando a distribuição é muito assimétrica, a média é desprezada e a moda e a mediana serão os valores mais representativos dos dados. É necessário que o aluno do Ensino Médio faça essa distinção do tipo de distribuição (simétrica ou assimétrica), para, assim, realizar uma análise correta dos dados coletados com a escolha correta do valor mais

representativo para o estudo da variabilidade. Por exemplo, os quartis.

É o que pretendemos mostrar nos próximos passos.

A análise esperada envolve as medidas de posição: Quartis.

Colocando os dados em ordem crescente, encontramos o primeiro quartil, a mediana e o terceiro quartil. Vejamos:

Tabela 13: Cálculo das medidas resumo

23	25	25	25	27	28	28	28
28	**29**	29	29	30	30	30	30
30	30	30	**30**	30	31	31	31
31	32	34	37	39	**40**	40	40
41	43	43	45	46	48	53	65

Fonte: Arquivo do pesquisador

Como a mediana é o que deixa 50% dos dados acima e abaixo, então basta pegarmos os quarenta dados e dividir por dois, assim:

40 / 2 = 20, ai então é só contar vinte termos a partir do primeiro e teremos o 30, que é a mediana e que está destacado no meio da tabela acima.

Com o primeiro quartil é o valor que deixa 25% dos dados acima e abaixo, então basta dividir os dados por quatro, assim teremos:

40 / 4 = 10, ai então é só contar 10 termos a partir do primeiro e teremos o 29, que é o primeiro quartil e que está destacado na tabela acima em primeiro lugar.

Para o terceiro quartil, basta pegar os 50% dos dados que ficaram acima da mediana e dividir por dois, então você novamente encontrará uma nova mediana mais que desta vez representará o terceiro quartil, assim:

20 / 2 = 10, então, basta contar 10 termos a partir da mediana e você encontrará o 40, que representará o terceiro quartil e está destacado na tabela acima na terceira posição.

Tabela 10 Medidas de posição A variável idade tem uma melhor distribuição quando as medidas de posição são tratadas. Logo, para esta análise, a melhor medida que explicaria o comportamento da variável idade seria a mediana. Verificaremos isso por meio do Box-plot.

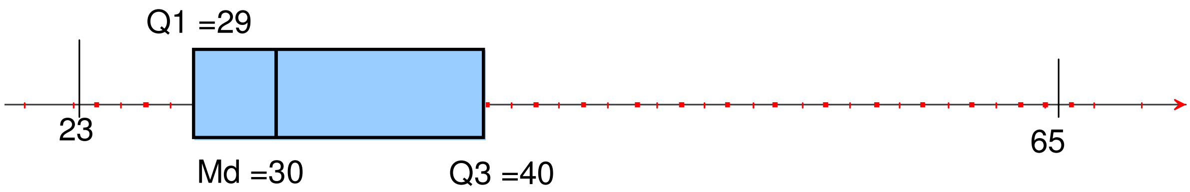


Gráfico 11: Box-plot

Fonte Bifi(2006)

Pelo box-plot (gráfico 11), podemos analisar que, para cada quartil, obteremos 25% da amostra. Nestes moldes, existe maior concentração dos dados entre 23 e 30 anos, de idade, isto é, 50% da amostra estão exatamente entre 23 e 30 anos, esperamos que o aluno do Ensino Médio tenha essa percepção de que a análise desta variável é mais representativa por meio da mediana.

Na análise dos níveis de mobilização citada por Robert (1998), o aluno, em um nível técnico, não relacionará as medidas encontradas e, provavelmente, nem fazer uma análise dos dados por meio das representações geométricas que acabamos de mostrar.

Caso aconteça o inverso, ou seja, se a análise relacional entre as medidas ocorrer e o aluno conseguir dar uma conclusão coerente dos resultados, segundo Robert (1998) este aluno estará em um nível disponível. Por outro lado, se esta análise ocorrer por intervenção do professor, no sentido de mostrar o caminho a ser percorrido e não fornecer a resposta, este aluno estará em um nível mobilizável (ROBERT, 1998).

2ª Estratégia de resolução: Cálculo das medidas por meio de uma tabela de distribuição de freqüência

De acordo com a Tabela de Distribuição de Freqüência com intervalo de classes, calcularemos as medidas de variação e separatrizes:

Para obter o número de classes, utilizaremos a seguinte “propriedade” ; o número de classes é aproximadamente igual ao valor $\sqrt{n}$.

Sabendo que a amplitude da amostra é de 42 anos, usaremos a seguinte regra:

Número de classes: $hc = \frac{AT}{K}$, (AT é a amplitude da amostra e ***n*** é o número de elementos da amostra). Então temos: $hc = \frac{42}{\sqrt{n}} \Rightarrow hc = \frac{42}{6{,}32} \Rightarrow hc = 6{,}64$.

Consideremos seis classes com intervalo de sete anos para cada classe

Tabela 14: Distribuição de freq. Com intervalo de classe: Idade

	f_i	f_a	X_i	$X_i.f_i$	$(x_i-\bar{x})^2.f_i$
23 ├ 30	12	12	26,5	318	867
30 ├ 37	15	27	33,5	502,5	33,75
37 ├ 44	8	35	40,5	324	242
44 ├ 51	3	38	47,5	142,5	468,75
51 ├ 58	1	39	54,5	54,5	380,25
58 ├ 65	1	40	61,5	61,5	702,25
Totais	***40***			***1403***	***2694***

Fonte: Arquivo do Pesquisador

Cálculo das medidas de variação

Calculo da média: $\bar{x} = \sum_{i}^{n} \frac{x_i.f_i}{n} \Rightarrow \bar{x} = \sum_{i}^{40} \frac{1403}{40} \cong 35{,}075$ anos

Cálculo do desvio–padrão: $s = \sqrt{\frac{\sum (x-\bar{x})^2.f_i}{n-1}} \Rightarrow s = \sqrt{\frac{2694}{39}} \cong 8{,}31$ anos.

Na representação geométrica dos cálculos da média e desvio-padrão, podemos, observar o seguinte:

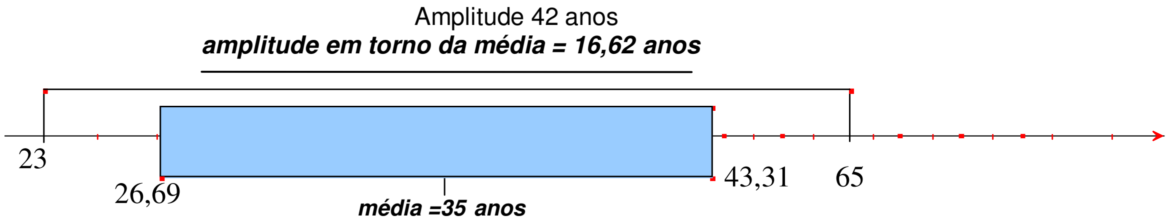


Gráfico 12: Amplitude em torno da média. Fonte Própria

No gráfico 12, a representação aponta para uma não-simetria na distribuição dos dados, com uma concentração que varia entre 26,69 a 43,31 anos em torno de uma média de aproximadamente 35 anos, representando 39,57% do total da amostra. Isso significa que na amostra de cada cem pessoas entrevistadas, podemos esperar que 39 estejam entre 27 e 43 anos de idade. Neste caso, observamos, por meio da representação geométrica que a distribuição é assimétrica, sendo assim, sugerimos que uma melhor análise poderia ser feita pelas medidas separatrizes e não pela associação entre média e desvio-padrão.

Cálculo das medidas separatrizes:

Para os cálculos dos quartis, utilizaremos um dos métodos apresentados aos alunos:

$$Pi = Li + \frac{\left[\frac{i.n}{100} - Faa\right]}{Fi}.hc$$

Pi = indicador do quartil que pretende-se encontrar, por exemplo, se queremos encontrar o primeiro quartil, então teremos P25, pois refere-se a 25% dos valores do conjunto. Se queremos encontrar a mediana, então teremos P50, pois refere-se a 50% dos valores do conjunto. E se queremos encontrar o terceiro quartil, então teremos P75, pois estaremos com 75% dos valores do conjunto.

Li = Limite inferior da classe em que estiver localizado o quartil pretendido.

i = o percentual referente ao quartil pretendido

N = número de elementos do conjunto

Fi = Freqüência simples

Hc = Amplitude da classe

Faa = Freqüência acumulada anterior

Para calcularmos o primeiro quartil ficará assim:

Tabela 15: Quartis

	f_i	f_a
23 ├ 30	12	12
30 ├ 37	15	27
37 ├ 44	8	35
44 ├ 51	3	38
51 ├ 58	1	39
58 ├ 65	1	40
Totais	***40***	

$$\begin{matrix} 30 & mediana & 37 \\ 12 & 20 & 27 \end{matrix} \quad \Rightarrow \quad \frac{med-30}{8}=\frac{7}{8} \Rightarrow Md=\frac{30+7.8}{15} \quad \Rightarrow$$

Obs : *quando* resolver $\frac{\text{i.n}}{100}$, o resultado obtido tem que ser observado no Fa, para poder localizar a que classe pertence o resultado, pois esta também é a classe onde estará localizado o quartil que desejamos. Para nosso caso teremos $\frac{50.40}{100}=20$, no Fa da tabela teremos, segunda classe pois o 20 está no intervalo entre 12 e 27, portanto 30 é o limite inferior e também a classe onde está localizado o segundo quartil.

$$mediana = 33{,}73\ anos$$

$$\begin{matrix} 23 & Q_1 & 30 \\ 0 & 10 & 12 \end{matrix} \quad \frac{Q_1-23}{10-0}=\frac{30-23}{12-0} \Rightarrow Q_1=\frac{23+7.10}{12}=28{,}83$$

Obs : *quando* resolver $\frac{\text{i.n}}{100}$, o resultado obtido tem que ser observado no Fa, para poder localizar a que classe pertence o resultado, pois esta também é a classe onde estará localizado o quartil que desejamos. Para nosso caso teremos $\frac{25.40}{100}=10$, no Fa da tabela teremos, primeira classe pois o 10 está no intervalo entre 10 e 12, portanto 23 é o limite inferior e também a classe onde está localizado o primeiro quartil.

37	Q_3	44
27	30	35

$$\frac{Q_3 - 37}{3} = \frac{7}{3} \Rightarrow Q_3 = \frac{37 + 7.3}{8} = 39{,}62$$

Obs : *quando* resolver $\frac{i.n}{100}$, o resultado obtido tem que ser observado no Fa, para poder localizar a que classe pertence o resultado, pois esta também é a classe onde estará localizado o quartil que desejamos. Para nosso caso teremos $\frac{75.40}{100} = 30$, no Fa da tabela teremos, terceira classe pois o 30 está no intervalo entre 27 e 35, portanto 37 é o limite inferior e também a classe onde está localizado o terceiro quartil.

Os dados calculados expressos no Box-plot trarão a representação geométrica da distribuição. Veja:

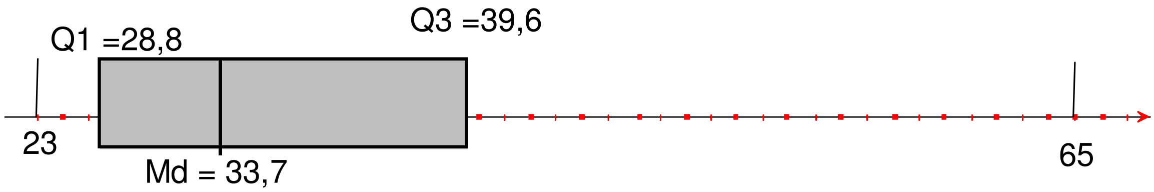


Gráfico 13: Representação dos quartis.

Fonte Bifi (2006)

Pelo Box-plot (Gráfico 13), verificamos que na análise para cada quartil obteremos 25% da amostra. Entretanto, verificamos que os resultados obtidos mostram uma estimativa do comportamento da amostra, que se diferencia de sua representação, quando são trabalhados na primeira estratégia. A escolha das estratégias na análise em questão será a critério do aluno, mas vale ressaltar que, em nossa análise, a posteriori, teremos de diagnosticar se o aluno tem a consciência de que a escolha da segunda estratégia mostra apenas uma estimativa dos resultados da análise. Nestes moldes, percebemos maior concentração dos dados entre 23 e 33,7 anos (próximos dos 34 anos), isto é, 50% da amostra estão exatamente entre 23 e 33,7 anos.

3ª estratégia: por meio de gráfico.

Utilizaremos a tabela da segunda estratégia para elaboração do gráfico, levando em consideração que o Histograma é o tipo de gráfico que mais se utiliza no ensino da análise exploratória de dados.

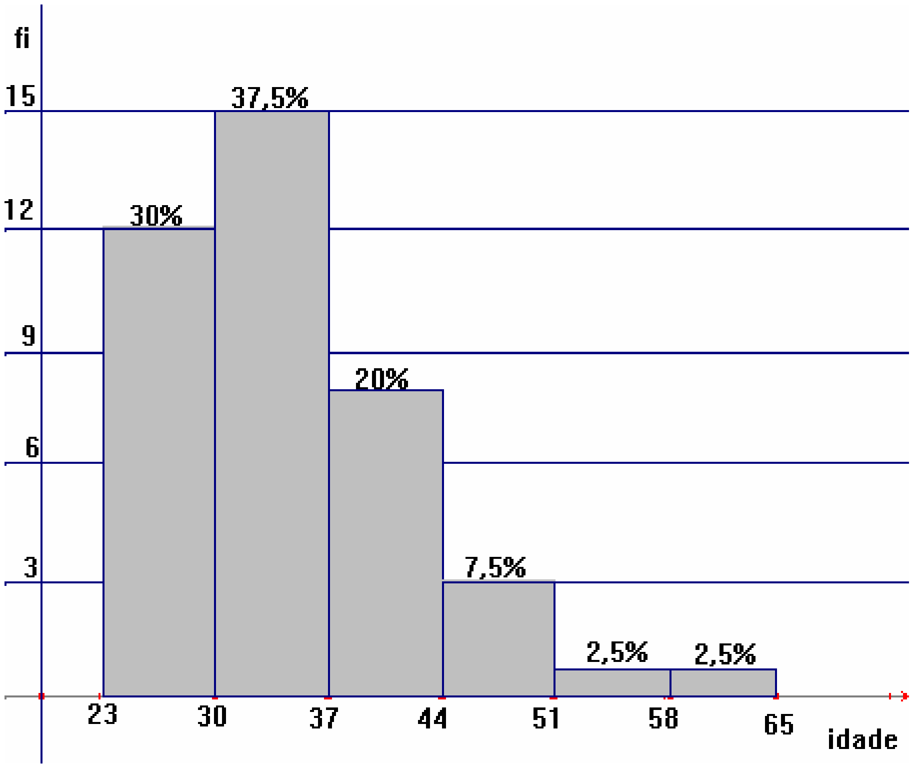


Gráfico 14: Histograma da variável idade.

Fonte: Bifi (2006)

O histograma consiste em retângulos justapostos com base nas faixas de valores da variável e com área igual à freqüência da respectiva faixa. Dividindo-se a coluna fi da Tabela 6 por 40, que é o total de elementos da amostra. Para facilitar a interpretação, colocamos em cada retângulo o valor percentual. Usaremos esta terceira estratégia para o cálculo das medidas de variação e separatrizes.

Para o cálculo das medidas de variação, representaremos como mostra a figura a seguir:

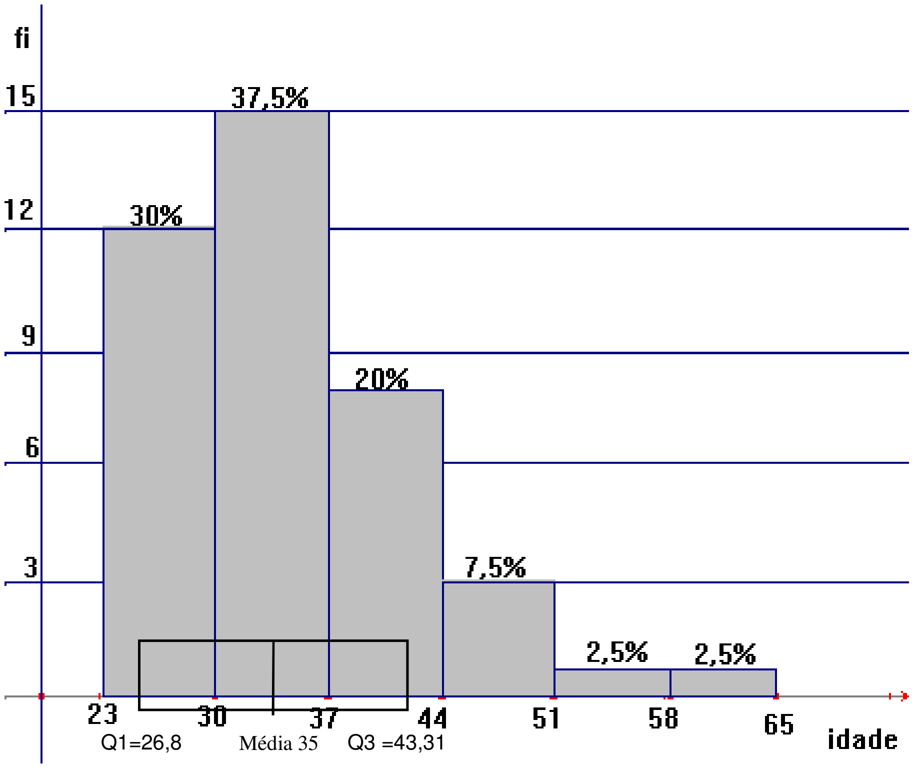


Gráfico 15: Estudo da média por meio do histograma.

Fonte: Bifi (2006)

Este tipo de representação também mostra a idéia de variação em torno da média, deixa visivelmente clara a relação da amplitude desta variabilidade com a amplitude da amostra. Os resultados possíveis e esperados são igualmente mostrados nas estratégias anteriores, mais ainda a que aborda a estratégia 2.

Já para os cálculos das medidas separatrizes, mostraremos o processo de como se chegar aos resultados. Calculemos a mediana da variável idade por meio do histograma. Inicialmente, identificamos o retângulo que deve conter a mediana. Uma simples soma das áreas resulta que a mediana pertence ao intervalo [30; 37[, já que até o valor 37 temos acumulado 67,5% (30% +37,5%) das observações dentro dessa faixa, Precisamos determinar um retângulo com área igual a 20%, pois é o valor que falta para atingir 50% (30% + 20% = 50%). Veja a ilustração abaixo, na qual o retângulo procurado está marcado com área mais escura, e os procedimentos finais para o cálculo.

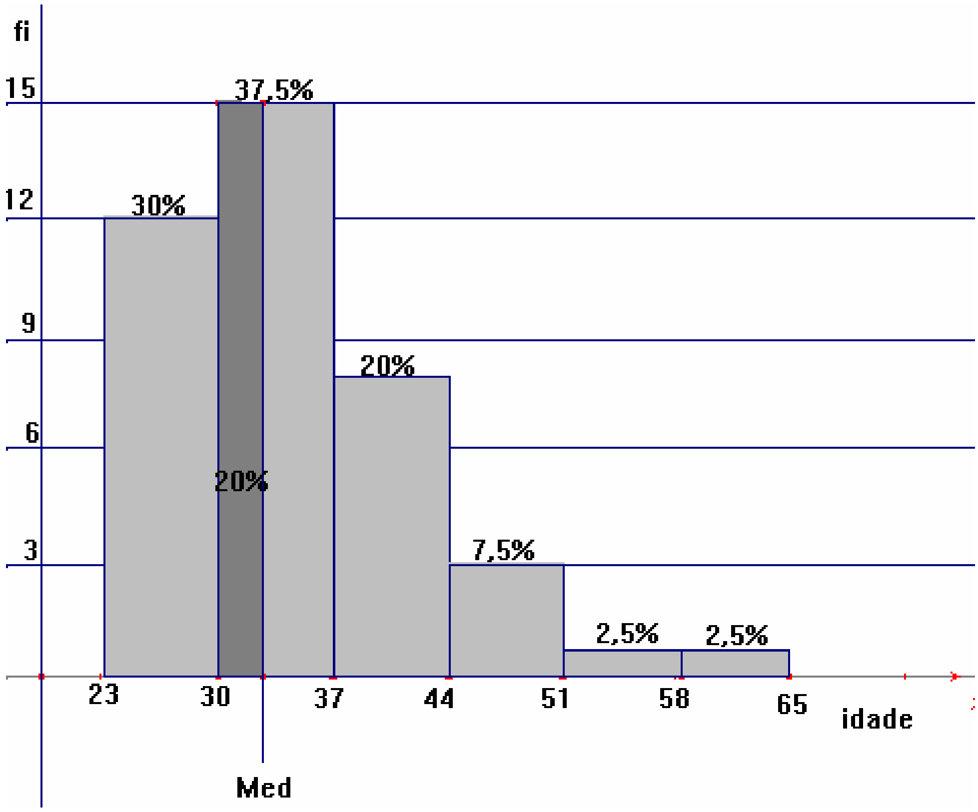


Gráfico 16: Estudo da mediana por meio do histograma.

Fonte: Bifi (2006)

Com o uso de proporções, estabelecemos a seguinte igualdade:

$$\frac{med - 30}{0,20} = \frac{37 - 30}{0,375} \Rightarrow md = 33,73$$

Este cálculo pode ser generalizado para situações em que o conjunto de dados é dividido em mais subgrupos. Um caso importante é o que dividimos o conjunto de dados em quatro subgrupos. Para tanto, deveremos determinar, além da mediana, dois valores, tais que 25% das observações ordenadas estejam abaixo de um deles e 75% abaixo do outro. Estes valores são o primeiro e terceiro quartis. O cálculo dos valores dos quartis poderá ser realizado de forma semelhante à descrita para a mediana, isto é, por meio do histograma.

Tomaremos a liberdade de não mostrar os procedimentos para tal cálculo por serem análogos ao anterior, focalizando a investigação dos níveis de funcionamento, os processos mostrados pela mediana serão suficientes para tal diagnóstico. É importante ressaltar as diversas formas de abordagem de um mesmo assunto para que, em nossa pesquisa, possamos diagnosticar os possíveis erros cometidos por alunos do Ensino Superior nas mais diversas formas de resolução que poderão surgir na aplicação da atividade.

Variável renda familiar (ROL)

Tabela 16: Renda familiar (dados fictícios)

300	380	387	400	400
406	490	500	540	554
600	630	700	700	760
770	800	850	860	890
890	1000	1000	1160	1180
1200	1200	1200	1340	1370
1400	1400	1420	1420	1500
1600	1600	1770	1770	1800

Fonte Bifi (2006)

Para o cálculo da média da renda familiar, usaremos a fórmula:

$\bar{x} = \sum \frac{x_i . f_I}{n}$. Neste caso, o valor da média da renda familiar será: $\bar{x} = \frac{39.018}{40} \cong 975{,}45\ reais$;

A variância será denotada pela fórmula: $s^2 = \frac{\sum (x-\bar{x})^2 . f_I}{n-1}$, logo o valor da variância será $s^2 = \frac{7813116}{39} \cong 200336{,}3$;

O desvio-padrão será a raiz quadrada da variância: $S = \sqrt{200336{,}3} \cong 447{,}58$;

O valor mínimo da amostra: 300 reais;

O valor máximo da amostra: 1,800 reais;

Amplitude da amostra: 1,500 reais;

Nas medidas separatrizes teremos:

O primeiro quartil (Q1): 596,25 reais;

A mediana será (md): 909,4 reais;

O terceiro quartil (Q3): 1.327,00 reais;

A moda é o elemento que aparece com mais freqüência dentro dessa amostra, será: 1200 reais;

Na amostra, a que a média salarial é de 975,45 reais aproximadamente e o desvio-padrão de quase 448 reais. Tendo uma amplitude amostral de 1.500 reais, percebemos que o coeficiente de variação em torno da média é de quase 60%.

Podemos analisar geometricamente a variabilidade em torno da média pela Figura 9:

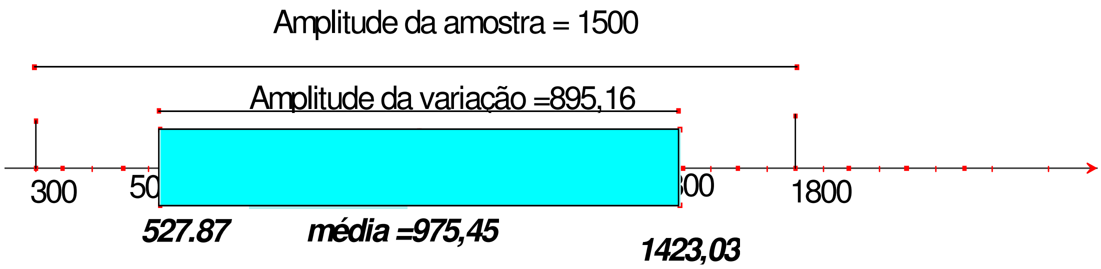


Gráfico 17: Variabilidade em torno da média.

Fonte Bifi

Pelo gráfico 17, percebemos a existência de uma concentração uniforme da amostra em torno da média, ou seja, uma concentração em renda mensal que varia entre 527,87 reais e 1.423,03 reais em uma média de 975,45 reais, representando 59,69% de toda a amostra. Esta análise geométrica dos dados coletados relata, com maior confiabilidade, o comportamento da amostra e qual tipo de distribuição ela apresenta (simétrica ou assimétrica). Neste caso, o aluno poderá definir qual o melhor valor para análise da variabilidade. Para tanto, vamos fazer um estudo dos quartis, conhecimento esperado dos alunos do Ensino Médio.

Utilizamos, para este estudo, o box-plot.

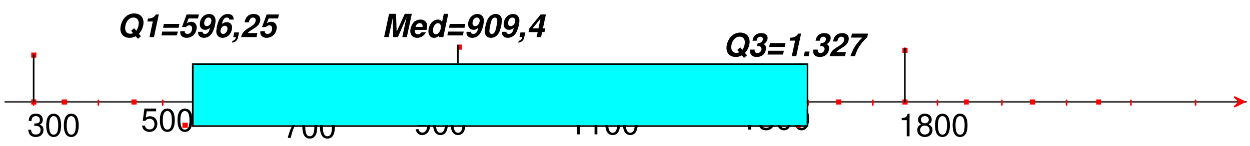


Gráfico 18: Variabilidade em torno da mediana.

Fonte Bifi (2006)

Verificamos pelo gráfico 18 que o Box-plot fornece uma análise mais detalhada e com maior precisão do comportamento dos dados coletados percebemos que com uma ligeira concentração à esquerda dos dados em relação à mediana. Nota-se maior concentração dos valores entre 300,00 e 909,4 reais, levando-nos a crer que a mediana seja o valor que melhor represente a variabilidade do conjunto. Pretendemos mostrar que, ao analisar a variabilidade de um conjunto de dados, não podemos simplesmente nos ater a um só tipo de medida, por exemplo, a média, mesmo que ela aparente ser um valor representativo. É imprescindível que o aluno do Ensino Médio, tenha habilidade e competência para analisar a variabilidade de conjunto de dados, escolhendo bem o valor que melhor represente o conjunto, cercando quase todas as possibilidades, para que suas conclusões futuras sobre o conjunto estudado tenham a confiabilidade que se espera do profissional nas atividades que lhe são atribuídas.

Segunda Parte

Nesta segunda parte, a distribuição apresentar-se-á na forma tabular e o motivo para esta representação é proposital para que possamos diagnosticar se os níveis de mobilização dos estudantes manifestam-se por meio de outra forma de representação dos dados. Assim, estaremos diagnosticando o uso pelos alunos de uma diversidade de representações, conforme Vergnaud (1998a).

Dessa forma, até por conta da mesma necessidade dos cálculos (média, quartis e desvio-padrão), acreditamos ser mais viável e fácil ao aluno, bastando ele completar a tabela com colunas auxiliares para encontrar os valores pedidos, porém não garantindo o sucesso na atividade.

Tabela 17: Distribuição de freqüência - qtde. carros/pessoa

Qtde de carros (xi)	***N. de pessoas (fi)***	$xi.fi$	$di = (xi - \overline{x})$	$di^2.fi$
1	10	10	1 – 2,27 = -1,27	16,129
2	25	50	2 – 2,27 = - 0,27	1,8225
3	15	45	3 – 2,27 = 0,73	7,9935
4	5	20	4 – 2,27 = 1,73	14,9645
Total	***55***	***125***		***40,9095***

Fonte Bifi (2006)

Comecemos os cálculos completando a Tabela 10 , fornecida na 2ª parte

1ª estratégia: Cálculo das medidas baseadas na distribuição de freqüência

1) Cálculo da média e desvio-padrão.

A média neste caso será $\overline{x} = \dfrac{\sum xi.fi}{n} \Rightarrow \overline{x} = \dfrac{125}{55} \cong 2,27$ carros por pessoa

O desvio padrão será $s = \sqrt{\dfrac{\sum di^2.fi}{n-1}} \Rightarrow s = \sqrt{\dfrac{40{,}9095}{54}} \cong 0{,}87$ carros por pessoa.

A análise esperada pelos alunos:

Nas medidas de variação, na amostra, temos que a média é de 2,27 carros por pessoa o desvio-padrão (0,87) carro por pessoa. Tendo uma amplitude amostral de três carros, percebemos que o coeficiente de variação em torno da média é de 58%. A amplitude em torno da média será quase de 1 carro. Uma forma de tentar “enxergar” o comportamento da amostra em relação à média seria mostrar uma representação geométrica, que indicaria quais os melhores caminhos a seguir:

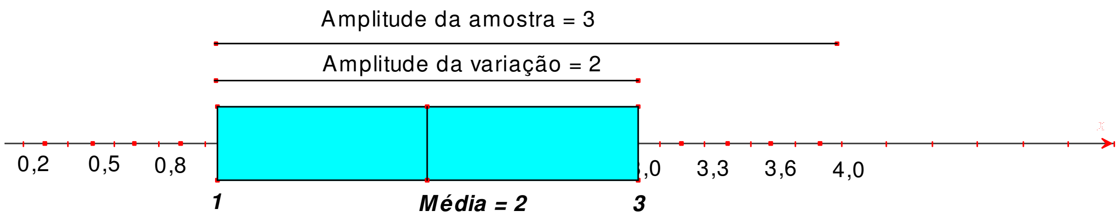


Gráfico 19: Amplitude em relação à média

Por esta disposição, podemos perceber que a dispersão dos dados não é alta, pois a média sendo $\cong 2$ e o desvio-padrão $\cong 1$, comparando com a amplitude total da amostra igual a 3, é fácil perceber que existe uma concentração à esquerda em relação ao conjunto todo e, sendo assim, é necessário um estudo dos quartis.

2) Cálculo da mediana, 1º quartil e 3º quartil.

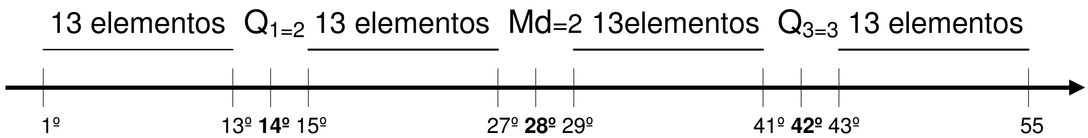
As medidas separatrizes também nos dão uma representatividade do comportamento da amostra e trarão os dados que parecerão mais adequados para a solução. Para resolver esta questão, utilizamos a seguinte estratégia:

Temos exatamente 55 valores para o nosso conjunto, se dividirmos o 55 por 4 teremos o seguinte:

$$55 \div 4 = 13 \text{ resto} = 3$$

Qtde de carros (xi)	*N. de pessoas (fi)*	*Fa*	
1	10	10	1º ao 10º
2	25	35	11º ao 35º
3	15	50	36º ao 50º
4	5	55	51º ao 55º
Total	***55***		

e traçarmos um reta e mostrar está divisão na reta ficaria assim:



$13 + 13 + 13 + 13 = 52$ elementos $+ Q_1 + Md + Q_3 = 55$ elementos

O valor da Mediana é 2; o valor do 1º quartil é 2; e o valor do 3º quartil é 3.

Podemos analisar que cerca da metade dos entrevistados ou 50% utiliza-se de dois carros ou menos, cerca de três quartos ou 75% dos entrevistados utilizam-se de três carros ou menos. Para a análise no box-plot, teremos :

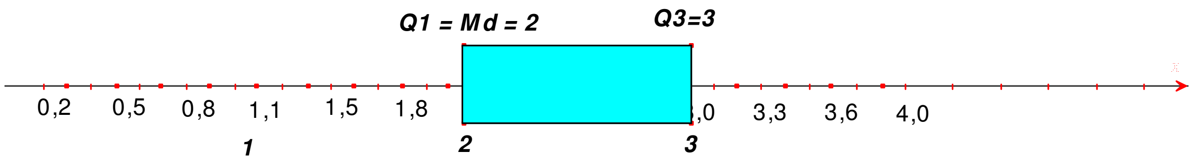


Gráfico 20: Representação da amplitude dos quartis.
Fonte: Bifi (2006)

Verificar que o primeiro quartil e a mediana coincidem e há uma concentração dos dados entre um e dois carros. Na análise do box-plot, fica claro que a mediana é o valor mais representativo ao o estudo da variabilidade do conjunto de dados.

2ª Estratégia Resolução por meio de gráfico

Na construção de gráficos estatísticos, percebemos que os professores do Ensino Médio, às vezes, dedicam pouco tempo ao ensino deste tópico. Levantamos a preocupação de que gráficos estatísticos podem ser um facilitador, trazendo informações resumidas, talvez, isso seja o suficiente para o tipo de análise ou, poderá ser um complicador, para distribuições com intervalo de classes, muitos dados são perdidos e a análise poderá não mostrar o que realmente se investiga. Por meio do gráfico abaixo, tentamos analisar o comportamento do conjunto de dados em estudo.

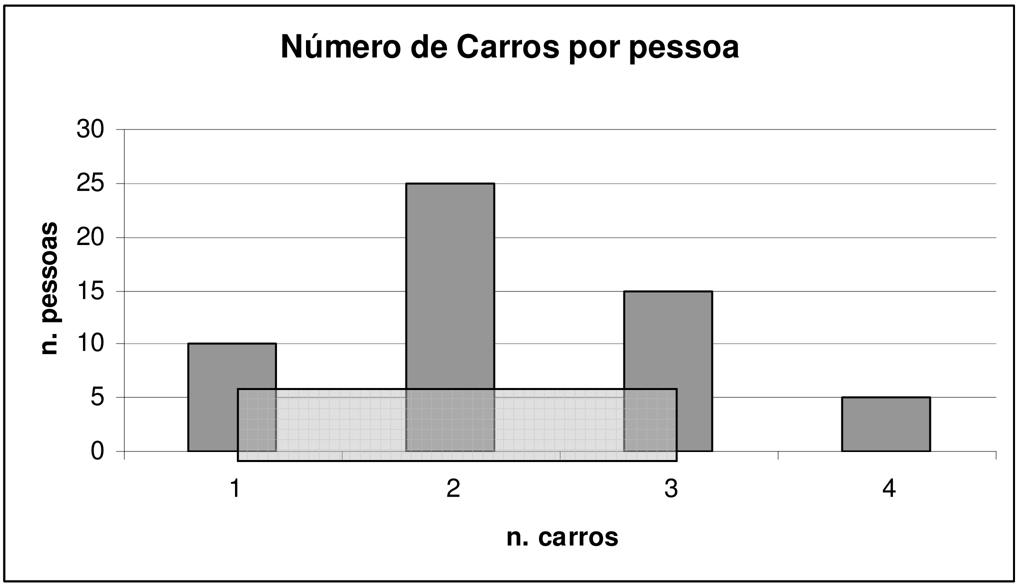


Gráfico 21: Estudo medidas centrais p/meio de gráficos.
Fonte: Bifi (2006)

Calculados os valores de variação, podemos representá-los no gráfico e ter uma noção clara do comportamento da variável e, por outro lado, é fácil perceber a existência de valores que estão distantes do intervalo da variação em torno da média. No caso, percebemos a necessidade de recorrer novamente às medidas separatrizes e verificar qual a melhor escolha para a representação da variabilidade. Vamos investigar a mediana no gráfico abaixo:

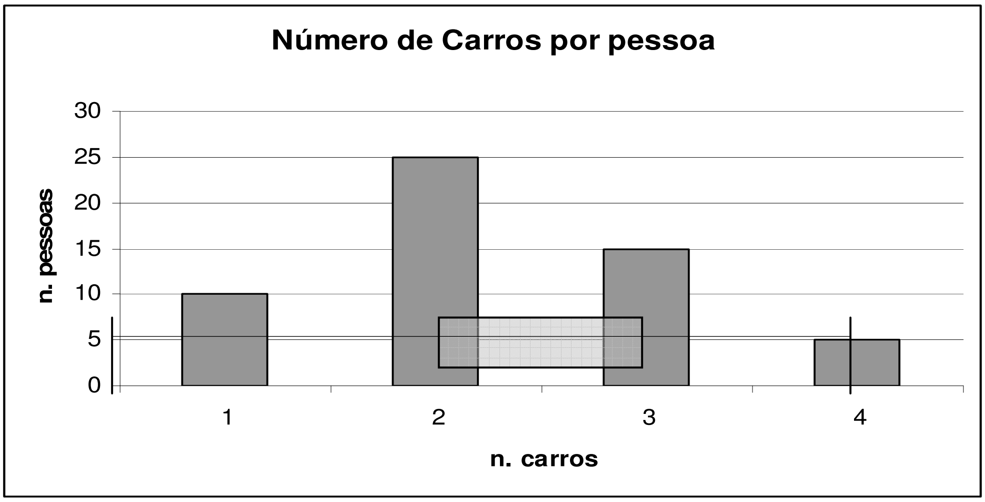


Gráfico 22: Estudo medidas separatrizes p/meio de gráf.
Fonte Bifi (2006).

Quando mostramos as medidas separatrizes, percebemos que o primeiro quartil e a mediana coincidem. Neste caso, existe maior concentração dos valores nos primeiros 50% da amostra. Na análise do aluno esperamos que, ele perceba a necessidade de se atribuir à mediana o melhor valor de representividade.

A Tabela 7 da segunda parte apresenta a distribuição com intervalo de classes, comecemos os cálculos, completando a Tabela 8, fornecida na Segunda parte.

Tabela 18: Distribuição de freqüência com intervalo de classe da Tabela 16

Tempo no trânsito	***N. de pessoas (fi)***	$P.médio(xi)$	***xi.fi***	$di = (xi - \bar{x})$	$di^2.fi$
0 ├ 2	10	1	10	1 – 4,33 = - 3,33	110,89
2 ├ 4	20	3	60	3 – 4,33 = -1,33	35,38
4 ├ 6	30	5	150	5 – 4,33 = 0,67	13,47
6 ├ 8	15	7	105	7 – 4,33 = 2,67	106,93
Total	***75***		***325***		***266,67***

Fonte Bifi(2006)

1ª estratégia: Cálculo das medidas com base na tabela

1) Cálculo da média e desvio-padrão .

Neste caso, a média será $\bar{x} = \dfrac{\sum xi.fi}{n} \Rightarrow \bar{x} = \dfrac{325}{75} \cong 4,33$ horas por pessoa;

O desvio padrão será $s = \sqrt{\dfrac{\sum di^2.fi}{n-1}} \Rightarrow s = \sqrt{\dfrac{266{,}67}{74}} = 1{,}89$ horas por pessoa.

Verificar o comportamento da variável com mais clareza quando é representado geometricamente e concluir se a média representa um bom valor de representação da variabilidade. Veja:

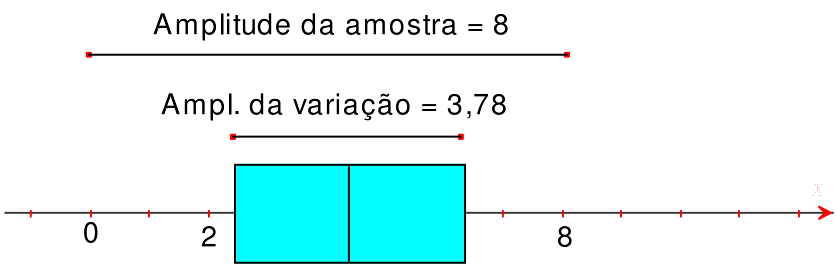


Gráfico 23: Cálculo da média a partir da tabela.
Fonte Bifi (2006)

1ª estratégia: Cálculo das medidas a partir da tabela.

A variável em torno da média apresenta uma amplitude de 3,78 horas, representando 47,25 da amplitude total da amostra. Não há segurança para afirmar se a média representa a variabilidade da amostra sem realizar os estudos das separatrizes, é o que esperamos dos alunos investigados.

2) Cálculo da mediana, 1º quartil e 3º quartil

Como a distribuição está com intervalos de classes, utilizamos as seguintes fórmulas:

Para a mediana:

$$md = l\inf + \left(\frac{\frac{n}{2} - \sum fant}{fi} \right).h$$

$l\inf$ = limite inferior à classe da mediana

n = número de elementos da amostra

$\sum fant$ = somatório da freqüência anterior à classe da mediana

fi = freqüência da classe da mediana

h = amplitude da classe

então

$$md = l\inf + \left(\frac{\frac{n}{2} + \sum fant}{fi} \right).h \Rightarrow md = 4 + \left(\frac{\frac{75}{2} - 30}{30} \right).2 \Rightarrow md = 4{,}5$$

Para a Moda: $mo = l\inf + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2} \right).h$

$l\inf$ = limite inferior à classe da mediana

h = amplitude da classe

$\Delta 1$ = diferença entre os valores do fi da classe modal com a anterior

$\Delta 2$ = diferença entre os valores do fi da classe modal com a posterior

então:

$$mo = l\inf + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2} \right).h \Rightarrow mo = 4 + \left(\frac{10}{10 + 15} \right).2 \Rightarrow mo = 4{,}8$$

Para o 1º e o 3º quartis, utilizamos a mesma fórmula da mediana, porém mudamos a fração de n para cada item:

1º quartil $Q1 = l\inf + \left(\frac{\frac{n}{4} + \sum fant}{fi} \right).h \Rightarrow md = 2 + \left(\frac{\frac{75}{4} - 10}{20} \right).2 \Rightarrow md = 2{,}88$

3º quartil: $Q1 = l\inf + \left(\frac{\frac{3n}{4} + \sum fant}{fi} \right).h \Rightarrow md = 4 + \left(\frac{\frac{3.75}{4} - 30}{30} \right).2 \Rightarrow md = 5{,}75$

Análise esperada dos alunos:

Nas medidas de variação, a média é de 4,33 horas por pessoa com desvio-padrão igual a 1,89 horas por pessoa. A amplitude total é de 8 horas por pessoa, com coeficiente de variação em torno de 47,25%. Vamos analisar estes dados por meio de uma representação geométrica, ou seja, o box-plot.

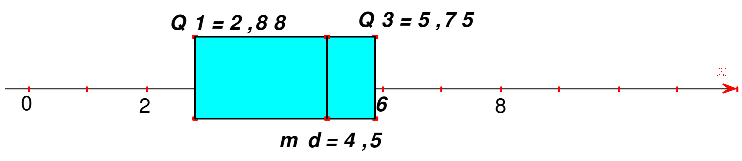


Gráfico 24: Variabilidade em torno da mediana.
Fonte Bifi (2006)

Mais uma vez, verificamos a concentração dos dados (à direita) da mediana, confirmando que a média, mais uma vez, não é um bom valor para análise dos dados e, sim, a mediana por ela explicar melhor o comportamento destes dados.

2ª estratégia: representação gráfica

Pelo gráfico abaixo, tentamos analisar o comportamento do conjunto de dados em estudo (Tabela 11). No gráfico, mostraremos os intervalos dos desvios em relação à média e, posteriormente, mostraremos os intervalos interquartílicos. Nessa primeira apresentação, mostraremos o intervalo dos desvios em relação à média.

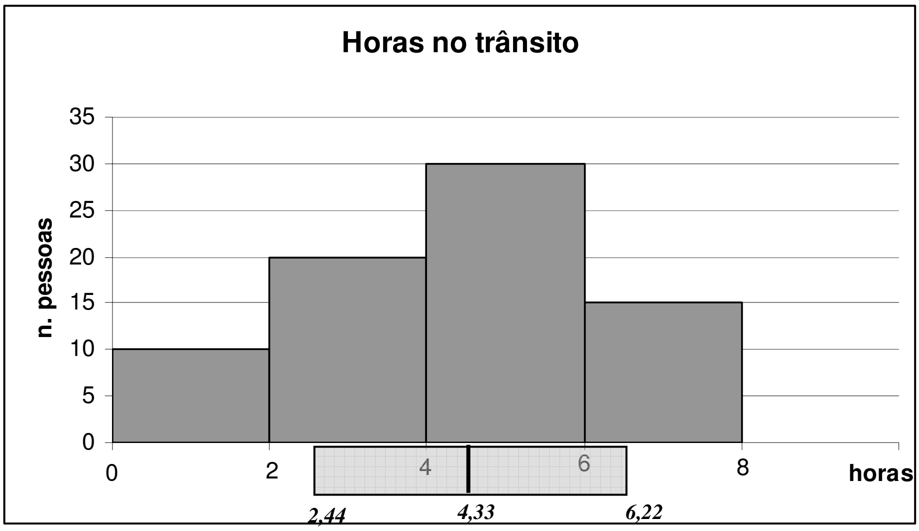


Gráfico 25: Análise das medidas centrais da Tabela 11 por meio de gráficos.

Fonte Bifi(2006)

Pelo gráfico 25, houve a concentração dos dados em torno da média e a partir daí podemos explicitar algumas conclusões pertinentes ao conjunto de dados em estudo. Por exemplo, que a média está na classe de 4 a 6, que os desvios apontam para no mínimo 2 horas e no máximo 6 horas de tempo no trânsito. Nosso próximo passo, será calcular a mediana por meio do próprio gráfico acima.

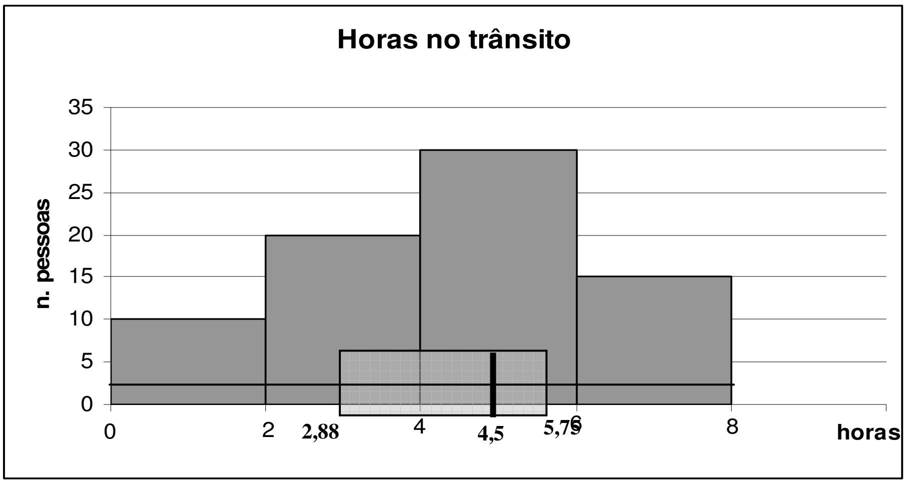


Gráfico 26: Análise das medidas separatrizes da Tab.11 p/meio de gráficos.
Fonte Bifi (2006)

Verificamos maior concentração à direita em relação à mediana. Neste caso, seria viável e mais seguro analisar o conjunto de dados pela mediana, pois ela representa melhor a amostra dos dados.

Terceira parte

A terceira parte pede para observar os dois gráficos (4 e 5) e, em seguida, pergunta-se: “Se você precisasse descrever esses dados para um cliente, como você os analisaria?”. Enfocamos também a necessidade de um outro tipo de representação de dados para diagnosticar as dificuldades que poderão surgir nesse tipo de representação.

Análise dos Gráficos

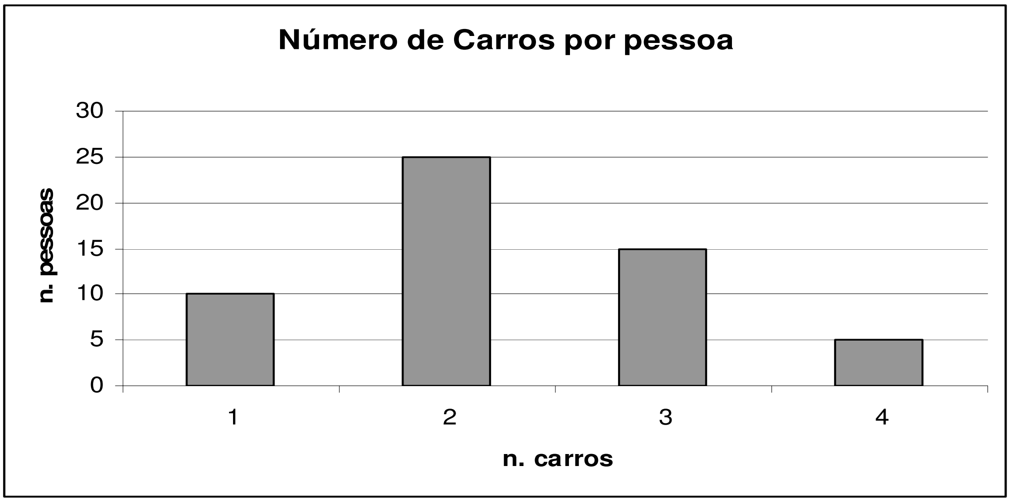


Gráfico 27: Retomada do Gráfico 8.
Fonte Bifi (2006)

De acordo com o gráfico, foram entrevistadas 55 pessoas. Destas, dez possuíam apenas um carro; 25, dois; 15, três e 5 cinco quatro carros. A média de carros por pessoa foi de quase dois carros por pessoa com desvio- padrão de aproximadamente 1.

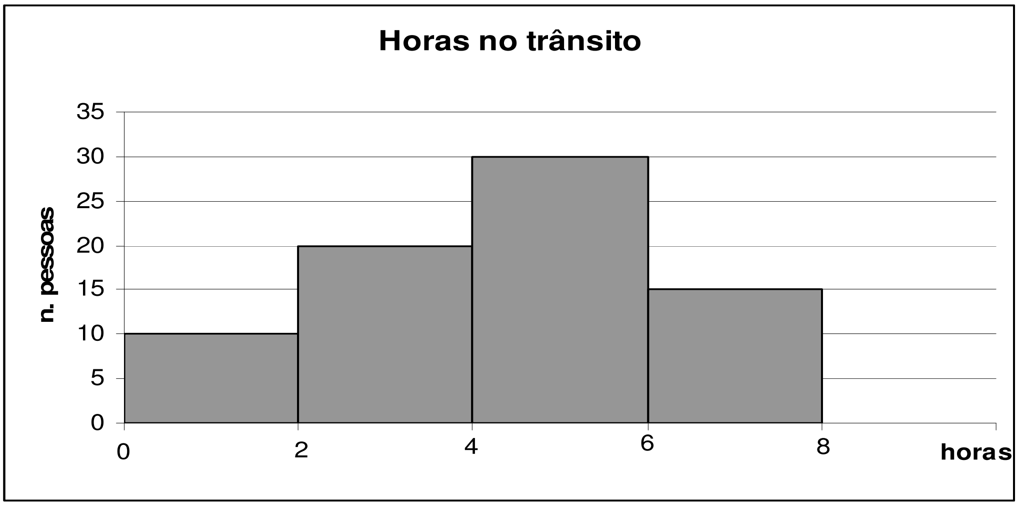


Gráfico 28: Retomada do Gráfico 9.
Fonte Bifi (2006)

Esta terceira e última parte, apresentaremos aos alunos dois gráficos que reforçarão nossa investigação dos níveis de funcionamento dos conceitos estatísticos, pois são passagens de estratégias que poderão surgir nas resoluções que se apresentarão pelos alunos. Esta terceira etapa necessária se, por acaso, nas etapas passadas, não surgirem estratégias desse tipo. Mas vale ressaltar que a resolução desta terceira etapa já foi abordada.

Estas etapas podem ser resumidas citando que os conhecimentos mínimos necessários são inúmeros, para que o aluno do Ensino Médio resolva as atividades propostas por meio das estratégias. Entre elas, estão os objetos matemáticos, já citados neste trabalho. Não podemos deixar de ressaltar que os resultados obtidos por meio dessas estratégias estão atrelados a cálculos algébricos. As dificuldades que poderão surgir por conta da Álgebra, podem ser um fator negativo para que nosso trabalho tenha êxito no objetivo traçado.

Independente das dificuldades que poderão surgir, as estratégias mostradas em nosso trabalho, nas três partes da atividade, permitirão ao aluno calcular as medidas de variação (média, desvio-padrão e coeficiente de variação) e, também, as medidas separatrizes (quartis), Isto força o aluno a uma análise que demandaria uma associação entre as medidas centro e as de variação, levando-o, assim, a um estudo da variabilidade em torno da média ou ainda, à uma associação entre as medidas separatrizes (mediana e quartis), porém a variabilidade seria em torno da mediana. No entanto, os dois casos estão associados com a amplitude total da amostra. A não associação entre estas medidas nos leva a crer que esse aluno encontra-se, segundo Robert (1998) em um nível técnico de mobilização desses conhecimentos.

A situação-problema apresentada nessa estratégia não sugere que o aluno faça associações entre as medidas, porém, em nossa pesquisa, esperamos encontrar alunos que as façam espontaneamente, caracterizando, segundo Robert (1998), um nível disponível de mobilização desse conhecimento.

Por outro lado, sem esta associação por parte do aluno e com a necessidade de algum tipo de intervenção do professor, sem que este dê a solução do problema e, sim, indique o caminho a ser percorrido, superando a dificuldade do aluno, segundo Robert (1998), classificamos em um nível mobilizável.

4.4 ANÁLISE DOS RESULTADOS

Voltaremos ao enunciado da atividade para que o leitor acompanhe melhor esta análise. Nossa atividade foi composta por três etapas: na primeira, apresentamos um banco de dados de variável discreta quantitativa; na segunda, duas tabelas representando uma distribuição de freqüências de variável discreta quantitativa (1ª tabela – quantidade de carros por pessoa) e variável contínua quantitativa (2ª tabela com intervalo de classes – tempo no trânsito por pessoa); e, na terceira, uma distribuição de freqüências representada graficamente. Para as duas primeiras partes da atividade, foi pedido que os alunos calculassem nas medidas de dispersão, a média e o desvio-padrão e nas medidas separatrizes, a mediana, moda e quartis.

Em seguida, solicitamos aos alunos que fizessem uma análise dos cálculos e relatassem quais das duas medidas encontradas representavam melhor a amostra. Para a última parte da atividade, pedimos que os alunos analisassem os dois gráficos apresentados e respondessem à seguinte questão: “Se você precisasse descrever esses dados para um cliente, como você os analisaria”.

Para aplicação das atividades, os alunos foram organizados em grupos que denominamos por grupo 1, grupo 2, grupo 3, grupo 4 e grupo 5 na seqüência de nossa análise. Os procedimentos descritos por grupo foram analisados de acordo com o descrito em nosso quadro teórico, no qual a educação estatística é retratada como a habilidade para interpretar, avaliar criticamente e, se necessário, discutir sobre: informações estatísticas, argumentos e mensagens (Gal 2002,p.28) conforme abordado no capítulo que tratou do quadro teórico.

Nesse contexto, procuramos identificar os níveis de mobilização dos conhecimentos pelos alunos (Robert, 1998), descritos como: técnico, mobilizável e disponível, conforme também é apresentado no capítulo do quadro teórico. Buscamos, assim, identificar um patamar não só em um campo de conhecimentos matemáticos, mas também no estatístico, caracterizado e sustentado por objetos matemáticos apresentados, de certa maneira, por meio de teoremas matemáticos e associados a diversos quadros, registros de representações e noções intuitivas.

No decorrer das atividades, percebemos que a maioria dos alunos conseguia calcular as atividades, cometendo poucos erros na utilização das fórmulas, mas o que chamou a atenção foi o fato de que os alunos para analisarem, calculavam tudo: tabela, média, moda, mediana, desvio-padrão, quartis; Box-plot, histograma e cometeram erros de ordem analítica.

Observamos que, na maioria dos grupos, os alunos são capazes de aplicar fórmulas para cálculo das medidas solicitadas, e conseguem uma análise parcial. Entretanto, não conseguem analisar com clareza o significado dos resultados obtidos.

As análises feitas pelos alunos, tanto na forma escrita como na oral não foram totalmente claras, mesmo assim, podemos perceber que já começa a surgir uma visão estatística, como se podem ver formas diferentes de resolução das atividades, pois eles estavam de fato buscando estratégias de resolução, portanto, ai não se caracterizava um contrato didático, ou seja, eles estavam buscando ou transferindo conhecimento disponível para ser utilizado.

Primeira Parte

Foi apresentada aos alunos uma tabela contendo dados brutos relativos à idade e à renda mensal de 40 pessoas. As questões colocadas foram:

1) Encontre, nas variáveis *idade* e *renda mensal*, a média e o desvio-padrão. Como você analisaria esses resultados? 2) Encontre, nas variáveis *idade* e *renda mensal*, a mediana, o 1º quartil e o 3º quartil. Como você analisaria esses resultados? 3) Se você precisasse explicar o *comportamento* da variável *idade* para um cliente, você usaria o item (1) ou o item (2)? Explique por quê.

Na nossa análise do grupos, podemos fazer as seguintes observações, segundo os preceitos por Robert(1998),descritos no capítulo quadro teórico, os grupos 1 e 4, optaram por resolver a primeira atividade sem intervalo de classe, enquanto os grupos 2, 3 e 5, fizeram a opção por resolver esta mesma atividade utilizando intervalo entre as classes, assim, podemos notar que duas estratégias para a manipulação dos cálculos foram utilizadas, porém não fizeram justificativa alguma a respeito de suas escolhas.

Os cálculos realizados com o uso da calculadora e a utilização do algoritmo escolhido pelo aluno para a manipulação dos resultados, demostrou que, quanto aos cálculos, todos dominavam perfeitamente, mais na resposta escrita, nem um dos grupos justificou a escolha do algoritmo utilizado, levando-nos a questionar o significado construído pelo aluno sobre estes conceitos.

No final da primeira parte, que envolvia média e desvio-padrão, nem um dos grupos paravam a atividade para observar os resultados encontrados, já partiam para o cálculo dos quartis, que com unanimidade utilizarão a fórmula como ferramenta de resolução, também construíram o box-plot e o histograma.

Somente no final após confrontarem o box-plot com o histograma, é que os grupos analisavam os resultados, surgiram as seguintes análises pelos grupos: (grupo 1), fez a seguinte análise, para a primeira atividade.

> Entre os clientes de 23 a 65 anos a idade média é de 33,7 anos, para os clientes que utilizam cartão de crédito, destes a maior parte estão concentrados entre os 23 e 40 anos.

média é de 35 anos, essa é uma idade boa em relação a idade obtida no histograma, esta está entre primeiro quartil e o terceiro quartil, 50% dos elementos do conjunto estão localizados entre o primeiro quartil e a média 35 anos.

A moda desse rol é de 30 anos, pelo o que se pode observar.

Note que este grupo não somente não prestou atenção no que estava sendo pedido no enunciado do problema como também ainda não sabe a diferença entre média e mediana, pois hora ele diz que a média é 33,7 anos e depois ele refere-se a média novamente como sendo 35 anos. Pelos cálculos que temos em mãos, podemos notar que ele, o (grupo 1) calculou corretamente a mediana e apresenta o resultado de 33,7 anos e, também, calculou corretamente a média 35,075 anos, mais no momento de mencionar estes existe muita confusão, mostrando que ainda não conseguiu assimilar tais conceitos.

Bakker (2004, apud Silva, 2007)[2] conta sobre sua experiência numa sala de 5 série do ensino fundamental, em que o professor fez uma pergunta que se assemelhava com Estatística e um aluno respondeu médiamodamediana, com se fosse uma única palavra.

> Este incidente exemplifica o que uma grande quantidade de pesquisas em educação estatística tem relatado: muito frequentemente 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 box-plots, 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 apud Silva, 2007).

[2] Tese em andamento, com defesa prevista para maio de 2007, com o título "Variação e variabilidade em Estatística: um estudo com professores de Matemática da Escola Básica".

No box-plot, ele utiliza (Md) referindo-se a mediana, mais no entanto o que utiliza como valor é o 35 que na verdade é o valor da média. No histograma podemos notar que foi construído sem problema, porém ele tenta confrontar o box-plot com o histograma, mais faz de maneira equivocada devido ao fato de não ter incorporado corretamente o conceito de média e de mediana. Quanto ao fato da comparação entre os dois gráficos é muito bom pois significa que ele estaria transferindo conceitos e mobilizando para uma situação nova, mas o fato de não ter assimilado conceitos fundamentais como média e mediana, não poderia dar certo tal confronto. Quanto ao fato de não ter mencionado nada sobre o desvio-padrão, mostra que além de não prestar atenção no enunciado da atividade, pois os cálculos e o algoritmo para tal foram aplicados corretamente, demostra que para este grupo o que importava era apenas calcular sem saber de fato se era necessário estes cálculos.

Neste contexto, supomos que estão em um nível técnico de mobilização dos conceitos abordados.

> Este nível corresponde para nós a dos focos em funcionamento indicados, isolados, colocando em jogo as aplicações imediatas de teoremas, propriedades, definições, fórmulas, etc. Ele contextualiza de maneira simples, sem etapas, sem trabalho preliminar de reconhecimento, sem adaptações. Isto concerne, sobretudo, e funcionamentos úteis (ferramentas) levando em conta definições (ROBERT, 1998 p. 165).

Grupo 2

> Vejamos análise (grupo 2), como a mediana é aproximadamente 33,7, então obtivemos 50% das idades dos clientes da empresa.
> No primeiro quartil temos 28,8, isto é 25% dos clientes aproximadamente têm menos do que 30 anos.
> A idade que aparece com mais freqüência é 30 anos, isto é, a moda.
> No histograma podemos analisar que de 30 à 37 anos, obtém-se mais cartões de crédito e pessoas com mais de 37 anos tem menos números de cartões.
> Com os nossos cálculos, podemos perceber que a maior concentração de cartões de créditos esta centralizada na média 35 anos.
> Quanto ao desvio-padrão, obtivemos (8,76), isto nos dá uma amplitude de (17,52) entorno da média que é de 35 anos, e uma amplitude total de 42 anos, isto significa que temos uma variação de aproximadamente 41% em torno da média, é o mesmo que dizer que em cada 100 pessoas 41 tem entre 26 e 43 anos.

Este grupo mostrou que embora cometam alguns equívocos de ordem analítica já estão em um nível de conceitualização diferente do (grupo 1), vejamos: quando ele (grupo2) fala da mediana no começo da análise, eles apontam para o valor 33,7 e mencionam que este é o centro da distribuição que 50%, dando a entender que sabiam que abaixo deste valor teríamos 50% dos dados e que acima deste valor os outros 50% restantes. Assim podemos entender que este conceito foi assimilado pelo grupo. Quando menciona que 28,8 é o primeiro quartil, e que aproximadamente 25% dos clientes tem menos de 30 anos, os alunos arredondam o valor, o que não é aprovado.

Quanto ao analisar o histograma, comete um equivoco por não ter levado em consideração os quartis, que o próprio grupo menciona Q1=28,8 e Q3= 39,6, mais também é verdade que pelo histograma nota-se que a maior concentração está entre 30 e 37 anos.

Quanto a calcular a média, o grupo fez perfeitamente e além do mais fez uma boa análise levando em consideração o desvio-padrão.

Desta forma, podemos observar que já existe uma alfabetização estatística, básica já que identifica o texto que está sendo lido, conhece os cálculos a serem abordados da Estatística e da Matemática e pela análise da variabilidade em torno da média, podemos notar que o grupo entendeu este conceito. Mas o grupo também mostrou um forte efeito do contrato didático quando para responder a primeira pergunta não havia a necessidade de construir todos os cálculos, isso significa estes cálculos devem servir para alguma coisa. Não consideram o fato de que os cálculos já efetuados podem ser suficientes para a análise solicitada.

Partindo para os cálculos das medidas separatrizes, não demonstrou dificuldades no desenvolvimento dos procedimentos, mas algumas dificuldades em redigir o texto de justificativa dos cálculos.

Como dissemos no capítulo quadro teórico, Robert (1998) classifica como nível mobilizável o caracterizado por um nível de fazeres em funcionamento mais amplo do que um conhecimento técnico. Por exemplo, a resolução de um problema proposto exige do aluno a adaptação de seus conhecimentos para aplicar o teorema adequado. O aluno precisará aplicar várias vezes o mesmo objeto matemático ou utilizar objetos distintos em etapas sucessivas ou, ainda, articular duas ou mais informações de natureza diferentes. O conhecimento mobilizado põe à prova um "fazer", que coloca em funcionamento uma justaposição de saberes dentro de um domínio dado, direcionado para uma organização.

Segundo a autora:

> Um saber é dito mobilizável, quando é bem identificado, é bem utilizado pelo aluno, mesmo que tenha sido necessária uma adaptação ao contexto particular (Robert, 1998, p. 166).

Neste contexto, supomos que este (grupo 2), para está atividade mostrou-se que estão em um nível mobilizável de mobilização de conhecimentos de média, desvio-padrão e separatrizes.

Grupo 3

> "Analise feita pelo (grupo 3), Entre os clientes de 23 à 65 anos a idade média é de 33,7 dos clientes que usam o cartão de crédito, destes onde a maior parte está concentrada é entre os 23 anos aos 40 anos.
> A média é de 35 anos, essa é uma idade boa em relação a idade obtida no histograma.
> Nesta parte está entre o primeiro e o terceiro quartil é aproximadamente 50% do total de clientes.
> A moda desse rol é de 30 anos, pelo que vimos nessa questão".

Podemos notar que, o grupo passou a resolver simplesmente, sem nenhum planejamento, e ainda podemos observar que os conceitos para este grupo ainda não fazem sentido, pois confunde média com mediana, fizeram todos os cálculos corretamente com exceção do box-plot que construiu de forma equivocada, onde utilizou a média no lugar da mediana.

Assim pelo que já foi exposto, consideramos que para está atividade o grupo se mostrou em um nível de mobilização técnico Segundo os preceitos de Robert(1998).

Grupo 4

O grupo 4 fez a opção por resolver a atividade sem intervalo de classe e como os demais grupos fizeram a tabela, calcularam a amplitude total, e embora não utilizassem o intervalo de classes fizeram os cálculos também para os intervalos entre as classes, calcularam a média a mediana, primeiro quartil, terceiro quartil, construíram o box-plot o histograma, tudo calculado perfeitamente e na hora da análise, nada é justificado é como se os dados não tivessem conexões, veja a análise.

> Com relação ao histograma, concluímos que a idade de 30 a 37 anos é onde está a maior concentração dos dados.
> Nas demais idades há uma concentração menor de usuários de cartão de crédito, isso quer dizer que de no mínimo 23 à mediana 30 se concentra a maior idade, 50%, da mediana 30 à terceiro quartil a 25% das idades, do terceiro quartil ao valor máximo 65 concentra outro 25%.

Note que este grupo tem um entendimento totalmente fragmentado, eles fazem todos os cálculos mais não sabem para que, cada resposta que este grupo apresenta, não é se referindo a nem um tipo de estudo, construíram o box-plot mais não conseguem fazer uma análise correta do seu resultado, não mencionam nada sobre o que foi pedido na introdução da atividade, embora tenha calculado corretamente a média não calculou o desvio-padrão, ficando assim insuficiente para a análise do que foi pedido, deixa uma marca de contrato didático.

Segundo Robert (1998), um conhecimento é caracterizado como nível técnico, quando, para resolver um problema, o aluno recorre às indicações isoladas, colocando em jogo aplicações imediatas de teoremas, propriedades, definições, fórmula, etc.

Trata-se de uma contextualização simples, local e sem adaptações. Por exemplo, se pedirmos para um aluno fazer uma representação gráfica de um conjunto de dados, ele aprendeu como elaborar o gráfico, e pode fazê-lo sem, no entanto, saber como interpretar de forma correta os dados ali representados, pois a interpretação de representações aponta para o estacionamento de relações complexas, indo além de simples aplicações de fórmulas e procedimentos.

Podemos por este protocolo enquadrar o (grupo 4), como nível técnico.

Grupo 5

Já o (grupo 5), fizeram os cálculos parcialmente, e além do que não mostrou nem um cálculo para a média e nem tão pouco para desvio-padrão, ficando assim, impossível, fazer qualquer análise desta atividade para este grupo, supondo, assim, que este grupo precisaria ser melhor preparado.

Nessa primeira etapa do protocolo, no que diz respeito à variável discreta, os alunos conseguem fazer a transformação necessária para a resolução do problema proposto, implicando, assim, um nível mobilizável de conhecimento e dentro da categoria operacional, ou seja, o algorítmico. Porém podemos perceber que, na segunda etapa do protocolo, essa mobilização não ocorre. Os alunos tentam manter o mesmo padrão de raciocínio para estabelecer estratégias de resolução, mais pelo o que foi mostrado não obtiveram sucesso, ficando assim para esta segunda etapa a classificação de nível técnico.

Para a segunda questão, o (gruo 1) seguiu o seguinte protocolo, construiu uma tabela, desta vez utilizando intervalo de classe, mostrado assim que também conhece este método, já que para a primeira atividade haviam montado a tabela sem intervalo de classe, calculou a média, a variância e o desvio-padrão, mediana e os quartis construiu o box-plot e o histograma. Fez a seguinte análise:

> Observa-se maior concentração abaixo da média, mesmo que abaixo da média a renda diminui, isso indica que a maioria das pessoas são de baixa renda, nota-se que a moda está entre 300 e 537, o que comprova a baixa renda da maioria, a maior concentração de estudo entre primeiro quartil e terceiro quartil não coincide com a maior concentração de pessoas. A menor concentração de pessoas são de renda superior a todos os outros o que comprova que muitos tem baixa renda e poucos possuem renda superior a 1500, em que mesmo com mais pessoas entre a mediana e terceiro quartil as que primeiro quartil e mediana, isso não significa que a maioria possui maior renda.

Podemos notar que pela resolução da atividade, que não existe dificuldade na aplicação do algoritmo, mais sim muitas dificuldades para analisar os resultados obtidos. Ele tenta criar ligações entre as partes mais o faz de forma equivocada, mas já se pode observar também que este grupo, embora não consiga visualizar os dados obtidos como um todo, já começa a despontar para um conhecimento maior. Ele menciona o valor 1500, mais não fala de amplitude, ele lê a tarefa consegue fazer vários cálculos de maneira correta mais não sabe analisar se tais cálculos seriam necessário, ele calcula a média e a mediana mais não diz qual delas é melhor para representar este conjunto de valores.

Segundo Gal (2002), em suas bases de alfabetização na Estatística, não há regras ou critérios para uma análise crítica de dados estatísticos. Por exemplo, o autor não defende que suas bases de alfabetização sigam uma ordem e que estas devam necessariamente passar pelos cálculos estatísticos ou matemáticos. Ao analisar uma tabela ou até mesmo um gráfico estatístico, pode o leitor usar seu senso crítico e intuitivo, e perceber, por uma análise visual, o que esses dados podem estar dizendo. (Grupo 1) poderia tentar dar resposta conclusiva por meio dessa análise visual, ou ,ainda, analisar por meio de uma apreensão perceptiva dos dados representados graficamente.

Nossa hipótese, feita por essa análise, mostrou que, mais uma vez, a necessidade de apresentar cálculos numéricos para justificar resultados foi muito forte para essa dupla. A matemática, que é uma das bases de Gal, manifestou-se com maior intensidade, e foi a ferramenta encontrada para justificar a análise dos dados.

Grupo 2

Análise para o (grupo 2): Seguiu o mesmo protocolo do (grupo 1), sua análise foi a seguinte:

> Analisando a renda familiar podemos perceber que a média R$975,00, a variância de 200336 e desvio-padrão 447,58.
> Como a mediana 909 e no primeiro quartil 596,25 e o terceiro quartil 1327, podemos analisar que 25% das famílias tem renda mínima de R$300,00 à R$596,25, e que R$1327,00 à 1800 representam os 25% que recebem mais, 50% das famílias recebem entre R$596,25 à R$1327,00, sendo que a mediana é de R$909,40, está nos dá uma visão através do box-plot, de onde está a maior concentração desta renda nos primeiro 50% ou após a mediana e pela observação existe uma maior concentração da renda abaixo da mediana, ou seja 50% das famílias recebem entre R$300,00 e R$906,40.

Note que para este grupo os resultados obtidos com os cálculos são utilizados de maneira clara, conseguem fazer uma análise que relacione um conceito com outro, mostrando segurança em suas respostas, mais ainda sim não perceber que alguns cálculos são desnecessários.

Grupo 3

Análise do (grupo 3), este grupo fez todos os cálculos, assim, como os outros grupos demonstrando que não têm dificuldade para tal, mais não fez a análise pedida, ficamos assim, impedidos de fazer nossa análise.

Grupo 4

Análise do (grupo 4), vejamos sua análise já que desta vez optaram por utilizar intervalo de classes, os cálculos estão quase todos corretos, tendo uma pequena discrepância apenas na mediana. Observa-se que a mediana é aproximadamente R$672,00 então 50% das rendas familiares varia de R$ 300,00 à R$672,00 e que os outros 50% são de renda maior ou igual a mediana, mesmo assim pelo gráfico temos uma concentração maior a esquerda da mediana.

Nota-se pelo (grupo 4) foi exatamente um relato dos valores encontrados, ou seja, o que percebemos é apenas que o grupo fez uma transcrição da linguagem matemática encontrada nos resultados para a linguagem coloquial, sem, contudo, fazer uma análise do significado desses valores. Assim, o efeito do contrato didático usual está novamente presente, já que grupo precisava, naquele momento, de uma resposta para o professor que acabara de realizar uma pergunta.

Nota-se que pela resolução da primeira atividade que o grupo sabe calcular a mediana, então este equivoco ocorrido nesta atividade poderá ser encarado como uma falha simples. O grupo algebricamente encontrou os resultados solicitados, e fez uma análise um pouco simplista e com algumas falhas ao redigir.

Grupo 5

Análise (grupo 5), Desta vez o grupo em questão resolveu toda a atividade proposta, fez a escolha dos algoritmos fez seus cálculos, não deixando dúvidas que o fazem muito bem, agora veremos a análise feita pelo grupo.

> Como a média é aproximadamente R$975,00 a renda média das famílias fica neste valor, o primeiro quartil é de aproximadamente R$596,25 e o terceiro quartil é de aproximadamente R$1327,00, Pelo box-plot podemos observar que a renda mínima é de R$300,00 e a máxima é de R$1800,00, dentro deste ainda podemos observar os quartis Q1 e Q3, entre eles ficam concentrados 50% das rendas que é o centro desta distribuição ou mediana R$909,4, no histograma quando comparado com os quartis podemos observar que entre Q1 e Q3, teremos uma amplitude de aproximadamente R$731,00, que nos da aproximadamente 49% dos dados observados no histograma.

Na análise deste grupo, podemos identificar uma ligação de idéias ou mesmo de conceitos que justificou sua análise.

A conclusão sobre os grupos investigados, referindo-se ao nível operacional, focalizando o processo algébrico nas duas primeiras etapas da atividade, é que os grupos se encontram no nível de conhecimento técnico para mobilizável, segundo Robert (1998). Em se tratando de nível analítico, os grupos conseguem explicar parcialmente os conceitos mobilizados nos cálculos não tiveram seu significado explicitado ou mesmo justificado pelos grupos, levando-nos a inferir que este conhecimento, se existente, permaneceu implícito. Mesmo com questionamentos durante a atividade que visavam proporcionar condições para que os grupos exteriorizassem o significado por eles atribuído aos valores calculados, os alunos permaneciam ligados somente aos valores numéricos, acreditando. Sendo assim, inferimos que as duplas classificam-se em nível técnico no contexto algébrico, segundo Robert (1998).

Grupo1

> usaria o item 1 e 2 por que pelo que entendemos é claro que não se pode existir 2,27 carros por pessoas, tem que ser representado por uma variável discreta, mais quando se trata de tempo no trânsito, as variáveis podem ser continuas.

Note que a associação que o grupo fez é muito boa, mais não consegue explicar qual é a melhor tabela 1 ou 2.

Grupo 2

> Pela tabela 1, a concentração de pessoas que tem carro está entre 45,45% à 27,27%, em média mais de 50% de pessoas tem carro. Em média há de 18,18% a 9,09% de pessoas que não tem carro. No gráfico seguinte mostra que a concentração de tempo no transito está entre 26,66% à 40%, isso, indica que eles

> permanecem no transito em média de 2 a 6 horas. Já 13,33% permanecem 1 hora, menos tempo que os outros 20% que permanecem aproximadamente até 8 horas. Portanto os dados são complementares, precisaríamos das duas tabelas.

Na análise feita por este grupo mais uma vez podemos ver que eles tem facilidades para os cálculos, mais muitas dificuldades para expressar suas justificativas.

Grupo 3

> serão necessários as duas tabelas, pois serão passados os dados quanto o tempo no transito e com as duas tabelas daria para se definir tudo nos melhores detalhes para o cliente.

Com está análise não se pode analisar nada sobre o nível em que o grupo se encontra para está atividade.

Grupo 4

o grupo apresentou apenas cálculos, mais não respondeu nossa indagação, novamente ficando impossível de tecermos uma análise.

Grupo 5

> serão necessário os dois itens pois serão passados dados completos e úteis ao cliente, tanto o tempo no transito quanto a quantidade de automóveis, assim, é necessário a informação completa.

Novamente podemos observar que através da resposta dada pelo grupo, não se pode fazer a análise esperada.

Conclusão da terceira parte: Esta tarefa permitiu observar que os alunos neste contexto ainda apresentam maiores dificuldade na interpretação dos resultados, é necessário que se trabalhe um pouco mais atividades deste tipo.

De um modo geral, ficou a idéia que para a aprendizagem ser profunda é necessário propor aos alunos de forma equilibrada tarefas cujas as características se complementem, Isso possibilitará a mobilização das capacidades de ordem superior e uma aprendizagem mais rica e estimulantes. Assim, para a terceira atividade os grupos se enquadram ainda no técnico pois novamente buscaram os cálculos e não fizeram boas interpretações.

CONSIDERAÇÕES FINAIS

Podemos observar pelos fatos históricos e atuais que a Estatística tem se destacado há várias décadas e vem se destacando ainda mais pela sociedade moderna sendo um dos principais temas de pesquisa em Educação Matemática. As dificuldades apresentadas por alunos, diagnosticadas em pesquisas que abordam o tema, em todas as faixas de, tornado-se molas propulsoras que incentivaram pesquisadores a investigar os possíveis fatores que influenciaram no processo ensino-aprendizagem.

Nossa preocupação e incentivo para a pesquisa surgiram a partir das atividades proposta por Bifi (2006), quando entrei em contato com este trabalho e ao observar as questões surgiram algumas indagações. São elas 1ª como o professor de matemática de Ensino Médio lida com o conteúdo de estatística? 2ª Qual a opinião que esse professor têm sobre a proposta de trabalhar com tais conteúdos no Ensino Médio? E 3ª O professor acredita que seus alunos sejam capazes de resolver situações-problema relacionados ao bloco tratamento da informação? Destas indagações surgiu então esta pesquisa, a relevância deste trabalho, além de estar atrelada à necessidade de buscar novas metodologias que auxiliem profissionais e pesquisadores da Educação Matemática em sua interface com a Educação Estatística é, ainda, um alerta de que o ensino da disciplina Estatística não pode ser encarado como somente um ramo do campo da Matemática, mas também mostrar que a Estatística tem suas particularidades, próprias da disciplina, e que é tão importante quanto a própria Matemática no processo de formação do cidadão de qualquer área profissional, tornando-o alfabetizado nesse componente.

Antes dos resultados finais de nossa pesquisa, faremos um apanhado de como tudo se deu. Primeiramente submetemos o professor colaborador a um questionário gravado em áudio e depois transcrito, este questionário tinha a finalidade de definir a partir de que ponto teríamos que trabalhar com o professor os conceitos estatísticos, feito este diagnostico e decidido quais dos conceitos de base da estatística efetivamente teríamos que trabalhar, foram marcados encontros que duravam em média 2h15 cada totalizando cinco encontros individuais e mais um ultimo encontro que seria na sala de aula do professor uma turma de alunos do terceiro ano do Ensino Médio, durante a preparação do professor feita pelo próprio pesquisador este professor em paralelo trabalhava estes conceitos com seus alunos. No final deste processo foram marcados os testes para seus alunos. Estes testes serviram para tentar responder nossa questão de pesquisa que se subdivide em três partes como já foi mostrado. Uma vez terminado o estudo os alunos comunicarão os resultados de sua investigação, tendo o cuidado de preparar argumentos para defender as opções que tomaram e as interpretações que fizeram ao longo do processo de investigação. Para responder a nossa questão de pesquisa, foi proposta uma situação-problema na forma de uma atividade diagnostica qualitativa, dividida em três etapas. Esperamos que esta atividade tenha permitido diagnosticar qual o nível de funcionamento dos conceitos, segundo os preceitos de Robert (1998) e Gal (2002), que trata da Alfabetização Estatística, constituiu nosso quadro teórico Estatístico. e, especificamente, aqueles ligados ao estudo da variabilidade, por parte dos alunos e, também, permitir identificar o(s) possível(is) erro(s) cometido(s) por estes alunos.

Para que pudéssemos analisar os resultados da pesquisa, dividimos as bases elencadas por Gal (2002) em dois níveis. O primeiro focou a Alfabetização, a Estatística e a Matemática, e o segundo focou a análise crítica e global. E, sendo assim, dentro desses níveis, investigamos a mobilização dos conhecimentos dos alunos segundo os preceitos de Robert (1998) quanto a Técnico, Mobilizável e Disponível.

Nos grupos investigados, não conseguimos identificar invariantes que justificasse possíveis dificuldades dos alunos nas duas primeira etapas da atividade, pois os grupos calculavam tudo a té mesmo o que não foi pedido demonstrando assim que não tinham dificuldade para com os cálculos, mas durante os relatórios onde deveriam justificar os cálculos. Somente o grupo 2 conseguiu relacionar os conceitos de forma coerente, mostrando assim que este grupo classifica-se em um nível mobilizável pelos preceitos de Robert (1998), entre os outro grupos em seus relatórios não fizeram uma relação entre os valores encontrados, apesar de estarem todos corretos. Os conceitos mobilizados nos cálculos não tiveram seu significado explicitado ou mesmo justificado pelas duplas, levando-nos a inferir que este conhecimento, se existente, permaneceu implícito. Os alunos permaneciam ligados somente aos valores numéricos, acreditando que estes eram auto-explicativos. Sendo assim, inferimos que as duplas classificam-se em nível técnico no contexto algébrico, segundo Robert (1998).

Durante as atividades percebemos, em alguns momentos, que os grupos investigados realizaram análises equivocadas de alguns conceitos, como, por exemplo, confundir média e mediana. Isso foi diagnosticado pelo fato de as duplas ao atribuírem, para qualquer banco de dados, a noção de simetria. Para elas, toda distribuição é simétrica, assim percebemos que, em toda a atividade, as duplas sentiram a necessidade de modelar a amostra para uma distribuição normal.

Podemos perceber que o contrato didático ainda é forte entre os grupos, pois durante a aplicação das atividades, percebemos, por parte dos alunos, a obrigatoriedade de, em primeiro lugar, calcular-se a média e o desvio-padrão, sem se dar conta de verificar, até por uma análise superficial da amostra, se realmente haveria a necessidade desses cálculos. Por exemplo, a última atividade que nós apresentamos para os grupos não exigia iniciar os cálculos pelas medidas de tendência central, e sim que fizessem uma análise visual do gráfico e, depois, tentassem explicar o comportamento dos dados informados pelos dois gráficos. Mas, no entanto, mostraram uma série de cálculos para tentar fazer uma justificativa. Fica evidente, nesse caso, a necessidade de que a alfabetização estatística deva ser contemplada na sua totalidade nas propostas de Gal (2002): a análise crítica e global dos dados coletados. Entendemos que esta análise não precisa ser necessariamente depois dos dados codificados, mas sim uma análise a priori da codificação dos dados, procurando um melhor caminho a seguir.

Nesta atividade à analise é que nenhum dos grupos saiu do nível técnico segundo Robert (1998)

Assim a idéia principal deste trabalho é de buscar mostrar que o professor de matemática traz falhas de sua formação e estas falhas com o passar do tempo vão se tornando cada vez mais grave. Veja por exemplo, o professor que nos ajudou na pesquisa, formado a pouco tempo mais segundo ele mesmo não aprendeu quase nada em estatística , e sabemos que para a sociedade moderna é fundamental para o indivíduo uma compreensão da estatística, imagine então para o professor formado antes do PCN de (1997), e que nunca mais teve a oportunidade de estudar ou fazer cursos de formação continuada para aprender sobre estatística certamente os mais prejudicados, serão os alunos deste professor que não aprenderão corretamente os conceitos necessário para interpretar as informações que nós são transmitidas a todo os instantes. Conseguimos fazer com que o professor envolvido melhora-se suas percepções sobre estatística e consequentemente, conseguimos fazer com que os alunos deste professor também pudesse expandir seus olhares a respeito dos conceitos básicos da estatística, nas análises dos grupos pudemos ver que existi um grupo que conseguiu uma percepção maior a respeito dos conceitos estocásticos elementares, mais certamente pesquisas futuras poderão apoiar-se nestes resultados para fazer com que mais e mais alunos e professores, também possa incorporar tais conceitos.

Notamos assim que o professor, quando recebe uma ajuda (tanto conceitual como pedagógica) tem condições de trabalhar os temas em questão com seus alunos. O que ficou evidente neste trabalho é que os alunos terão as mesmas dificuldades que seus mestres. No caso, a análise interpretativa dos resultados obtidos com os cálculos.

O desenrolar deste trabalho, proporcionou-nos novas perspectivas para o ensino da Estatística. Uma delas é a necessidade de elaborar programas eficientes de treinamento de professores para poder trabalhar melhor com este tópico tão importante, a necessidade da elaboração de uma seqüência didática que permita ao aluno vivenciar as fases necessárias para a construção de um conceito e sua mobilização.

REFERÊNCIAS BIBLIOGRÁFICAS

BATANERO, C. (2001). *Didáctica de la Estadística, Granada*: Grupo de Investigación en Educación Estadística. ISBN 84-699-4295-6.

__________. (En prensa). *Aleatoriedad, Modelización, Simulación. Presentado en las X Jornadas sobre el Aprendizaje y la Enseñanza de las Matemáticas.* Zaragoza, 2001.

__________. *Los Retos de la Cultura Estadistica.* Jornadas Interamericanas de Ensênaza de la Estadistica. Buenos Aires, 2002. Conferencia inaugural.

BATANERO, C. ;Y GODINO, J. D. *Análisis de Datos Y Su Didática.* Reprografia de la Faculdad de Ciências Univerdidad de Granada, 2001b.

BROUSSEAU, G. *Les Obstacles épistemologiques et les problémes em Matématiques.* RECHERCHES EM DIDATIQUES DES MATHEMÁTIQUES. La Pensée Sauvage-Editions, v.4.2, pp. 165-198. Grenoble, 1983.

BRASIL – Secretaria da Educação. Parâmetros Currículares Nacionais. Matemática. Brasília: MEC/SEF, 1998.

BIFI, CALOS RICARDO. Estatística em um Curso de Administração de Empresas (2006).

BARBETA, PEDRO ALBERTO. Estatística Aplicada a Ciências Sociais – 5 Edição. Editora da UFSC- 2003.

CARVALHO, C. & CÉSAR, M. Interações entre Pares e Estatísticas: Contributos para o Estudo do Conhecimento Instrumental e Relacional. QUADRANTE Revista Teórica e de Investigação. Vol 10, N. 1 – 2001.

CASTRO, Lauro S. Viveiros. Pontos de Estatística. 17 ed. Rio de Janeiro: Científica, 1975.

CHIAVENATO, Idalberto. *Administração de Empresas: Uma abordagem contigencial:* 3 Ed. São Paulo. Makron Books, 1995.

COUTINHO, C.Q.S. (2001). *Introduction aux situations aléatoires dès le Collège: de la modélisation à la simulation d'expériences de Bernoulli dans l'environnement informatique Cabri-géomètre II.* Tese de doutorado. Université Joseph Fourier. França.

CROSSEN, C. *O Fundo falso das pesquisas: a ciência das verdades torcidas.* Rio de Janeiro: Ed. Revan. (1996)

CURCIO, F.R. (1987). *Comprehension of mathematical relationships expressed in graphs.* Journal for Research in Mathematics Educations, 18, 382-393.

________. (1989). *Developing graph comprehension.* Reston, VA: N.C.T.M.

DOWNING, Douglas; CLARK, Jeffrey. *Estatística Aplicada. Série essencial. São Paulo: Saraiva, 1998.*

DRUCKER, Peter F. *Administrando em tempos de grandes mudanças 5ª Ed.* São Paulo: Pioneira, 1998.

Escola Básica 2.3 de SÃO JULIÃO da BARRA.

www. Educ.fc. ul.pt/docentes/jponte/mem/textos/sousa02.pdf.

GAL, I (2002) *Adult´s Statistical Literacy: Meanings, Components, Responsibilities* - Appeared in: Internacional Statistical Review, 2002,70 (1), 1-25.

________ (1995) *Statistical tools and Statistical Literacy: The case of the average.* Teaching Statistics, 17(3), 97-99.

GARFIELD, JEAN, Universidade de Minnesota, Jornal de Educação Estatística volume 10, n.3 (2002)

HAND, D. J. (1998). Breaking Misconceptions *- Statistics and its Relationship to Mathematics. The Statistician, 47 (2), 245-250.*

HEID, M. K et al. *Advanced Mathematical Thinking: Implications of various perspectives on advanced mathematical thinking for mathematics education reform*, acessado em 05 de setembro de 2004 em http://www.DefSOnat.psu.edu/stafftmlklmkh2lAMTheid.00f

LOPES, C. (2004). *Literária estatística e INAF 2002. In Fonseca M. C. (org) Letramento no Brasil – habilidades matemáticas. Ed Global. Pp. 187-197.*

LOPES .C. A. E. A probabilidade e a estatística no Ensino Fundamental: uma análise curricular. Dissertação (Mestrado em Educação Matemática) Faculdade de Educação. Universidade de Campinas. Campinas, 1998.

MACHLINE, Claude et al. *Manual de Administração da produção*, 8 Ed. Rio de Janeiro: FGV, 1994, V.2.

MILONE, Giuseppe; ANGELINI, Flávio. *Estatística Geral, Vol. 1 e 2*, São Paulo: Atlas, 1995.

MIRADOR, *Enciclopédia Internacional*, 8V. São Paulo: Britânia, 1989.

MOORE (1997). *New Pedagogy and new content: the case of statistics. Internacional Statistical Review, 65(2), 123-137.*

MINISTÉRIO DA EDUCAÇÃO. *Secretaria da Educação Média e Tecnológica. Parâmetros Curriculares Nacionais. Brasília: Ministério da Educação, 1999.*

NOVAES, Diva Valério. *A Mobilização de Conceitos estatísticos: Estudo exploratório com Alunos de um Curso de Tecnologia em Turismo.* Dissertação de Mestrado em Educação Matemática – PUC/SP – 2004.

PEREIRA, B. B. (1997). *Estatística: A tecnologia da Ciência.* Boletim da Associação Brasileira de Estatística, ano XIII, no. 37, 2º Quadrimestre.

PARKER, M.,& LEINHARDT,G. (1995). *Percent: A priviledged proportion.* Review of Educational Research, 65(4), 421-481.

POTTER, A. M. *Statistics for Sociologists: Teaching Techniques that Work.* Teaching Sociology, v. , 23, july, 1995.

ROBERT, A. *Outis D'analyse dês Contenus Mathématiques á ensiigner au lycée á l'Université.* RECHERCHES EM DIDACTIQUE DÊS MATÉMATIQUES, Vol. 18, nº2, pp. 139-190. 1998

ROITER, K., PETOCZ, P. *Introductory statistics courses-A new way of thinking.* Journal of Statistics Education., v. 4, n. 2, l996.

RUBERG, S. J. e MASON, R. L.. *Increasing public awareness of Statistics as a science and profession starting in high school.* The American Statistician, 42 (3), 167-170. (1988)

SALIBY, E. *Software para Simulação, [S.l.:s.n.].* jun. 2000. Disponível em: <http://www.coppead.ufrj.br/pesquisa/cel/po_psal.htm>. Acesso em: 09 jun. 2000;

SCHANK, R. *Using Simulators to Teach. [S.l.:s.n.],* jul. 2000. Disponível em: <http://www.ils.nwu.edu/~e_for_e/nodes/NODE-125-pg.html>. Acesso em: 20 jul. 2000.

SILVA C., CAZORLA, IRENE, BRITO, FERREIRA M. R. *Concepções e Atitudes em relação à Estatística.* In: Conferência Internacional "Experiências e Perspectivas do Ensino da Estatística: Desafio para o século XXI", 1999. Anais da Conferência Internacional "Experiências e Perspectivas do Ensino da Estatística: Desafio para o século XXI", 1999 p. 18-29, 1999.

SILVA, C. B. *Atitudes em relação à Estatística: Um Estudo com alunos de graduação.* Dissertação de Mestrado. UNICAMP/SP, 2000

SFARD, A. & LINCHEVSKI, L. (1994a). *The gains and the pitfalls of reitification: The case of algebra.* Educational Studies in Mathematics, 26, 191-228.

SNEE, R. D. (1988). *Mathematics is only one tool that Statistician use*. The College Mathematics Journal, 19, 30-32.

SPRENT, P. (1998). *Statistics and Mathematics - Trouble at the Interface?* The Statistician, 47 (2), 239-244.

STENVENSON, Willian J. *Estatística Aplicada a Administração. São Paulo.* Harbra, 1986.

STUART, M. (1995). *Changing the teaching of Statistics.* The Statistician, 44 (1), 45-54.

TOLEDO, Luciano Geraldo; OVALLE, Ivo Izidoro. *Estatística Básica.* 2 Ed. São Paulo: Atlas, 1995.

Triviños, A.N.S, Introdução à pesquisa em Ciências Sociais: a pesquisa qualitativa em educação São Paulo: Atlas, 1987.

VENDRAMINI, C.M.M., SILVA, C. B., CAZORLA, I. M. & BRITO, M. R. F. (2000). Análise da relação entre os aspectos afetivos e cognitivos com desempenho em Estatística. Anais do 14º Simpósio Nacional de probabilidade e Estatística. Caxambu, Minas Gerais, 536.

VENDRAMINI, C. M. M. *Implicações das Atitudes e das Habilidades Matemáticas na Aprendizagem dos Conceitos Estatísticos.* Tese de Doutorado, UNICAMP/SP, 2000.

VENDRAMINI, C.M.M., SILVA, M. C., & CANALE. M. (2003). *Análise de itens de uma prova de raciocínio estatístico.* Artigo submetido à revista Pscicologia em Estudo, Maringá.

VENDRAMINI, C. M. M., CHENTA, V. C., & SILVA, L. S. (2004). *Leitura de dados estatísticos: Um estudo com alunos do ensino fundamental.* Texto não publicado.

VENDRAMINI, Claudete Maria Medeiros & Brito, Márcia Regina Ferreira de. Relações entre atividade, conceito e utilidade da estatística. Psicol. Esc.,jun. 2001, vol.5, nº.1, p.59-73. ISSN 1413-8557

VERGNAUD, G. *A Comprehensive Theory of Representation for Mathematics Education.* IMB V17, N2, pp. 167 – 181, 1998.

VLAHOS, Kiriakos. *Minimizando o risco quanto a incerteza. In: Dominando a Administração:* Financial Times: São Paulo: Makron Brooks, 1999.

WAINER, H. (1992). *Understanding Graphs and Tables*. Educational Researcher, 21(I), 14 –23.

WALLMAN, K. K. (1993). *Enhancing Statistical Literacy: Enriching our society.* Journal of the American Statistical Association, 88, 1-8.

WERKEMA, Maria Cristina C. *Como estabelecer conclusões com confiança: entendendo a inferência estatística.* Belo Horizonte: Fundação Christiano Ottoni, Escola de Engenharia – UFMG, 1996.