

Pós-Graduação em Ensino de Ciências Naturais e Matemática

REVISTA

DYNAMIS

ESTUDO DE PROPOSTAS DIDÁTICAS PARA O CONCEITO DE VARIABILIDADE ESTATÍSTICA: UTILIZANDO “PAPEL E LÁPIS” E O SOFTWARE “R”

Study of Didactic Proposals for the Concept of Statistics Variability: using "paper and pencil" and software "R"

Ailton Paulo de Oliveira Júnior

Doutor e Pós Doutor em Educação

Programa de Pós-Graduação em Ensino e História das Ciências e da Matemática – PEHCM

Universidade Federal do ABC

ailton.junior@ufabc.edu.br

Flávia Helena Pereira

Mestranda do Programa de Pós-Graduação em Educação – PPGED

Universidade Federal do Triângulo Mineiro

fh_pereira18@yahoo.com.br

Resumo

Acreditamos que enfatizar o pensamento estatístico deve ser considerado para que o aluno entenda que o processo estatístico tem como elementos básicos, a necessidade de dados, bem como a importância da análise, quantificação e explicação da variabilidade destes dados. O presente trabalho teve como objetivo avaliar o entendimento de alunos da Educação Superior do conceito de variabilidade estatística utilizando a representação gráfica das distribuições de frequência e dos conceitos das medidas de tendência central e de dispersão. Foram utilizadas duas metodologias de ensino (“utilizando papel milimetrado e lápis” e “utilizando o *software* R”) em sequências didáticas para a apresentação dos conceitos básicos introdutórios de medidas que são necessárias para melhor entendimento da distribuição de frequência. Verificou-se que o uso do *software* foi aceito pelos participantes e que também é importante considerar o uso do papel e lápis na construção dos gráficos. Ao realizar comparações de dois ou mais conjuntos de dados e examinar seus gráficos na mesma escala permitiu especular de que forma esta variação explica os dados. O grau de compreensão dos participantes sobre a variabilidade convergiu para a maneira pela qual perceberam a distribuição de dados.

Palavras-chave: Ensino de Estatística; Variabilidade; Propostas Didáticas; Educação Superior.

Abstract

We believe that emphasizing statistical thinking should be considered so that the student understands that the statistical process has as basic elements, the need for data, as well as the importance of the analysis, quantification and explanation of the variability of these data. The present work had as objective to evaluate the understanding of Higher Education students of the concept of statistical variability using the graphic representation of frequency distributions and the concepts of measures of central tendency and dispersion. Two teaching methodologies ("using graph paper and pencil" and "using software R") were used in didactic sequences to present the basic introductory concepts of measures that are necessary for a better understanding of the frequency distribution. It was found that the use of the software was accepted by the participants and that it is also important to consider the use of paper and pencil in the construction of the graphics. By comparing two or more datasets and examining their graphs on the same scale, we can speculate how this variation explains the data. The participants' degree of understanding of variability converged to the way in which they perceived the data distribution.

Keywords: Teaching statistics; Variability; Didactic proposals; Higher Education.

1 INTRODUÇÃO

Por vezes nos deparamos com informações em nosso cotidiano (seja na escola, no trabalho, mesmo que seja assistindo o jornal na TV) sem saber que podemos estruturá-las a fim de estabelecer relações entre as informações apresentadas ou seus dados.

A sociedade de hoje exige que as pessoas que são competentes para analisar, compreender e interpretar informações de vários eventos que acontecem ao seu redor entenda que várias situações envolvem elementos de incerteza ou variabilidade (por exemplo, eventos climáticos, eventos desportivos, jogos de azar, previsão financeira, seguros de vida ou danos). Como resultado, questões estatísticas têm mostrado presença no currículo da matemática escolar em muitos países, desde o ensino primário ao nível universitário.

Portanto, existem conceitos e medidas que nos permitem estabelecer maneiras mais práticas e fáceis de interpretar essas informações, e se forem estudadas adequadamente podem ser aplicados de alguma forma em situações em que vivemos, ou estudamos.

Desta forma, a variabilidade é o coração da estatística (MOORE, 1990) e indispensáveis para a sua existência (WATSON et al., 2003).

E Wild e Pfannkuch (1999) propõem o estudo da variabilidade como um dos tipos de pensamento estatístico relacionados com a aprendizagem e a tomada de decisão.

Este conceito está relacionado com muitas ideias estatísticas fundamentais, sendo que Garfield e Ben-Zvi (2008) observaram que a compreensão das ideias de dispersão e variabilidade dos dados é componente chave para a compreensão do conceito de distribuição e é essencial para fazer inferências estatísticas.

Konold e Pollatsek (2004) recomendam que o centro de análise de dados (sinal) e variabilidade (ruído) seja considerado em conjunto, porque ambas as ideias são necessárias para dar sentido aos dados.

Para Watson, Callingham e Kelly (2007) a compreensão da variabilidade está relacionada com a percepção de incerteza, ou seja, com as mudanças e valores atípicos esperados e inesperados que os dados apresentam.

A compreensão da variabilidade não é livre de dificuldades (GARFIELD et al., 2005), e expressam que não tem sido contemplada de modo curricular e didático, além de ser difícil para os alunos equilibrarem as visões informais e formais na abordagem do conceito. Estes apresentam dificuldade em passar da ideia intuitiva de que os dados variam (diferenças entre os dados) para lidar formalmente com essa variação por meio de uma medida (amplitude total dos dados e o desvio padrão) e ainda mais para combinar ambos em uma interpretação coerente.

Não menos substancial para a Educação Estatística nos parece a intervenção de produtos de tecnologia computacional, por seu potencial para promover a parte qualitativa da própria análise estatística e contribuir para o desenvolvimento conceitual.

Em particular, isso está relacionado a habilidades como a mobilização de diferentes objetos estatísticos ou matemáticos através de representações numéricas, tabulares, gráficas e simbólicas.

De acordo com Ponte, Brocardo e Oliveira (2006) as novas tecnologias não são apenas uma ferramenta útil para o trabalho em Estatística, mas constituem como um importante elemento na prática desse campo.

Ben-Zvi (2011) ao refletir sobre os elementos-chave para o ambiente de aprendizagem para o raciocínio estatístico revela que um desses elementos é a integração de ferramentas tecnológicas no ensino. Para o autor, existe uma gama de ferramentas tecnológicas que podem auxiliar no desenvolvimento do raciocínio e da compreensão dos estudantes sobre a Estatística, como computadores, calculadoras gráficas, Internet, *software* estatístico e aplicativos da Web.

Castro L., Gutiérrez F. e Hugues G. (2014) apresentaram proposta que aborda a variabilidade estatística através de uma sequência de atividades apoiadas pelo uso da tecnologia computacional, considerando que é um tema central da Educação Estatística e que, na sua promoção, a tecnologia pode desempenhar um importante papel didático.

Portanto, há de se considerar que, por meio da tecnologia, esse processo pode ser facilitado, pois muitas ferramentas computacionais trazem importantes ganhos para o desenvolvimento do ensino e a aprendizagem da estatística.

2 UTILIZANDO O *SOFTWARE* “R” NO ENSINO DE ESTATÍSTICA

Caso pergunte a uma pessoa que tenha feito um curso de estatística sobre a sua impressão da matéria, possivelmente irá ouvir que a mesma foi desenvolvida com o uso de uma série de fórmulas que tinham de memorizar e cálculos tediosos para chegar a um resultado. No entanto, com a ajuda da tecnologia, poderia ter sido minimizado este trabalho e tempo para a manipulação dos cálculos com a ajuda de *softwares* ou planilhas eletrônicas.

Batanero, Godino e Cañizares (2005), salientam que os computadores oferecem aos estudantes uma variedade de ferramentas para simulação que auxiliam na exploração de conceitos e ideias que de outra forma seriam limitados ou mesmo inviáveis.

Borovcnik e Kapadia (2009) defendem a simulação enquanto estratégia de ensino e afirmam que, quando associada ao uso da tecnologia, auxilia na redução de cálculos técnicos, ao mesmo tempo em que possibilita ao aluno dedicar atenção aos conceitos em discussão.

Existe consenso entre educadores, de que disciplinas de estatística devem ser acompanhadas por algum tipo de tecnologia para diminuir a realização de cálculos manuais e permitir aos alunos o acesso a conjuntos de dados de casos práticos e situações reais (BECKER, 1996; MALTBY, 2001; SCHUYTEN; THAS, 2007; VERZANI, 2008; GOULD, 2010).

O que, no entanto, ainda não há consenso é sobre o *software* estatístico apropriado para este fim (VERZANI, 2008). Assim, uma possibilidade é o uso de *softwares* acessíveis e sem custo como, por exemplo, o R (R DEVELOPMENT CORE TEAM, 2011).

O R é uma linguagem e um ambiente para computação estatística. É um projeto *open source*, baseado no conceito de *software* livre e pode ser usado sem custos de licença, tanto para estudantes quanto para universidades e além de possuir versões para *Windows*, *MacOS*, *GNU/Linux* e *Unix*, têm uma extensa coleção de pacotes adicionais, também gratuitos.

O “R” como é conhecido por seus usuários, é um programa de domínio público e uma ferramenta importante que facilita a elaboração de diversos tipos de gráficos, no qual o usuário possui controle sobre o que cria.

O processo de instalação é rápido e fácil, com a vantagem de ser gratuito, podendo ser realizado por qualquer usuário da internet.

Além do “R” ser um ambiente de programação, pode desenvolver novas implementações como a manipulação dos gráficos e geração de estatísticas básicas para apresentação e análise dos dados coletados neste trabalho.

3 PROCEDIMENTOS METODOLÓGICOS

Tendo em vista o tema, o objetivo fixado deste trabalho foi o de avaliar a aprendizagem do conceito de variabilidade estatística utilizando a representação gráfica das distribuições de frequência e as medidas de tendência central e de dispersão. Para tanto, foram utilizadas duas metodologias de ensino (“utilizando papel milimetrado e lápis” e “utilizando o software R”) em uma sequência didática para a apresentação dos conceitos básicos introdutórios de medidas que são necessárias para o melhor entendimento da distribuição de frequência.

3.1 AMOSTRA

As atividades foram desenvolvidas em dois encontros de 4 horas em junho de 2016 com 16 participantes, quais sejam: bolsistas do PET (Programa de Educação Tutorial) e do PIBID (Programa Institucional de Bolsas de Iniciação à Docência) da Universidade Federal do Triângulo Mineiro - UFTM; e alunos do Mestrado em Educação da UFTM participantes do Grupo de estudos em Educação Estatística e Matemática - GEEM.

3.2 INSTRUMENTO

Foram elaborados e aplicados os seguintes instrumentos:

- (1) Uma sequência didática utilizando a construção gráfica manual de uma distribuição de frequências e as respectivas fórmulas das medidas de tendência central e de dispersão (utilizando papel milimetrado e lápis);
- (2) Uma sequência didática utilizando o *software* “R” para mostrar a representação gráfica das distribuições de frequência, bem como a geração das medidas de tendência central e de dispersão para apreensão do conceito de variabilidade;
- (3) Questionário para avaliação da atividade no tocante às sequências didáticas desenvolvidas.

E ainda apoiamo-nos na definição de sequência didática de Pannuti (2003) que é uma modalidade organizativa que se constitui numa série de ações planejadas e orientadas com o objetivo de promover uma aprendizagem específica e definida. Estas ações são sequências de forma a oferecer desafios com o grau de complexidade crescente, para que se possam colocar em movimento as habilidades dos alunos, superando-as e atingindo novos níveis de aprendizagem.

3.3 COLETA DE DADOS

As atividades foram desenvolvidas e conseqüentemente os dados coletados em junho de 2016, em um dos laboratórios de informática da universidade.

Antes do início das atividades foi apresentado o objetivo do estudo, bem como a voluntariedade e anonimato em relação à participação na pesquisa.

Explicamos aos participantes da pesquisa que utilizaríamos duas metodologias de ensino, sendo a primeira com a utilização de papel milimetrado e lápis para a solução de um problema estatístico e para outro problema seria utilizado o *software* R.

Aqueles que aceitaram participar foram orientados a não identificar o material e assinar o termo de consentimento livre e esclarecido (TCLE).

As atividades foram realizadas e ao final da mesma foi distribuído instrumento (questionário) para que os participantes avaliassem a proposta.

3.4 ASPECTOS ÉTICOS DA PESQUISA

Este trabalho é contemplado pela aprovação do projeto junto ao Comitê de Ética em Pesquisa (CEP) da Universidade Federal do Triângulo Mineiro sob o Certificado de Apresentação para Apreciação Ética – CAAE, n. 38076114.9.0000.5154. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

3.5 MÉTODO

Pretendeu-se avaliar o entendimento dos participantes da pesquisa pelo desenvolvimento das atividades, considerando as duas sequências didáticas que abordaram a construção gráfica de uma distribuição de frequências e os conceitos das medidas de tendência central e de dispersão.

Portanto, para avaliar e apresentar os resultados obtidos na sequência didática, tomamos Garfield e Ben-Zvi (2005), que propõem um quadro de sete categorias ou "componentes de um modelo epistemológico" que organizou o conjunto de ideias e ações necessárias para uma compreensão completa de variabilidade. Selecionamos as seguintes três primeiras categorias como as mais convenientes para analisar o nível cognitivo do conceito de variabilidade:

- 1) Desenvolver ideias intuitivas de variabilidade, que se referem a ela como uma característica onipresente nos fenômenos e dados obtidos a partir deles. Além disso, há fenômenos e conjuntos de dados que têm mais variabilidade do que outros e, finalmente, é uma característica da variabilidade de um conjunto de dados e não de os elementos isolados.
- 2) Descrever e representar a variabilidade: Quando restrita a um conjunto de dados, a variabilidade é a dispersão destes que pode ser representado graficamente para mostrar seus aspectos. A variabilidade pode ser medida de diferentes maneiras e expressa por um número (amplitude, desvio médio ou desvio-padrão).
- 3) Usar a variabilidade para fazer comparações, onde as medidas de tendência central e de

dispersão permitem a interpretação dos dados e, conseqüentemente, a tomada de decisão.

Para o estudo das questões do questionário, quanto à forma de abordagem, pretendeu-se realizar uma reflexão a partir dos dados experimentais e respostas ao questionário de avaliação das atividades, na tentativa de explicar o contexto no âmbito dos participantes.

Para o estudo em questão, quanto à finalidade, é considerada como uma pesquisa básica ou fundamental com o objetivo de adquirir conhecimentos sobre como os participantes se relacionam com o conceito de variabilidade dos dados.

E, quanto à sua natureza a consideramos como uma pesquisa experimental onde pode-se participar ativamente na condução do processo, ao poder selecionar as variáveis estudadas, definindo a forma de controle sobre elas e observando os efeitos sobre o objeto de estudo, em condições pré-estabelecidas.

3.6 PROCEDIMENTOS

Foram construídos gráficos (Histograma e Polígono de Frequências) utilizando papel milimetrado e lápis num problema proposto (sequência didática) e outro problema utilizando o *software* “R”.

Com a geração dos gráficos, pretendemos estudar a variabilidade dos dados, considerando a distribuição de dados e apoiadas pelas medidas de tendência central e de dispersão. No *software* R pode-se plotar o histograma e o polígono de frequências ao utilizar a função *density* (Figura 1).

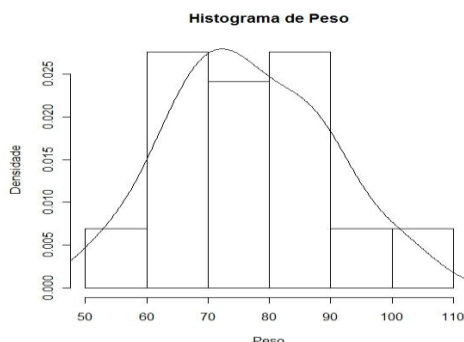
Figura 1 - Utilização da função *density* para gerar o polígono pela densidade de frequência no *software* “R”.

```
##histograma pela densidade de frequência  
hist(P,freq=FALSE)  
h1=density(P) ##obtendo a densidade dos dados  
lines(h1) ##adicinar um gráfico de linhas
```

Fonte: Elaborado pelos autores.

A Figura 2 apresenta exemplo de resultado gráfico (Histograma e Polígono de Frequências) utilizando o *software* R.

Figura 2 - Polígono pela densidade de frequência gerada pelo *software* “R”.



Fonte: Elaborado pelos autores.

Ao estudarmos a distribuição de frequências de uma variável quantitativa, seja em apenas um grupo ou comparando vários grupos, verificamos basicamente três características: tendência central; variabilidade; e forma.

A tendência central da distribuição de frequências de uma variável é caracterizada pelo valor (ou faixa de valores) típico da variável. Uma das maneiras de representar o que é típico é através do valor mais frequente da variável, chamado “moda”. Ou, no caso da tabela de frequências, a classe de maior frequência, chamada de classe modal. No histograma, esta classe corresponde àquela com barra mais alta ("pico").

Para descrever adequadamente a distribuição de frequências de uma variável quantitativa, além da informação do valor representativo da variável (tendência central), é necessário dizer também o quanto estes valores variam, ou seja, o quão disperso eles são. De fato, somente a informação sobre a tendência central de um conjunto de dados não consegue representá-lo adequadamente.

A distribuição de frequências de uma variável pode ter várias formas, mas existem três formas básicas (simétrica; assimétrica à esquerda; e assimétrica à direita); podendo ser visualizadas pela construção de histogramas e o respectivo polígono de frequências.

Assim, ao definir a assimetria de uma distribuição, pode-se referir ao lado onde está a concentração dos dados ou ao lado onde está faltando dados (cauda). As duas denominações são alternativas.

Em alguns casos, apenas o conhecimento da forma da distribuição de frequências de uma variável já nos fornece uma boa informação sobre o comportamento dessa variável.

E ainda no apoiamos em Snell e Peterson (1992), que identificou três grandes áreas em que a tecnologia pode ser útil no ensino da estatística: (1) Reduz os cálculos manuais; (2) Facilita a análise gráfica dos dados; (3) Ilustra os conceitos pela simulação.

4 ESTUDO DO DESENVOLVIMENTO DAS ATIVIDADES

Iniciamos a sequência didática propondo uma atividade “roteiro”, que foi entregue aos participantes para que resolvessem um problema com a participação dos pesquisadores.

A atividade apresentada na Figura 3 apresenta o problema inicial proposto aos participantes trazendo um assunto atual e contextualizado ao povo brasileiro, na medida em que foi realizado no ano de 2016 os Jogos Olímpicos na cidade do Rio de Janeiro.

Foram construídos o histograma (dados agrupados) e o polígono de frequências dos dados, bem como calculado os valores da média, desvio padrão, medidas separatrizes e os coeficientes de variação, assimetria e curtose. Então,

Na sequência discutimos os resultados obtidos das medidas estatísticas e das representações gráficos. O objetivo foi que os participantes pudessem apreender os conceitos que advém destas medidas e representações gráficas.

Figura 3 - Problema “roteiro” para desenvolvimento da atividade.

Em cerca de 2500 a.C., os gregos realizavam festivais esportivos em honra a Zeus, no santuário de Olimpia, o que originou o termo olimpíada. As Olimpíadas acontecem de quatro em quatro anos, onde atletas de centenas de países se reúnem num país sede para disputarem um conjunto de modalidades esportivas. A própria bandeira olímpica representa essa união de povos e raças, pois é formada por cinco anéis entrelaçados, representando os cinco continentes e suas cores. A paz, a amizade e o bom relacionamento entre os povos são os princípios dos Jogos Olímpicos que tiveram origem na Grécia Antiga. A primeira Olimpíada ocorreu no estádio Panathenaic em Atenas em 1896, na época da era moderna em Atenas, com a participação de 14 países. A Tabela 1 apresenta lista dos países com maior número de medalhas em toda história das Olimpíadas.

Tabela 1 - Lista dos países com maior número de medalhas na história das Olimpíadas.

Países	Medalhas	Países	Medalhas
Estados Unidos	2548	Coreia do Sul	260
União Soviética	1204	Alemanha Ocidental	243
Grã-Bretanha	802	Bulgária	218
França	730	Cuba	208
Alemanha	719	Dinamarca	171
Itália	628	Tchecoslováquia	168
Suécia	604	Bélgica	144
Alemanha Oriental	519	Equipe Alemã Unida	137
China	518	Equipe Unificada	135
Rússia	496	Espanha	133
Austrália	477	Grécia	110
Hungria	464	Brasil	108
Finlândia	455	Ucrânia	101
Noruega	447	Iugoslávia	94
Canadá	405	Nova Zelândia	87
Japão	397	Bielorrússia	75
Países Baixos	332	Quênia	75
Suíça	308	África do Sul	70
Romênia	293	Argentina	70
Áustria	287	Jamaica	67
Polônia	275	México	62

Fonte: https://pt.wikipedia.org/wiki/Quadro_de_medalhas_dos_Jogos_Ol%C3%ADmpicos.

A partir da Tabela 1, considere a variável “número de medalhas dos países que mais conquistaram na história das Olimpíadas”:

2548	1204	802	730	719	628	604	519	518	496	477	464	455	447	405	397	332	308	293	287	275
260	243	218	208	171	168	144	137	135	133	110	108	101	94	87	75	75	70	70	67	62

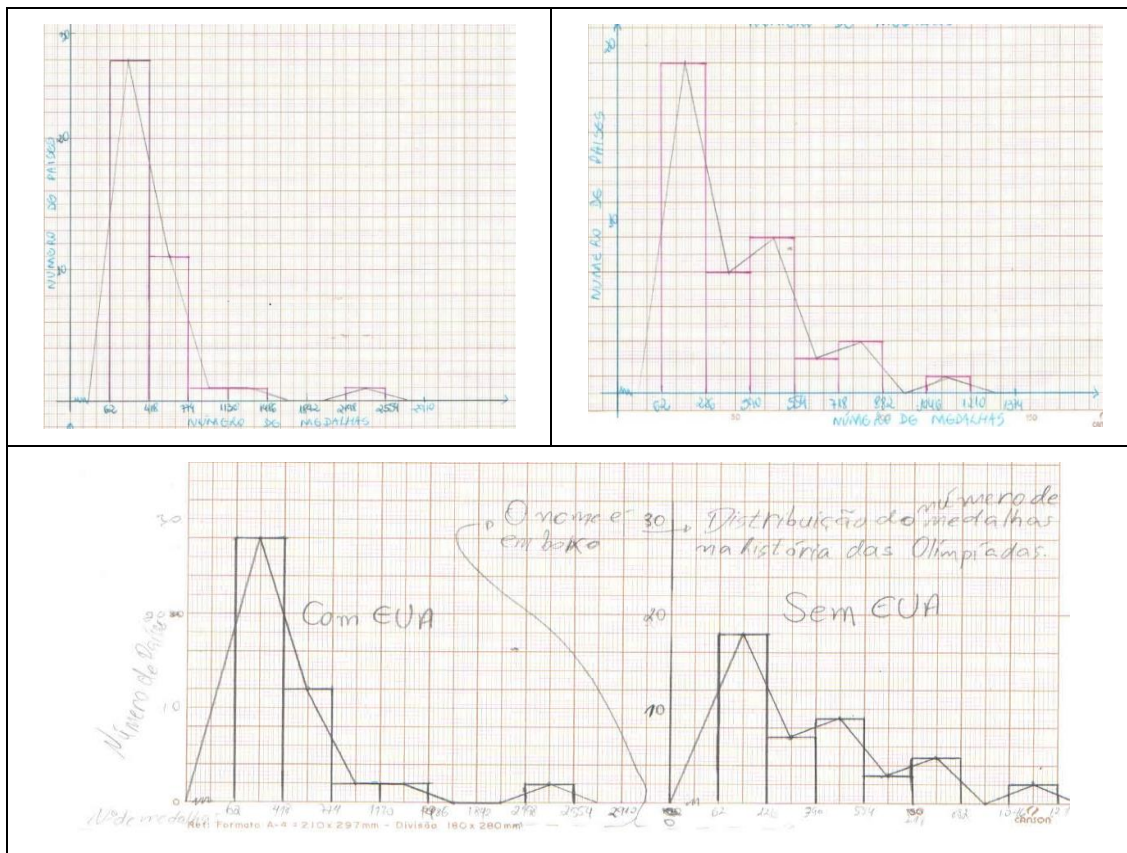
Assim,

- 1) Construa uma distribuição de frequência considerando o número de medalhas conquistadas pelos principais países na história das Olimpíadas, utilizando o método tradicional.
- 2) Esboce o histograma de frequências e o polígono de frequências utilizando o método tradicional e interprete o seu comportamento.
- 3) Interprete a dispersão dos dados e também as medidas de assimetria e curtose e interprete o seu comportamento.

Fonte: Elaborado pelos autores.

A Figura 4 apresenta a solução de dois participantes da atividade em que foi utilizado o papel milimetrado para a construção das distribuições de frequência com todos os dados apresentados na Figura 6 e com a remoção do número de medalhas conquistadas em Olimpíadas pelos EUA.

Figura 4 - Gráficos comparativos da distribuição de frequências em relação à distribuição das medalhas por países na história das Olimpíadas.



Fonte: Elaborado pelos participantes.

Nesta atividade os alunos puderam observar a diferença entre as distribuições de dados e a forma do polígono de frequência quando utilizamos todos os dados e quando removido o número de medalhas referente aos EUA.

O desenvolvimento desta sequência possibilitou aos participantes a visualização do gráfico e determinação de medidas estatísticas. E a partir desta atividade puderam perceber que o tratamento dos dados é de extrema importância para a apropriação do conhecimento do seu comportamento.

Após este momento iniciamos o processo de desenvolvimento das atividades com o *software R*, com a mesma atividade.

Consideramos aqui que com a análise gráfica dos dados, a tecnologia pode auxiliar através de uma variedade de gráficos que permitem que os alunos consigam investigar conceitos-chave através da visualização.

É possível construir, histogramas, diagramas de caixa, gráficos de dispersão, polígonos de frequência, distribuições de probabilidade, distribuições de amostragem, etc.

Também é possível manipular os dados, a fim de observar as mudanças nas representações gráficas, conseguindo assim remover o caráter abstrato para muitos conceitos.

Assim, os participantes receberam um roteiro (Figura 5) com o passo a passo para a utilização do *software R* na resolução da atividade proposta. Também foi orientado que os participantes o baixassem no computador e receberam as orientações sobre como utilizá-lo.

Figura 5 - Roteiro para auxiliar no desenvolvimento da sequência didática utilizando o software “R”.

Apostila passo a passo para a utilização do R na resolução das atividades propostas:

Iniciaremos fazendo o download do programa que podemos encontrar no link:
 1- <https://www.r-project.org/>
 2- CRAN -> Mirror(e.g., Brazil-University of Sao Paulo, Sao Paulo) ? Download and Install R for Windows.
 3- Install R for the first time
 4- Download R 3.2.3 for Windows

Na resolução das atividades:

Para iniciar a distribuição de freqüências é necessário digitar a amostra de dados criando um vetor através dos comandos:
V<-c(126, 133, 149, 163, 187, 195, 196, 203, 232)
 (neste comando, a primeira letra é o nome do vetor, e os dados da amostra devem ser separados por vírgula e espaço e estar entre parênteses, sempre um à frente do outro, independente do tamanho da amostra)

Para gerar a distribuição de freqüências utiliza-se o comando: **table(V)**
 (observe que o R nos mostra a freqüência absoluta, mas não divide a amostra em intervalos de classes)

Para encontrar o número de elementos da série de dados do vetor digitamos: **length(V)**

Para encontrar o valor máximo digitamos: **max(V)**

Para encontrar o valor mínimo digitamos: **min(V)**

A função **range(V)** nos mostra os valores máximo e mínimo

Para calcularmos a amplitude digitamos: **Amplitude = max(V) - min(V)**
 Em seguida digitamos abaixo: **Amplitude**

Para criar o histograma de freqüências, devem ser utilizados os seguintes comandos:
hist(V,main="Nome do histograma",xlab="legenda do eixo x",ylab="legenda do eixo y")
 (na função *hist* serão inseridos o vetor *V*, o nome do histograma dado pelo comando *main* e as legendas dos eixos *x* e *y* dados pelos comandos *xlab* e *ylab*)

Para criar o polígono de freqüências devemos transformar o histograma criado em um vetor, o qual daremos o nome *h*. Utilizaremos os seguintes comandos:
h=hist(V,main="Nome do histograma",xlab="legenda do eixo x",ylab="legenda do eixo y")

Para traçar o polígono de freqüências no histograma devemos utilizar os comandos:
lines(c(min(h\$breaks), h\$midpoints, max(h\$breaks)), c(0,h\$counts, 0), type = "l")

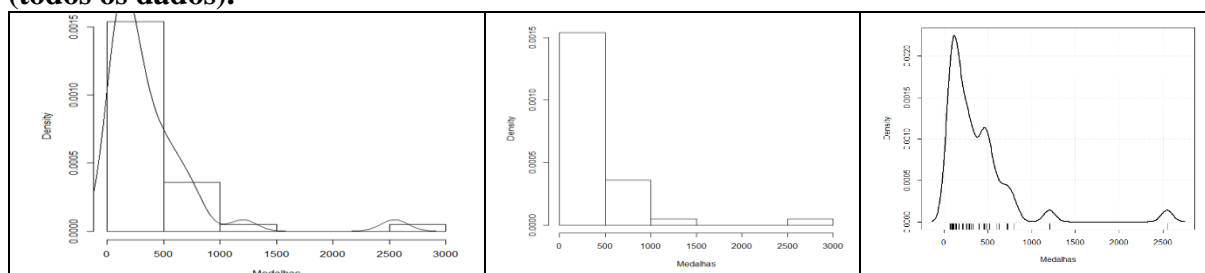
A função **summary(V)** nos mostra a média, mediana, 1º quartil, 3º quartil, mínimo e máximo da variável.

Para o desvio padrão devemos digitar **sd(V)**
 Para encontrar os percentis, digitamos **quantile(V,seq(0.10,0.9,0.1))**

Fonte: Elaborado pelos autores.

E com a utilização do software “R”, inicialmente foram gerados gráficos (histograma e polígono de freqüências), figura 6, considerando as medalhas de todos os países constantes da tabela apresentada na Figura 3.

Figura 6 - Gráficos da atividade com Medalhas dos países na história dos Jogos Olímpicos (todos os dados).



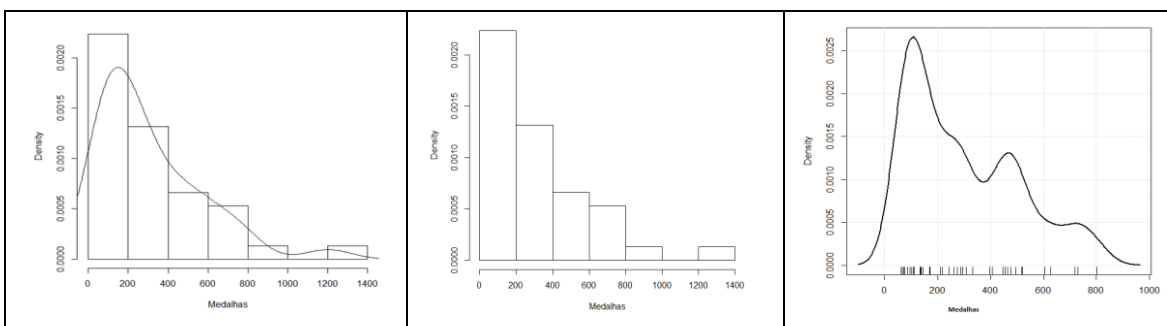
Fonte: Elaborado pelos autores a saída do software R.

Pode-se observar que há assimetria no comportamento dos dados e percebe-se também que há uma grande dispersão dos dados em torno das medidas de tendência central.

Por este fato, foi removido o número de medalhas referente aos EUA (Figura 7). E foi possível analisar o comportamento do polígono com base no histograma, a qual mostra o comportamento dos dados, indicando que há uma menor dispersão dos dados.

O que motivou a continuidade da atividade, com a retiradas de observações que se mostraram visualmente discrepantes, foi acreditar intuitivamente que apresentaria uma menor dispersão dos dados.

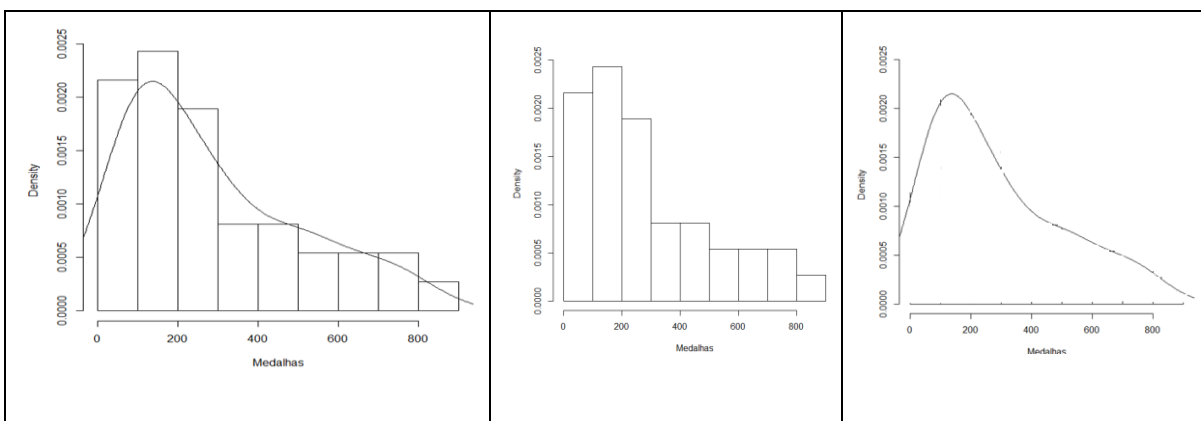
Figura 7 - Gráficos da atividade com Medalhas dos países na história dos Jogos Olímpicos (exceto Estados Unidos).



Fonte: Elaborado pelos autores a saída do *software* R.

E ainda foi proposto retirar o número de medalhas dos Estados Unidos e da antiga União Soviética (Figura 8).

Figura 8 - Gráficos da atividade com Medalhas dos países na história dos Jogos Olímpicos (exceto Estados Unidos e União Soviética).



Fonte: Elaborado pelos autores a saída do *software* R.

A continuidade da atividade mostra-se útil na medida em que se percebe que os dados se mostram menos assimétricos, indicando uma aparência próxima à curva da distribuição Normal de dados (em forma de sino) ou simétrica, onde os valores da média, da mediana e da moda são iguais.

Acreditamos que com esta atividade, pode-se examinar a variabilidade dentro de um grupo (observando como os dados variam dentro de um ou mais conjuntos de dados) e a variabilidade entre os grupos (a variabilidade das medidas usadas para resumir e comparar conjuntos de dados) e distinguir entre esses dois tipos da variabilidade.

Além disso, podemos listar os seguintes aspectos para dar suporte à importância desta atividade:

1. Os gráficos de dados mostram como estes variam e podem revelar padrões para ajudar a descobrir o comportamento da distribuição dos dados e identificar aspectos que possam causar problemas para a sua identificação;
2. Diferentes gráficos podem revelar diferentes aspectos da variabilidade de um conjunto de dados, por isso é importante estudar mais do que um único gráfico de um conjunto de dados.

Acreditamos ser essencial solicitar aos participantes desenhar gráficos de distribuições para o mesmo conjunto de dados e/ou diferentes conjuntos de dados, já que com este procedimento, permite-se visualizar a variabilidade dos dados.

Após trabalhar com histogramas e polígonos de frequência, pode-se observar que a distribuição dos dados pode assumir várias formas. E em algumas distribuições tendem a agrupar os dados mais para uma parte da distribuição do que a outro (assimetria).

Mas, também é possível analisar as distribuições dos dados de modo a obter medidas numéricas ou estatísticas, podendo nos ajudar na análise das características dos dados. Dois desses recursos são de particular importância para a tomada de decisão: a tendência central e de dispersão.

Portanto, no *software* R, também pode-se obter medidas de tendência central e de dispersão ao utilizar as funções *summary* e *quantile* (Figura 9) e que dão suporte ao estudo do conceito de variabilidade.

Figura 9 - Funções *summary* e *quantile* para gerar medidas de tendência central e de dispersão no *software* “R”.

```
> numSummary(EX4[,"X802"], statistics=c("mean", "sd", "IQR", "quantiles"),  
+ quantiles=c(0,.25,.5,.75,1))  
> quantile(EX4[,"X802"],seq(.1,.9,.1))
```

Fonte: Elaborado pelos autores.

A Tabela 1 apresenta as medidas de tendência central e de dispersão dos dados apresentados na Figura 3 e que foram gerados a partir do *software* R com as funções apresentadas na Figura 9.

Tabela 1 - Medidas estatísticas (tendência central e de dispersão) geradas pelo software “R” para o problema do número de medalhas.

Medidas estatísticas	Todos os dados	Exceto EUA	Exceto EUA e União Soviética
n	42	41	40
Média	372,48	319,41	297,30
Desvio	424,87	252,63	211,88
Mínimo	62,00	62,00	62,00
Percentil 10	71,50	71,00	70,50
1º Quartil	109,50	109,00	108,50
Mediana	267,50	260,00	251,50
3º Quartil	481,75	470,50	461,75
Percentil 90	726,70	700,80	625,60
Máximo	2548,00	1204,00	802,00
IQR*	186,13	180,75	176,63

*Intervalo Interquartilício.

Fonte: Dados organizados pelos autores a partir da saída do software R.

Ao contrário das medidas de tendência central que medem a representação em um único ponto, as medidas de dispersão medem o grau de separação e alienação que tem uma variável estatística em torno de uma medida de posição ou de tendência central. Entendem-se as medidas de dispersão ou de variação como o grau em que os dados numéricos tendem a se distribuir ao entorno de um valor central ou de tendência central.

Portanto, na Tabela 1, quando se retira o número de medalhas dos Estados Unidos e posteriormente o número de medalhas da antiga União Soviética, pode-se observar que:

- 1) Há uma diminuição do valor médio de medalhas por país: ($\bar{x}_t = 372,48$ medalhas; $\bar{x}_{(t-EUA)} = 319,41$ medalhas; $\bar{x}_{(t-EUA-URSS)} = 297,30$ medalhas);
- 2) Há uma diminuição do desvio padrão do número de medalhas, ou seja, a variação dos dados observados em torno do valor médio: ($S_t = 424,87$ medalhas; $S_{t-EUA} = 252,63$ medalhas; $S_{t-EUA-URSS} = 211,88$ medalhas);
- 3) Outra medida importante no estudo da variabilidade é o valor máximo observado, onde também pode-se observar que há um valor menor: ($P_{100\%(t)} = 2548$ medalhas; $P_{100\%(t-EUA)} = 1204$ medalhas; $P_{100\%(t-EUA-URSS)} = 802$ medalhas);
- 4) Há também diminuição do valor observado do Intervalo Interquartilício (IQR). E considerando que esta medida, que indica que a variabilidade dos dados, tendo a propriedade de ser não negativa e de ser tanto maior quanto maior for a variabilidade presente nos dados; pode-se dizer que o comportamento das distribuições em estudo indica que há uma menor variabilidade.

Desta forma, o valor das medidas indica que com a sucessiva retirada do número de medalhas dos Estados Unidos e da União Soviética, os dados se apresentam mais homogêneos, indicando que o número de medalhas destes dois países altera o comportamento da distribuição dos dados.

Não menos importante para a educação estatística é indicado neste estudo a intervenção da tecnologia computacional devido ao seu potencial para promover a parte qualitativa da própria análise estatística e contribuir para o desenvolvimento conceitual.

Em particular, refere-se a mobilizar recursos, como diferentes objetos estatísticos ou matemáticos por meio de representações numéricas (medidas de tendência central e de

dispersão), tabular (representação resumo dos dados), gráficas (fotografia do comportamento dos dados) e simbólicas (conceitos).

Que por sua vez pode gerar ambientes favoráveis para o surgimento de variadas experiências de aprendizagem através de uma abordagem concreta e visual, permitindo, intuitivamente, iniciar o contato com conceitos e a descoberta de relações complexas entre esses tipos de objetos.

E assim, no segundo encontro, com a participação de mais dois componentes, foi proposto que as atividades fossem realizadas em duplas, sendo que cada dupla recebeu um problema diferente. Os problemas foram elaborados pensando em cobrir as diferentes áreas e cursos em que a universidade oferece para a comunidade.

Os problemas foram resolvidos com papel e lápis e com a utilização do *software* R. Os participantes construíram a distribuição de frequências, o polígono de frequência e puderam interpretar a dispersão dos dados.

Como exemplo do que foi realizado neste segundo momento, a Figura 10 apresenta problema que foi proposto aos participantes, abordando o nível de glicose em alunos de uma escola pública.

Figura 10 - Atividade proposta no segundo momento da atividade.

“O diabetes se caracteriza por uma alteração na produção do hormônio insulina pelo pâncreas ou uma resistência à ação da insulina pelo organismo. É a insulina que ajuda o organismo a transformar o açúcar (ou glicose) em energia, promovendo assim um bom funcionamento do corpo humano. Na falta da insulina, a glicose se concentra no sangue e apenas uma pequena parte passa para o interior da célula, de uma maneira que esta célula fique como se estivesse subnutrida. Nesta situação, a pessoa pode se sentir cansada, indisposta, comer muito, ter muita fome, acordar a noite para urinar, urinar várias vezes durante o dia e ter muita sede. A dificuldade de cicatrização, o acometimento da visão, de circulação e a impotência sexual masculina são algumas das complicações quando o diabetes não é controlado. O diabetes é a segunda doença crônica mais comum na infância. “Antigamente era uma doença apenas dos adultos, mas com o crescimento do índice de obesidade infantil, associada a uma vida sedentária e aos maus hábitos alimentares, os casos de diabetes aumentaram consideravelmente entre meninos e meninas.”

Fonte: <http://br.guiainfantil.com/diabetes.html>

Assim, considere os seguintes dados referentes ao nível de glicose de 60 crianças de uma escola pública do município de Uberaba, Minas Gerais:

**56 61 57 77 62 75 63 55 64 60 60 57 61 57 67 62 69 67 68 59
72 65 61 68 73 65 62 75 80 66 61 69 76 72 57 75 68 83 64
69 64 66 74 65 76 65 58 65 64 65 60 65 80 66 80 68 55 66 71**

a) Construa uma distribuição de frequência utilizando o método tradicional.
b) Esboce o histograma de frequências e o polígono de frequências utilizando o método tradicional e interprete o seu comportamento.
c) Interprete a dispersão dos dados e também as medidas de assimetria e curtose e interprete o seu comportamento.

Fonte: Elaborado pelos autores.

No *software* R pode-se plotar medidas de tendência central e de dispersão ao utilizar as funções *summary* e *quantile* (Figura 9) e que dão suporte ao estudo do conceito de variabilidade.

A Tabela 2 apresenta as medidas de tendência central e de dispersão dos dados que foram gerados a partir do *software* R com as funções apresentadas na Figura 10.

Tabela 2 - Medidas de tendência central e de dispersão geradas pelo *software* “R” para o problema do nível de glicose.

n	39
Média	284,359
Desvio	197,985
Mínimo	62
1º Quartil	109
Percentil 10	70,5
Mediana	243
3º Quartil	451
IQR*	342
Percentil 90	604
Máximo	730

*Intervalo Interquartilico.

Fonte: Dados organizados pelos autores da saída do software R.

Após o desenvolvimento da sequência didática foi aplicado instrumento para avaliar a aceitação, ou não, das metodologias de ensino e solicitar aos participantes que apresentem outras maneiras de aplicação dos conceitos em situações do cotidiano. Portanto, foi perguntado qual foi o método preferido (utilizando papel milimetrado e lápis ou “utilizando a tecnologia”) no desenvolvimento da atividade e solicitado que fosse justificada a opção.

A metade dos participantes das atividades que responderam ao instrumento de pesquisa considera que a utilização da tecnologia, ou seja, utilizar o *software* “R” para a solução dos problemas propostos contribui de forma efetiva para o desenvolvimento da apreensão dos conteúdos abordados. Estes depoimentos são apresentados a seguir:

Fica visualmente mais compreensível. **(PIBID 1 – 19 anos)**

Eu preferi o método utilizando a tecnologia, pois, os resultados são mais exatos e o gráfico fica visualmente melhor de entender. **(PIBID 2 – 20 anos)**

Porque aumentou meu interesse pelo desenvolvimento da oficina, além disso, o uso da tecnologia desperta e motiva o aprendizado. **(PIBID 3 – 23 anos)**

Porque usando a tecnologia a resolução fica mais clara, e a visualização fica melhor. **(PIBID 4 – 19 anos)**

Utilizando o recurso do *software* R é uma aula diferente, em outro ambiente. Esta locomoção atrai a curiosidade do aluno, no meu caso, facilita a construção do gráfico a partir dos dados encontrados. Outro fator interessante foi referente ao final da atividade onde se calculava o exercício usando o “R” e então podíamos desenhar o gráfico na folha, utilizando os dados encontrados no *software*. **(PET 1 – 21 anos)**

Pela rapidez com que você consegue obter o gráfico e as informações necessárias para calcular os coeficientes, também pela exatidão que você consegue o gráfico. Acredito que também seria interessante calcular os coeficientes no *software* e mostrar a localização da média, moda, no gráfico. **(PET 2 – 21 anos)**

Cabe ainda considerar que a média de idade deste grupo é de 20,5 anos, com desvio padrão 1,5 anos, ou seja, de acordo com o coeficiente de variação, a dispersão das idades é considerada baixa (7,3%). Observa-se que o grupo, em média, é jovem.

Outro fator é que a composição deste grupo é de 4 participantes do PIBID Matemática e 2 (dois) do PET Matemática, ou seja, alunos em formação inicial. A distribuição do sexo é 50% para homens e mulheres.

O depoimento dos participantes vai ao encontro do que afirma Levy (2008), ou seja, que as tecnologias da informação e comunicação atualmente estabelecem uma nova forma de pensar sobre o mundo que vem substituindo princípios, valores, processos, produtos e instrumentos que mediam a ação do homem com o meio.

Além disso, corroboram o referencial teórico trazido neste texto por Ponte, Brocardo e Oliveira (2006), Ben-Zvi (2011) e Castro L., Gutiérrez F. e Hugues G. (2014), quando afirmam que a integração de ferramentas tecnológicas no ensino da variabilidade estatística desempenha um papel didático importante, contribuindo para o desenvolvimento do raciocínio e da compreensão dos estudantes.

Na sociedade atual, não é mais possível negar que as novas tecnologias estão presentes nas experiências diárias dos indivíduos e a escola não pode ficar à margem dessas vivências dos alunos.

Hoje, o aprendizado é diferente das gerações anteriores e a reforma no ensino trouxe mudanças nas disciplinas e nos conteúdos. Por outro lado, são muitas as mudanças e as transformações que ocorreram nas últimas décadas na sociedade brasileira e estas estão se refletindo na escola, onde os professores não podem mais ficar indiferentes a estas mudanças, redirecionamentos e reconstruções.

Observa-se ainda que a outra metade dos participantes declarou que utilizar a tecnologia, no caso, o *software* “R”, contribui de forma positiva para o desenvolvimento da apreensão dos conteúdos abordados, mas destacam também ser importante trabalhar os conceitos utilizando papel milimetrado e lápis. Estes depoimentos são apresentados a seguir:

O uso do *software* facilitou bastante, com ele calculando todos os valores e gerando o gráfico, mesmo se tiver que fazer o gráfico manualmente, o entendimento da questão já está mais claro. Mesmo para quem ainda não cursou a disciplina, o entendimento foi facilitado com o uso do *software*. **(PIBID 5 – 36 anos)**

A visualização é mais fácil, os dados são calculados mais rapidamente, mas a explicação dos conceitos pode ser feita utilizando papel milimetrado e lápis. Concluindo, é interessante trabalhar os dois métodos em conjunto. **(PET 3 – 25 anos)**

Acho que os dois métodos são importantes para a compreensão de todos os processos para calcular os dados, porém o uso da tecnologia facilita e muito o cálculo desses valores, pois, muitas vezes ao realizar o mesmo, podemos calcular erroneamente, comprometendo assim todo o resultado. **(PET 4 – 22 anos)**

Porque a tecnologia auxiliou na rapidez para a resolução dos exercícios e possibilitou concretizar os conceitos estatísticos que são muito abstratos. **(Mestrando em Educação 1 – 38 anos)**

É mais rápido e nos ensina a ter domínio do *software*, mas para saber interpretar os dados é preciso ter domínio utilizando papel milimetrado e lápis. **(Mestrando em Educação 2 – 30 anos) 30 ANOS**

O método utilizando a tecnologia é bastante eficiente, no entanto, acredito que a união dos métodos traria um resultado melhor. **(Mestrando em Educação 3 – 25 anos)**

Observamos que a média de idade deste grupo é de 29,3 anos, com desvio padrão 6,5 anos, ou seja, de acordo com o coeficiente de variação, a dispersão das idades é considerada moderada (22,2%).

Destacamos que a diferença média das idades entre os grupos é de $(29,3 - 20,5) = 8,8$ anos. Assim, percebe-se que grupo que considera a associação das duas metodologias de ensino é 8,8 anos mais velho. Percebe-se que é um grupo mais experiente e com vivências maiores no dia a dia e em sua formação.

Outro fator é que a composição deste grupo é de 3 alunos do Mestrado em Educação, 1 aluno do PIBID Matemática (com 36 anos) e 2 alunos do PET Matemática, ou seja, participantes tanto em formação inicial quanto em formação continuada. A maioria é do sexo feminino (66,67%).

Concordamos ainda com Tornaghi et al. (2010) ao afirmar que a tecnologia não é apenas a parceira que nos permite fazer as mesmas coisas de forma mais divertida ou eficiente. Ela traz, também, novos conhecimentos e novas necessidades para a escola; e, naturalmente, atreladas a tantas “novidades”, surgem, igualmente, novas possibilidades, como formas inusitadas de registro, de leitura e de trabalho em parceria.

Além disso, é necessário avançar para além do uso do papel e lápis para realizar cálculos e desenhar gráficos e tabelas, sendo desejável o uso das novas tecnologias em situações de ensino, incluindo os *softwares* específicos que é vista como ferramenta pedagógica frequentemente recomendada (BATANERO; GODINO; FLORES, 2001; COBO, 2003).

Assim, o depoimento destes alunos nos leva a pensar na existência de bom senso na relação do ensino dito tradicional (utilizando papel milimetrado e lápis) e o ensino com suporte tecnológico. Não podemos levá-la a extremos, onde, por exemplo, todos tenham computadores para utilizar em sala de aula, não sendo obrigatório o ensino da letra cursiva. É preciso considerar uma mescla entre um tradicional método de ensinar, ou seja, utilizando papel e lápis, e os avanços da tecnologia, como a utilização de um *software*.

Devemos também considerar como ideal um ensino usando diversos meios, ou seja, no qual todos os meios deveriam ter oportunidade, desde os mais modestos até os mais elaborados. Desde a utilização do quadro, dos mapas e das transparências de retroprojeter até as antenas de satélite de televisão. No processo ensino e aprendizagem deveriam também estar presentes todas as linguagens, ou seja, desde a palavra falada e escrita até as imagens e sons, passando pelas linguagens matemáticas, gestuais e simbólicas (SANCHO, 2001, p. 136).

Acreditamos que devemos aplicar um diferencial tecnológico em sala de aula, sem deixar de ser educativo, mantendo assim o aluno interessado e o professor com o controle, não só da sala de aula, mas também das ferramentas à disposição para levar ao conhecimento.

Além disso, a percepção da variabilidade começa com uma reflexão sobre o problema proposto e de que forma pode-se encontrar a sua solução, ou mesmo mais claramente, na observação do seu comportamento.

Um maior nível de especificidade pode ser dado para organizar os dados ou esboçar gráficos, para gradualmente, dar lugar a uma concepção mais formal e assim quantificar a variabilidade, descrevê-la e enfim explicá-la.

5 CONSIDERAÇÕES FINAIS

O grau de compreensão dos participantes sobre a variabilidade convergiu para a maneira pela qual perceberam a distribuição de dados; no entanto, o problema proposto também foi importante para a compreensão do comportamento dos dados.

Após a realização da pesquisa e análise de resultados, verificou-se que o uso do *software* foi aceito pelos participantes. Mas, não devemos considerar a possibilidade de descartar o uso do método tradicional (utilizando papel milimetrado e lápis) em sala de aula, uma vez que, é possível unir as duas metodologias para a apresentação dos conceitos.

Sabemos que a tecnologia nos traz diversas ferramentas e opções para diversificarmos a maneira de ensinar estatística, no caso desta pesquisa, os conceitos de dispersão de dados, assimetria e curtose de dados. Assim, a partir dos resultados obtidos neste trabalho, apresentamos o seguinte resumo:

- Todos os participantes afirmaram que gostaram de realizar as atividades;
- Todos os participantes afirmaram que após a realização das atividades, compreenderam os conceitos da dispersão de dados e os conceitos da assimetria de dados;
- Somente um dos participantes afirmou que não compreendeu o conceito de curtose de dados e justificou que não ficou claro o conceito;
- Somente um dos participantes afirmou que se os conceitos matemáticos fossem trabalhados de maneira divertida, seu interesse não aumentaria;
- Todos os participantes preferiram o método em que foi utilizado o *software* R na solução dos problemas.

Consideramos, portanto, que as tecnologias de informação e comunicação estão presentes, e o educador tem a oportunidade de levar para a sala de aula e proporcionar ao seu aluno um ensino atual e de qualidade, porém, é preciso antes de tudo saber o que está fazendo, o porquê da utilização, se tem fundamento e se irá servir para o conteúdo que será trabalhado, pois de nada adianta utilizar destes meios apenas porque está na moda ou porque é orientação da instituição.

Também não podemos apenas utilizar um meio de aprendizagem, devemos ter um equilíbrio e saber sempre quando utilizá-los, diversificando as formas de dar aula e de realizar atividades.

Este trabalho corrobora que o mundo contemporâneo, está marcado pelos avanços na comunicação, na informática e por outras tantas transformações tecnológicas e científicas. Essas transformações intervêm nas várias esferas da vida social, provocando mudanças econômicas, sociais, políticas, culturais, afetando, também escolas e o exercício profissional da docência.

Isto se reflete nos tipos de atividades propostas em sala de aula, onde a educação se depara com o duplo desafio: adaptar-se aos avanços das tecnologias e orientar o caminho de todos para o domínio e a apropriação crítica desses novos meios.

A atividade pode mostrar o auxílio da informática no ensino de Estatística. Por um lado, o seu apoio na redução dos cálculos para ensinar de uma maneira tradicional, e por outro lado, a visualização de conceitos. E a possibilidade de simulações, que se mostrou vital na geração de situações de ensino que se tornaria difícil realizá-los em sala de aula por outros métodos.

Os participantes concluíram que não devem realizar todas as atividades utilizando a tecnologia. Além disso, acreditam que deve ser utilizada de forma equilibrada com outras técnicas de ensino, como o papel milimetrado e o lápis.

6 RECONHECIMENTOS

Os autores agradecem aos alunos e alunas participantes. Agradecem também o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) na forma de bolsa de Iniciação Científica junto ao Programa de Educação Tutorial (PET) em curso de Licenciatura em Matemática.

REFERÊNCIAS

BATANERO, C.; GODINO, J. D.; CAÑIZARES, M. J. Simulation as a tool to train pre-service school teachers. In: ICMI African Regional Conference, 1., 2005. *Proceedings...* Johannesburg: ICMI, 2005.

BATANERO, C.; GODINO, J. D.; FLORES, P. El análisis didáctico del contenido matemático como recurso en la formación de profesores de Matemáticas. In: OLIVIER, A.; NEWSTEAD, K. (Eds.). *Proceedings of International Conference for the Psychology of Mathematics Education*, 2001, v. 22, Stellenbosch, África do Sul. Disponível em: <<http://www.ugr.es/local/batanero>>. Acesso em: 21 nov. 2017.

BECKER, B. J. A look at the literature (and other resources) on teaching Statistics. *Journal of Educational and Behavioral Statistics*, v. 21, p. 71-90, 1996.

BEN-ZVI, D. Statistical reasoning learning environment. *EM TEIA – Revista de Educação Matemática e Tecnológica Ibero-americana*, Recife, v. 2, n. 2, p. 1-13, 2011.

BOROVCNIK, M.; KAPADIA, R. Research and Developments in Probability. *Education International – Electronic Journal of Mathematics*, New York, v. 4, n. 3, p. 111-130, 2009.

CASTRO L., F. J.; GUTIÉRREZ F., G.; HUGUES G., E. *Un acercamiento a la variabilidad estadística usando tecnología computacional. Tecnología Computacional en la Enseñanza de las Matemáticas*. Universidad Autónoma de Nuevo León y Asociación Mexicana de Investigadores en Uso de la Tecnología en Educación Matemática A. C., Monterrey, México, 2014. Disponível em: <http://pmme.mat.uson.mx/Articulos/Hugues2014Un_acercamiento_variabilidad.pdf>. Acesso em: 08 abr. 2018.

COBO, B. *Significados de las medidas de posición central para los estudiantes de secundaria*. 2003. 303f. Tese (Doutorado em Educação) - Granada, Espanha, Departamento de Didáctica de la Matemática, Universidade de Granada, 2003. Disponível em: <<http://www.ugr.es/~batanero/proyecto.html>>. Acesso em: 15 nov. 2017.

GARFIELD, J.; BEN-ZVI, D. A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, v. 4, n. 1, p. 92-99, 2005.

_____. Developing students' statistical reasoning: Connecting research and teaching. New York: Springer, 2008.

GARFIELD, J.; ALIAGA, M.; COBB, G.; CUFF, C.; GOULD, R.; LOCK, R.; MOORE, T.; ROOSMAN, A.; STEPHENSON, B.; UTTS, J.; VELLEMAN, P.; WITMER, J. Guidelines for assessment and instruction in statistics education (GAISE) College report. Alexandria, VA: American Statistical Association, 2005. Disponível em: <<http://www.amstat.org/education/gaise/GAISECollege.pdf>>. Acesso em: 08 abr. 2018.

GOULD, R. Statistics and the Modern Student. *International Statistical Review*, v. 78, n. 2, p. 297–315, 2010.

KONOLD, C.; POLLATSEK, A. Conceptualizing an average as a stable feature of a noisy process. In: BEN-ZVI, D.; GARFIELD, J. (Eds.). *The challenge of developing statistical literacy reasoning and thinking*. Netherlands: Kluwer Academic Publishers, 2004. p. 169-199.

LEVY, P. *Cibercultura*. Trad. Carlos Irineu da Costa. São Paulo: Editora 34, 2008.

MALTBY, J. Learning statistics by computer software is cheating. *Journal of Computer Assisted Learning*, v. 17, n. 3, p. 329-330, 2001.

MOORE, D. Uncertainty. In: STEEN, L. A. (Ed.). *On the shoulders of giants: New approaches to numeracy*. Washington, DC: National Academy Press, 1990. p. 95-137.

PANNUTI, M. R. V. *Caminhos da prática pedagógica*. TVE Brasil. Rio de Janeiro, 2004. Disponível em: <<http://tvebrasil.com.br/SAUTO/boletins2004/ei/text1.htm>>. Acesso em: 08 abr. 2018.

PONTE, J. P.; BROCARD, J.; OLIVEIRA, H. *Investigações Matemáticas em sala de aula*. Belo Horizonte: Autêntica, 2006.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2011. Disponível em: <<https://www.r-project.org/>>. Acesso em: 28 mai. 2017.

SANCHO, J. M. (Org.) *Para uma tecnologia educacional*. 2. ed. Porto Alegre: Artmed, 2001.

SCHUYTEN, G.; THAS, O. Statistical thinking in computer-based learning environments. *International Statistical Review*, v. 75, n. 3, p. 365-371, 2007.

SNELL L.; PETERSON W. *Does the computer help us understanding statistics? In Statistics for the Twenty First Century*. Ed. Florence, G. and Gordon, S. MAA Notes, n. 26, 1992.

TORNAGHI, A. J. da C.; COSTA, A. J da; BRITO, M. E. B.; ALMEIDA, M. E. B. de A. *Tecnologias na educação: ensinando e aprendendo com as TIC: guia do cursista*. 2. ed. Brasília: Secretaria de Educação à Distância, 2010.

VERZANI, J. Using R in Introductory Statistics Courses with the pmg Graphical User Interface. *Journal of Statistics Education*, v. 16, n. 1, p. 1-17, 2008.

WATSON, J. M.; CALLINGHAM, R. A.; KELLY, B. A. Students' appreciation of expectation and variation as a foundation for statistical understanding. *Mathematical Thinking and Learning*, v. 9, n. 2, p. 83-130, 2007.

WATSON, J.; KELLY, B.; CALLINGHAM, R.; SHAUGHNESSY, M. The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology*, v. 34, n. 1, p. 1-29, 2003.

WILD, D. J.; PFANNKUCH, M. Statistical thinking in empirical enquiry. *International Statistical Review*, v. 67, p. 223-265, 1999.