

O Uso de Simulação para o Ensino de Estatística Inferencial: o caso do Teorema Central do Limite

Fernando Frei¹ 

UNESP, FCLAssis, Departamento de Ciências Biológicas, Assis, SP, Brasil

Resumo

Os conceitos de Estatística Inferencial compõem a base pela qual a prática da Estatística é desenvolvida em grande parte dos cursos de graduação. O Teorema Central do Limite é o pilar central para a sustentação de análises paramétricas e, portanto, de parte expressiva da Estatística Inferencial. O objetivo deste artigo é apresentar um módulo de simulação usando planilha eletrônica para o aprendizado interativo que ilustra os conceitos e a validade do Teorema Central do Limite para alunos do curso de Graduação em Ciências Biológicas como uma atividade formadora. Os resultados apontam que a simulação tem a capacidade de promover o aprendizado ativo com o uso de tecnologia para desenvolver o entendimento conceitual do teorema. Ademais, a atividade de simulação auxiliou o docente a entender dificuldades apresentadas durante o processo o que possibilita propor novas estratégias para o ensino do referido teorema.

Palavras-chave: Simulação; Teorema Central do Limite; Estatística; Planilha Eletrônica.

The Use of Simulation for Inferential Statistics Teaching: The Case of the Central Limit Theorem

Abstract

The concepts of Inferential Statistics form the basis upon which the practice of Statistics is developed in most undergraduate courses. The Central Limit Theorem is the core pillar for the sustenance of parametric analyses and, therefore, it is a significant part of Inferential Statistics. The aim of this paper is to present a simulation module using electronic spreadsheets for interactive learning that illustrates the concepts and validity of the Central Limit Theorem for undergraduate Biology students as a training activity. The results show that simulation is able to promote active learning through the use of technology to develop the conceptual understanding of the theorem. Besides, the simulation activity helped the teacher understand difficulties presented during the process, which makes it possible to propose new strategies for teaching the aforementioned theorem.

Keywords: Simulation; Central Limit Theorem; Statistic; Spreadsheet.

Submetido em: 20/11/2020

Aceito em: 27/01/2021

Publicado em: 02/02/2021

¹ Doutor em Saúde Pública pela FSP-USP. Professor do Programa de Pós-Graduação em Biociências da FCLAssis - UNESP. Av. Dom Antônio, 2100, CEP 19806.900, Assis - SP. E-mail: fernando.frei@unesp.br.

El uso de la simulación para la enseñanza de la estadística inferencial: el caso del teorema del límite central

Resumen

Los conceptos de Estadísticas Inferenciales constituyen la base por la cual la práctica de Estadística se desarrolla en la mayoría de los cursos de grado. El Teorema de Límite Central es el pilar central para el apoyo a los análisis paramétricos y, por lo tanto, a una parte significativa de las estadísticas inferenciales. El objetivo de este artículo es presentar un módulo de simulación utilizando una hoja de cálculo para el aprendizaje interactivo que ilustre los conceptos y la validez del Teorema de Límite Central para estudiantes de ciencias biológicas como una actividad formativa. Los resultados indican que la simulación tiene la capacidad de promover el aprendizaje activo con el uso de la tecnología para desarrollar la comprensión conceptual del teorema. Además, la actividad de simulación ayudó al profesor a entender las dificultades presentadas durante el proceso, lo que permite proponer nuevas estrategias para la enseñanza de este teorema.

Palabras clave: Simulación; Teorema del límite central; Estadística; Hoja de cálculo electrónica.

1. Introdução

O campo da tecnologia digital educacional aplicada ao ensino da Estatística é recente uma vez que os primeiros computadores pessoais estavam disponíveis nas universidades e nas residências nas décadas de 1980 e 1990 (BEN-ZVI, 2000). As aulas de estatística de hoje podem ser ministradas em laboratório com alunos trabalhando em computadores, em computadores portáteis na escola, em suas casas e até em movimento com seus *smartphones* (CHANCE et al., 2007).

Como enfatiza Tishkovskaya e Lancaster (2012), diferente dos métodos em que a prioridade eram os cálculos tediosos e a memorização de fórmulas, o ensino da Estatística na atualidade apresenta como metas a alfabetização estatística, ou seja, a necessidade de adquirir capacidades para avaliar e interpretar o conjunto crescente de informações a que todos são submetidos diariamente, a competência para raciocinar sobre problemas do mundo real e as habilidades para aprendizagem ativa.

Métodos de aprendizagem ativa fazem parte da abordagem construtivista que podem incluir aprendizagem baseada em problemas, aprendizagem colaborativa e aplicações de tecnologia em simulações, entre outras. Neste sentido, a tecnologia também ampliou as possibilidades de visualização para que os estudantes possam explorar e analisar dados e conceitos estatísticos, permitindo-lhes concentrar-se na interpretação dos resultados e na compreensão de conceitos e não em mecânica computacional (TISHKOVSKAYA; LANCASTER, 2012).

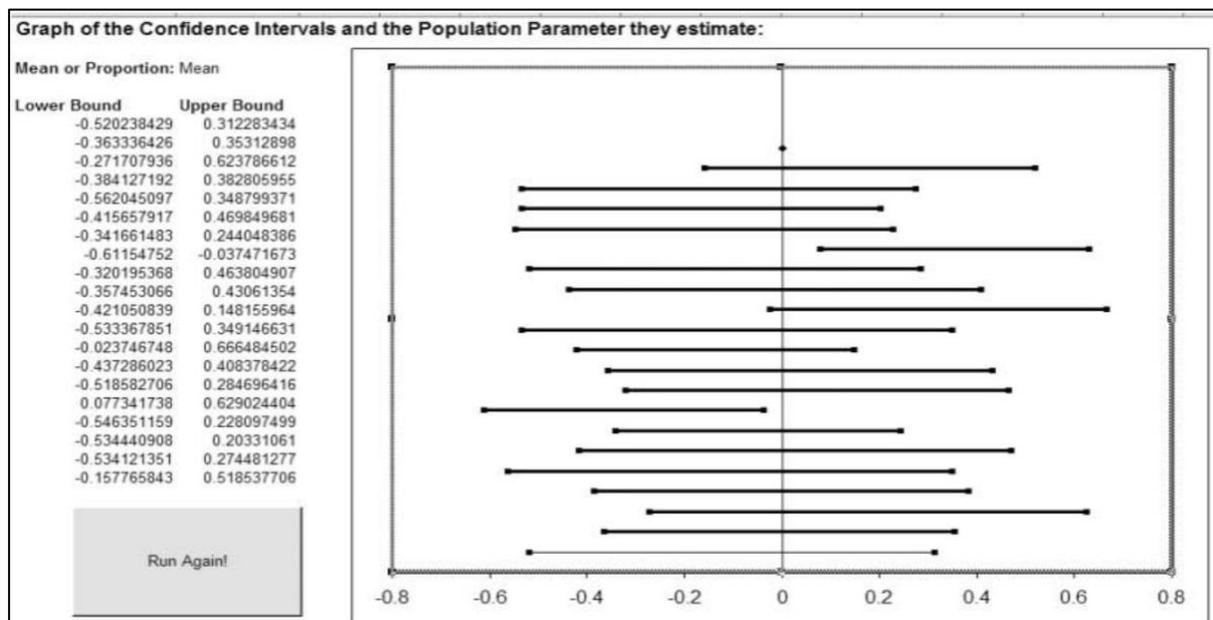
Por outro lado, mesmo com a utilização de programas computacionais que permitem que os estudantes executem tarefas de rotina de análise de dados com o objetivo de melhorar a aprendizagem, conceitos abstratos ainda podem ser difíceis de serem compreendidos (MILLS, 2002).

Para suplementar as atribuições de análise de dados fornecendo aos estudantes experiências adicionais, outras estratégias metodológicas podem contribuir para a melhoria do ensino e aprendizagem da Estatística nas mais diversas áreas e níveis de ensino. Entre elas, pode-se destacar o uso de simulações. Como destacam Barr e Scott (2011), ao utilizar simulações, pode-se relevar uma gama de ideias mal interpretadas que são fundamentais para o conhecimento estatístico básico, como os de aleatoriedade e distribuição de probabilidade. Assim, a simulação é baseada na utilização de técnicas as quais possibilitam reproduzir o funcionamento de, praticamente qualquer operação ou processo do mundo real.

Os conceitos de probabilidade e distribuições de probabilidade, entre as quais a distribuição Binomial e distribuição Normal, são muitas vezes apresentados por meio de simulações. Ferreira; Kataoka e Karrer (2014), utilizaram a linguagem R como ferramenta para simular dados e introduzir conceitos de probabilidade, enquanto Delmas; Garfield e Chance (1999) utilizaram um programa denominado de *Sampling Distributions* para simular populações com distribuição normal.

Autores como Kennedy et al. (1990) apud Mills (2002), utilizaram a simulação para o ensino de Intervalo de Confiança para a média populacional. Para esta situação os autores sugerem fornecer aos estudantes uma população finita para que eles possam sortear uma amostra e calcular seu respectivo Intervalo de Confiança de 95% para a média. Essa estratégia permite comparar os diversos intervalos de maneira a verificar quantos contêm a verdadeira média populacional. De forma semelhante, Hagtvedt; Jones e Jones (2007) utilizam a simulação para discutir conceitos relativos ao Intervalo de Confiança para a média populacional de forma que cada aluno pudesse calcular simultaneamente n intervalos de confiança e observar graficamente seus resultados (Figura 1).

Figura 1. Tela dos resultados da simulação de n Intervalos de Confiança para média populacional



Fonte: Hagtvedt, R.; Jones, G. T.; Jones, 2008.

Neumann; Neumann e Hood (2011) e Chandrakantha (2014) apresentam um conjunto de diferentes simulações para o ensino de conceitos de tipos de gráficos para dados qualitativos, tabelas de distribuição de frequência, correlação e regressão, teste de hipóteses, análise de variância entre outros conceitos estatísticos.

Como exposto, o uso da simulação para aprendizagem de conceitos e técnicas estatísticas é grande e variada e se mostra como uma importante ferramenta didática. Neste sentido, o uso de planilhas eletrônicas desempenha um importante papel.

2. Planilhas eletrônicas como ferramentas para ensino de estatística

As planilhas de cálculos para computadores pessoais surgiram em meados no ano de 1979, um aplicativo projetado para ajudar nas tarefas de contabilidade. Desde então, a diversidade de aplicações é evidenciada por seu reaparecimento contínuo em diversas publicações. Em nenhum lugar sua aplicação se torna mais marcante do que no campo da educação. Desde o ensino fundamental até o ensino superior, a planilha eletrônica tem aumentado gradualmente sua importância como uma ferramenta para o ensino e aprendizagem (BAKER; SUGDEN, 2003; PACE; BARCHARD, 2006).

Por outro lado, especificamente no caso do ensino da Estatística, diversos softwares estatísticos dedicados têm sido utilizados como ferramenta para a aprendizagem, no entanto, como enfatizam Pace e Barchard (2006), esses programas podem não ser a melhor estratégia tecnológica pois os alunos podem realizar análises estatísticas simples e complexas sem ver as fórmulas ou entender os conceitos envolvidos.

Em países desenvolvidos o uso de Tecnologias Digitais da Informação e Comunicação - TDIC são integrados de forma mais rápida no ensino e aprendizagem, no entanto o mesmo não ocorre em países em desenvolvimento em que o custo financeiro dessas tecnologias ainda representa barreiras. No caso do ensino de estatística, alternativas com custos acessíveis devem ser priorizados. Uma forma pela qual é possível implementar tecnologias é pelo uso de planilhas eletrônicas (CHAAMWE; SHUMBA, 2016). As planilhas eletrônicas são fáceis de usar, permitem que os alunos vislumbrem os cálculos para que possam entender como as fórmulas funcionam. No entanto, uma vantagem significativa das planilhas eletrônicas é que os alunos podem ver simultaneamente os dados brutos, as fórmulas e os resultados estatísticos (PACE; BARCHARD, 2006).

Nash e Quon (1996) apud Chaamwe e Shumba (2016), Warner e Meehan (2001), Hunt, 2005 e Hunt, 2007, apresentam as principais vantagens do uso de planilhas eletrônicas no ensino:

- 1) O uso generalizado e conhecimento de planilha pode economizar custos de aquisição de softwares estatísticos dedicados;
- 2) Os professores podem preparar antecipadamente modelos para que os estudantes possam seguir e realizar cálculos particulares;
- 3) Os cálculos da planilha eletrônica oferecem uma atualização imediata dos resultados quando os dados são alterados;
- 4) As atuais planilhas eletrônicas oferecem um rol de análises estatísticas pré-definidas;
- 5) As planilhas eletrônicas são uma ferramenta útil e com facilidades para entrada, edição e manipulação de dados, antes mesmo do uso de pacotes estatísticos dedicados;
- 6) Facilidade de comunicação entre professor e aluno. A planilha eletrônica propicia que o professor possa inserir textos que indiquem diretamente o erro e oferece propostas para a correção na própria planilha. Essa operação didática é rara em softwares estatísticos;
- 7) Com planilhas eletrônicas é possível desenvolver atividades práticas individuais de forma a evitar ou minimizar plágio por parte dos alunos.

As características supracitadas das planilhas eletrônicas favorecem a proposição de atividades mais ativas por meio de simulações com o objetivo de tornar o ensino e aprendizagem de estatística mais envolvente e efetiva (BARR; SCOTT, 2011; CHANDRAKANTHA, 2014; HANEY, 2015).

3. Teorema Central do Limite

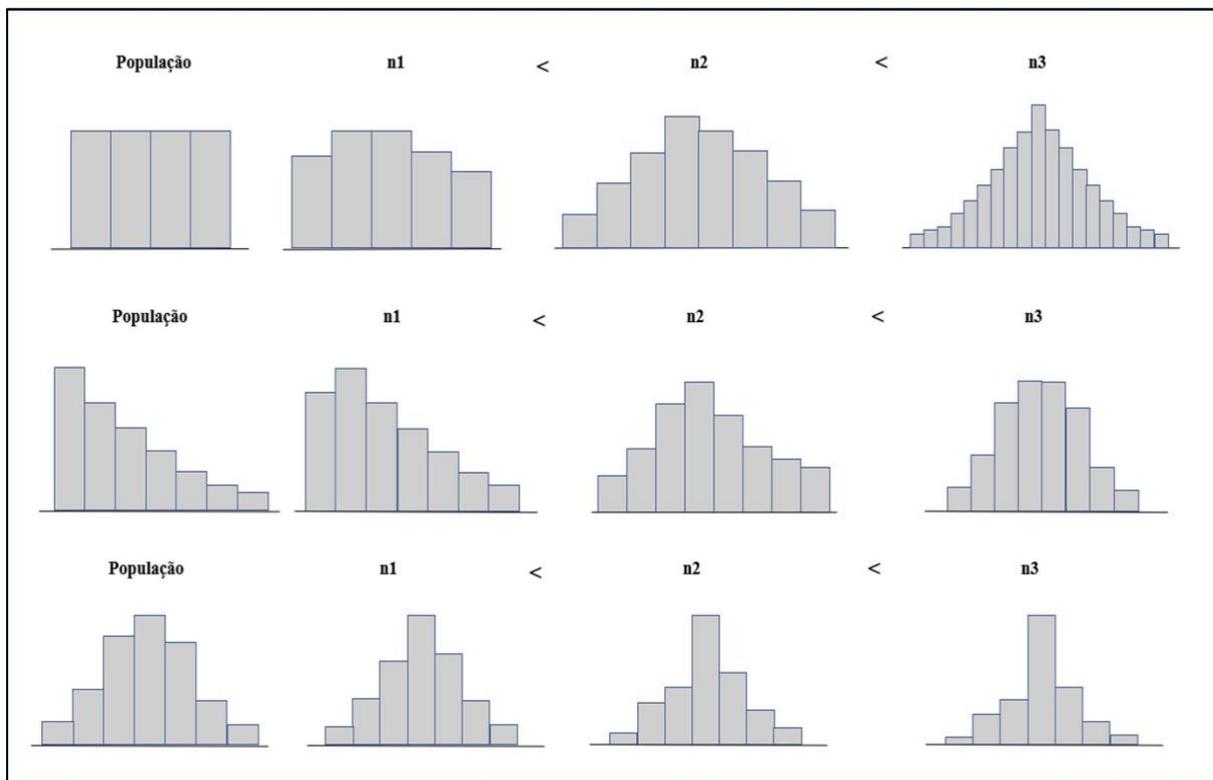
Os conceitos de Estatística Inferencial compõem a base pela qual a prática da Estatística é desenvolvida, assim, grande parte dos cursos introdutórios de Estatística tem como objetivo o ensino de fundamentos e aplicações de inferência (MAURER; LOCK, 2016). No entanto, a Estatística Inferencial é um tema complexo para o ensino e aprendizado. Pesquisadores de diversas áreas usam de forma inadequada as ferramentas inferenciais em muitas oportunidades (ERICKSON, 2006).

A Estatística Inferencial permite apresentar conclusões sobre as populações baseadas em amostras. De modo geral a mensuração de variáveis de interesse em toda a população é impraticável. A Estatística Inferencial permite apresentar conclusões sobre as populações baseadas em amostras. De modo geral a mensuração de variáveis de interesse em toda a população é impraticável. Não é possível medir o diâmetro do caule de todas as árvores da espécie chamada de Tipuana (*Tipuana tipu*), o peso de todos os peixes da espécie Tilápia (*Oreochromis spp.*), ou mesmo a taxa de desempregados na população de brasileiros. Desta forma, quando utilizamos uma ou mais amostras, não podemos ter 100% de certeza que nossos resultados podem ser extrapolados para a população da qual as amostras foram extraídas. Essa incerteza decorre das variações naturais dos objetos mensurados e das variações naturais entre as diversas amostras possíveis, o chamado erro amostral (ISLAM, 2018).

Como ressalta Taylor (2018), da mesma maneira que os cursos introdutórios de Estatística abordam com especial atenção à inferência, o Teorema Central do Limite, longe de ser um conceito matemático abstrato, é o alicerce sobre o qual se situam as estatísticas inferenciais paramétricas, as quais apresentam como pressuposto básico a existência da Distribuição de Gauss – Distribuição Normal na distribuição das médias amostrais. Essa suposição pode trazer restrições ao uso da estatística paramétrica, no entanto, o Teorema Central do Limite apresenta como resultado que para amostras aleatórias e independentes de tamanho n , a distribuição das médias amostrais se aproxima da Distribuição Normal à medida que n aumenta, independentemente da forma da distribuição da população em foco (Figura 2). O Teorema Central do Limite apresenta uma longa história, para o caso específico da distribuição de probabilidade de Bernoulli, uma versão do teorema foi obtida por De Moivre na

década de 1739. A primeira tentativa de um teorema mais geral foi feita por Laplace em 1810, mas os resultados mais definitivos foram obtidos no segundo quarto do século XX (Trosset, 2011, p. 125).

Figura 2. Representação das distribuições das médias amostrais.



Fonte: autor.

Ainda pelo Teorema Central do Limite, a distribuição das médias destas amostras tende para uma distribuição normal com média $\bar{x} = \mu$ (média das médias igual à média da população) e com desvio padrão $\frac{\sigma}{\sqrt{n}}$. Uma regra comum, é que a aproximação para a Distribuição Normal pode ser usada se $n > 30$. Frequentemente, a aproximação para a Distribuição Normal funciona com n ainda menor (TROSSET, 2011, p. 122).

Assim, a suposição de normalidade pode ser satisfeita sempre que tomarmos amostras grandes para a realização de inferências estatísticas. No entanto, em contraste com a o Teorema Central do Limite, Chance; Delmas e Garfield (2004) observaram que muitos estudantes acreditavam que as distribuições amostrais deveriam se parecer mais com a população à medida que o tamanho da amostra aumenta.

Como enfatiza Ruggieri (2016), livros textos utilizam simulações e outros autores apresentam sugestões do uso de programas computacionais para o ensino do Teorema Central do Limite (MAURER; LOCK, 2016; DINOVI; CHRISTOU e SANCHEZ, 2008; HAGTVEDT; JONES e JONES, 2007).

O principal objetivo deste estudo foi aferir os efeitos do processo de simulação com o uso de planilha eletrônica interativa para auxiliar no ensino do Teorema Central do Limite.

4. Método

Nesta pesquisa foi utilizado o método fundamentado no aprendizado da descoberta. Esse tipo de aprendizagem é baseado na premissa de que o conhecimento que os alunos encontram para si mesmos tem mais valor do que o conhecimento que foi simplesmente explicado (KLAHR; NIGAM, 2004). O objetivo do aprendizado da descoberta é envolver os estudantes em atividades que os levam a “descobrir” um princípio. No entanto, os alunos nem sempre entendem o princípio que o professor pretendia que eles descobrissem pois os alunos têm dificuldades em mudar suas crenças pré-estabelecidas, mesmo se os dados contradizem essas crenças (DE JONG; VAN JOOLIGEN, 1998; MILLS, 2002). Por essa razão, tem sido sugerido que tanto o aprendizado orientado quanto a descoberta, ou uma combinação de ambos são métodos de ensino eficazes (DELMAS; GARFIELD E CHANCE, 1999; LANE; PERES, 2006).

Com o aprendizado de descoberta guiada, os alunos recebem um conjunto de atividades que os levam a um objetivo predeterminado (LANE; PERES, 2006). É importante destacar que a atividade de simulação proposta não é necessariamente avaliativa, mas prioritariamente formativa, no sentido que apresenta conceitos subjacentes ao Teorema Central do Limite. As análises dos resultados foram desenvolvidas por meio de análise descritiva – gráficos e tabelas.

4.1 Sujeitos

A metodologia de simulação com planilha eletrônica foi aplicada em uma disciplina de estatística inferencial para 40 estudantes de graduação de um curso de Ciências Biológicas da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, São Paulo, Brasil. A carga horária para o curso foi de 60 horas durante um semestre.

4.2 Atividades

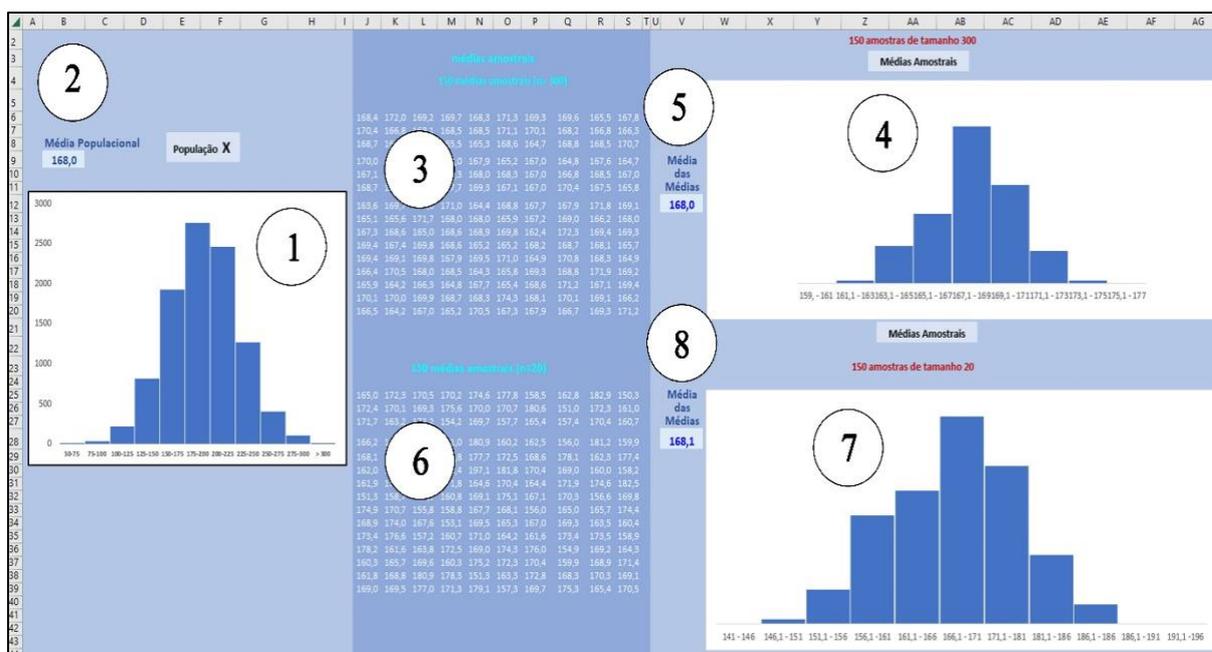
A atividade tem como objetivo realizar simulações para apresentar aos alunos as principais características do Teorema Central do Limite -TCL. Não foi realizada qualquer aula

ou disponibilizado qualquer material didático com os conceitos do TCL antes da atividade de simulação objeto deste artigo. Assim, os estudantes não tiveram qualquer contato com os conceitos do teorema.

Foi disponibilizado como parte da atividade, um vídeo de 3 minutos com instruções dos procedimentos a serem efetuados na planilha eletrônica Microsoft Excel®. A planilha eletrônica pode ser executada gratuitamente na plataforma Excel Online. Nesta planilha eletrônica diversas simulações para três populações distintas podem ser observadas. Para efeito de uma melhor visualização, a planilha é apresentada neste trabalho em figuras diferentes.

A primeira simulação é referente a população com distribuição Normal, suas respectivas amostras e médias amostrais (Figura 3) – A simbologia circular com números na figura não faz parte da planilha eletrônica e serve como legenda explicativa para o artigo em pauta.

Figura 3. Simulação – População com distribuição Normal.



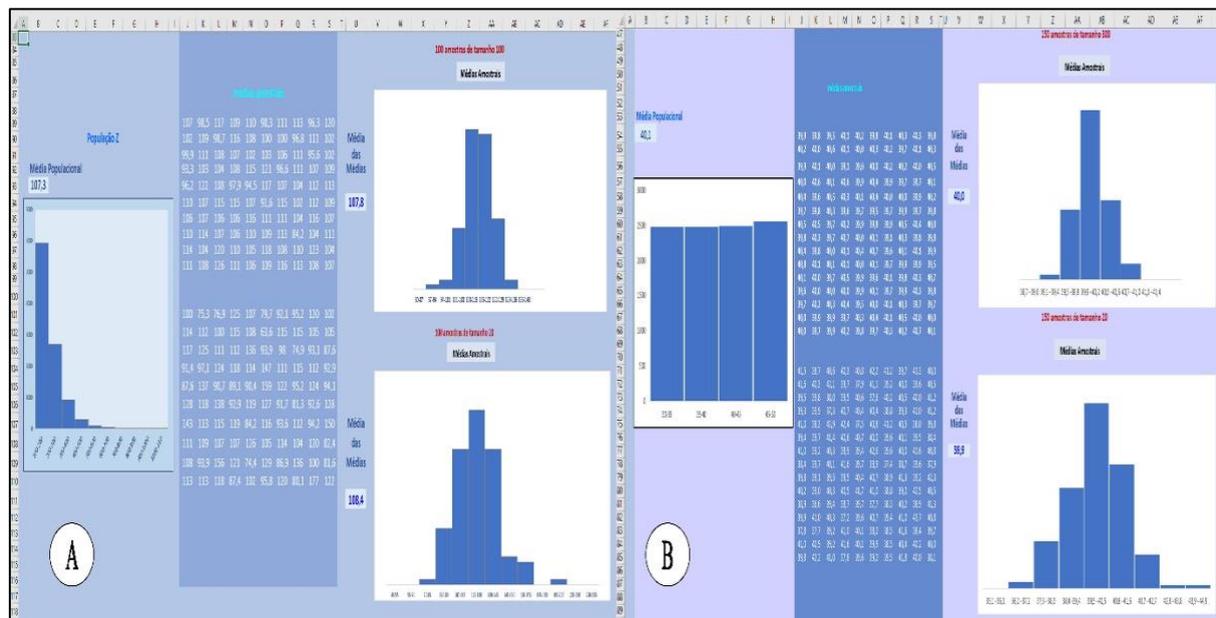
Fonte: autor.

Legenda:

- 1 – Histograma de uma população com variável simulada distribuída pela Normal
- 2 – Média Populacional
- 3 – Médias amostrais de 150 amostras de tamanho n = 300
- 4 – Histograma da distribuição das médias das 150 amostras de tamanho n= 300
- 5 – Médias das 150 médias amostrais de tamanho n = 300
- 6 – Médias amostrais de 150 amostras de tamanho n = 20
- 7 – Histograma da distribuição das médias das 150 amostras de tamanho n= 20
- 8 – Médias das 150 médias amostrais de tamanho n = 20

A Figura 4 apresenta duas outras simulações para duas distintas populações – população com distribuição Assimétrica (A) e Uniforme (B). Ambas apresentam as mesmas características operacionais que a distribuição apresentada na Figura 3.

Figura 4. Simulação – População com distribuição Assimétrica (A) e Uniforme (B).



Fonte: Autor.

4.3 Procedimentos

Para o desenvolvimento da atividade proposta, após a visualização do vídeo, os estudantes receberam a planilha eletrônica e acionaram a tecla F9 para que a planilha executasse todos os cálculos e a confecção dos histogramas para cada uma das três distribuições de forma automática. Esse processo de simulação foi realizado diversas vezes (de dez a quinze vezes) pelos estudantes de forma individual para que os conceitos do Teorema Central do Limite pudessem ser visualizados. Assim, ao acionar a tecla F9, 150 amostras de tamanho $n = 300$ e 150 amostras de tamanho $n = 20$ são selecionadas aleatoriamente da população, suas médias são calculadas e dispostas em duas outras colunas, um histograma para cada distribuição de médias é confeccionado e calcula-se a média das médias, as quais são também dispostas na planilha (figura 3). Os procedimentos foram realizados durante o tempo de duas horas em laboratório de informática.

4.4 Instrumento de coleta de dados

Concomitantemente ao processo de simulação, os alunos responderam individualmente a um questionário (Anexo I) para que o professor pudesse aferir os efeitos do processo de simulação.

Como mencionado anteriormente, a atividade se reveste de características formativas e por essa razão as respostas dos alunos não foram classificadas como corretas ou incorretas, mas buscou-se verificar se esses mesmos alunos conseguiram através da simulação obter bases intuitivas para facilitar a aprendizagem dos principais conceitos do Teorema Central do Limite. Assim, as respostas no questionário foram classificadas como:

Apropriada – No sentido de apropriar-se dos conceitos subjacentes ao teorema.

Parcialmente Apropriada.

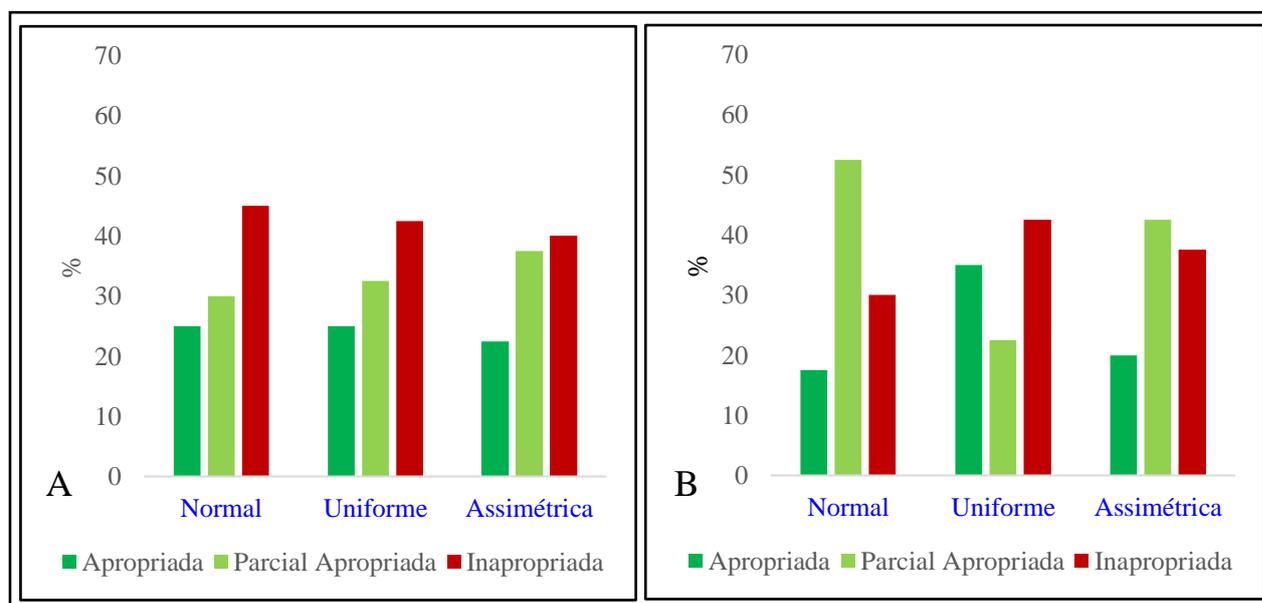
Inapropriada.

5. Resultados e Discussão

Os resultados são baseados nas questões e buscam evidenciar como o processo de simulação pode contribuir para a aprendizagem dos conceitos do TCL, mas também propiciar *feedbacks* ao professor para que o mesmo possa corrigir, minimizar e sanar equívocos no processo de ensino e aprendizagem.

No processo de simulação desenvolvido pelos estudantes foi possível realizar sorteios, cálculos e representações gráficas das três populações. Para o primeiro conjunto de simulações – pertinente às populações com Distribuição Normal, Distribuição Uniforme e Distribuição Assimétrica, buscou-se evidenciar que a distribuição das médias, independentemente das distribuições populacionais, apresenta para amostras de tamanho $n = 300$, distribuições Normais representadas pelos histogramas. Essa propriedade característica foi observada de forma apropriada e parcialmente apropriada por pouco mais da metade dos alunos como mostra a Figura 5 A. Resultado semelhante foi obtido quando a simulação com as distribuições para tamanhos amostrais menores ($n = 20$) foi avaliada (Figura 5 B). Nesta última simulação, muitos dos histogramas relativos a distribuições das médias apresentam forma pouco características da distribuição Normal, o que foi percebido pelos estudantes.

Figura 5. Resultados para tamanhos amostrais $n = 300$ (A) e $n = 20$ (B) quanto à forma das distribuições das médias.



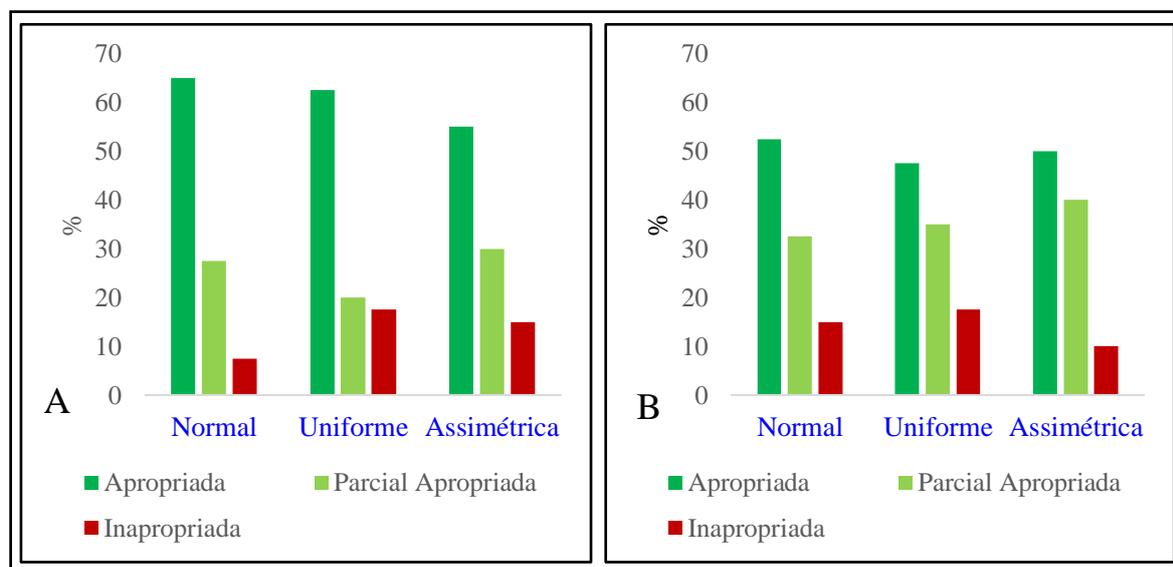
Fonte: Autor.

Os resultados das categorias apropriada e parcialmente apropriada, para as três distribuições no caso de amostras de tamanho 300 apresentam percentagem entre 50% e 60% o que pode ser considerada satisfatória dado que é a primeira situação a que os estudantes são expostos em um curso inferencial. Autores como Schwartz et al. (1997) apud Delmas; Garfield e Chance, (1999) apresentam mudanças modestas de 15% em alunos que experimentaram simulação no ensino de distribuições amostrais e Delmas; Garfield e Chance (1999) literalmente reportam despondimento ao verificar que muitos dos alunos apresentam mal-entendidos após o uso de programa computacional de simulação sobre Distribuições de Amostragem. Ainda com relação a apropriação dos conceitos relativos ao TCL, um grupo maior de alunos observou características adequadas da forma das distribuições das médias amostrais para amostras de tamanho 20 (Figura 5 B). Deve estar claro que a maior parte desta percentagem é resultado da categoria parcialmente apropriada.

Pesquisa análoga desenvolvida por Ruggieri (2016) para o ensino do Teorema Central do Limite com o uso de simulações mostra que de 21 alunos 9 (42%) responderam corretamente a questões sobre conceitos do teorema.

No caso da avaliação das médias das médias e sua proximidade com a média da população, os estudantes apresentam respostas adequadas ou parcialmente adequadas nas duas situações amostrais, amostras de tamanho 300 e tamanho 20 (Figuras 6 A e B).

Figura 6. Resultados para tamanhos amostrais $n = 300$ (A) e $n = 20$ (B) quanto proximidade das médias das médias em relação à média população.



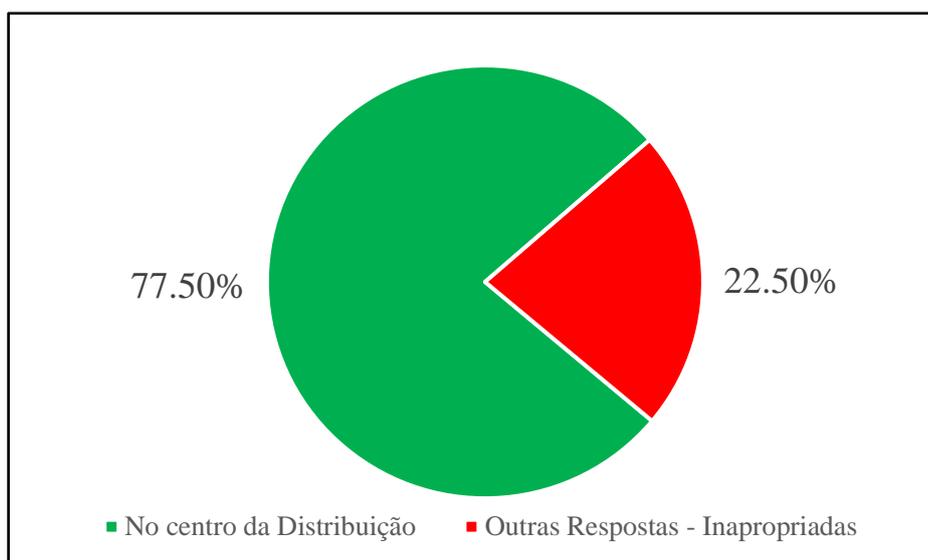
Fonte: Autor.

Na atividade seguinte da simulação os alunos avaliaram a variabilidade das distribuições das médias para amostras de tamanho $n = 300$ e $n = 20$. Neste caso, a variabilidade é incrementada de amostras maiores para amostras menores e essa situação foi também observada pela grande maioria dos estudantes – 75%, 67,5% e 70% para as distribuições Normal, Uniforme e Assimétrica respectivamente.

Por fim, a simulação propiciou identificar pela maioria dos alunos a provável localização de uma média amostral na distribuição das médias o que permite estabelecer quão próxima ou distante está sua média amostral em relação ao parâmetro de interesse (Figura 7). Resultado semelhante foi obtido por Brussolo (2018) em pesquisa envolvendo o ensino do Teorema Central do Limite que simula amostras de tamanhos que variam de 5, 20, 30, 50 e 100. Chandrakantha (2014), usa planilha eletrônica para simular resultados da distribuição da média amostral para diferentes tamanhos amostrais de distribuições de probabilidade Uniforme, Normal e Qui-Quadrado e avalia o procedimento como positivo entre os alunos para desenvolver os conceitos do Teorema Central do Limite. Ao usar o Excel como instrumento de simulação, Patron; Smith e Bold (2009) demonstram em seu trabalho que os alunos foram capazes de aprender os conceitos envolvendo o TCL, enquanto praticavam habilidades básicas de planilha, mas também recursos mais avançados da planilha eletrônica. Hagtvedt; Jones e Jones (2007) chamam a atenção que muitos alunos, especialmente aqueles com uma tendência para aprendizagem visual, foram beneficiados pelo software de simulação para o aprendizado

sobre os conceitos de distribuição de amostragem e Teorema Central do Limite. Essa atividade pode permitir mais facilmente o ensino de como o nível de confiança desempenha um papel de importância na construção de intervalos de confiança e testes de hipóteses em aulas e atividades subsequentes.

Figura 7. Respostas – Onde encontra-se sua média amostral.



Fonte: Autor.

O conjunto de respostas dos estudantes podem indicar outras questões de importância para o ensino e aprendizagem da Estatística. Ao avaliar a percentagem das categorias inapropriadas, verifica-se que a maior dificuldade é relativa à identificação das formas das distribuições de médias (Figura 5). Essa situação pode ilustrar obstáculos na aprendizagem em cursos de estatística inferencial em que os conceitos do pensamento aleatório são fundamentais, mas ao mesmo tempo, pode ser para a maioria desses mesmos estudantes, um raciocínio incomum dado que grande parte de seu treinamento foi dado pelo raciocínio determinístico da matemática. Pfannkuch e Brown (1996) observam que estudantes apresentam dificuldades para interpretar situações envolvendo probabilidade, o que indica a tendência para o pensamento excessivamente determinista, revelando dificuldades do conceito de variação. Essa condição pode decorrer da influência do ensino da Matemática em detrimento ao ensino da Estatística nos vários níveis de ensino precedentes ao terceiro grau, e por consequência a influência por um modo de pensar determinista (SHI; HE e TAO, 2009). Desta forma, a análise dos resultados das respostas dos estudantes para a forma das distribuições das médias amostrais sugere novas estratégias de ensino em aulas e atividades anteriores.

6. Considerações finais

A proposta de simulação com o uso de planilha eletrônica para o ensino do Teorema Central do Limite apresenta algumas limitações. Entre elas o desenho experimental que não contempla grupo controle. Também é importante salientar que a abordagem foi feita em uma única disciplina de um curso específico de Ciências Biológicas. Desta maneira, possíveis vantagens devem ser analisadas sob uma perspectiva específica.

A despeito das limitações descritas, as simulações podem desempenhar um caráter significativo no ensino e aprendizagem de processos estatísticos inferenciais. As ferramentas tecnológicas devem ser usadas para auxiliar os alunos a visualizar e explorar dados, não apenas para seguir algoritmos com fins pré-determinados.

A simulação com o uso de planilha eletrônica desempenhou um importante papel no ensino do TCL na medida em que cada aluno teve a oportunidade de estudar seu próprio e único conjunto de distribuições amostrais, promoveu o aprendizado ativo na sala de aula, utilizou a tecnologia para desenvolver o entendimento conceitual e análise de dados. A simulação também auxiliou o docente a entender dificuldades apresentadas durante a atividade proposta o que possibilita propor novas estratégias para o ensino do TCL.

Por meio da simulação, os estudantes podem confiar nos conceitos aprendidos dado que construíram o saber pela experiência, diferente da informação apresentada de forma pronta e passiva. A planilha eletrônica oferece aos alunos controle sobre o processo de simulação o que propicia o acesso a qualquer momento no ritmo desejado.

Simulações utilizando ferramentas tecnológicas podem estimular o interesse dos alunos e cultivar suas habilidades experimentais de forma a promover o aprendizado ativo. Neste sentido, existe um interesse crescente na avaliação e uso da tecnologia digital em especial na área de simulação para o ensino e aprendizado da Estatística.

7. Referências

BAKER, J.; SUGDEN, S. Spreadsheets in Education –The First 25 Years. **Spreadsheets in Education**, 1(1), 2003. Disponível em <<http://epublications.bond.edu.au/ejsie/vol1/iss1/2>> Acessado em: 01 de maio de 2019.

BARR, G. D., SCOTT, L. Teaching statistics in a spreadsheet environment using simulation. **Spreadsheets in Education**, v.4, n.3, p.1-16, 2001.

BEN-ZVI, D. Toward understanding the role of technological tools in statistical learning. **Mathematical Thinking and Learning**, v.2, n.1, p.127-155, 2000.

BRUSSOLO, M. E. Understanding the Central Limit Theorem the Easy Way: A Simulation Experiment. **2nd Innovative and Creative Education and Teaching International Conference (ICETIC2018)**, Badajoz, Spain, 20–22 June 2018.

CHAAMWE, N.; SHUMBA, L. ICT Integrated Learning: Using Spreadsheets as Tools for e-Learning, A Case of Statistics in Microsoft Excel. **International Journal of Information and Education Technology**, v.6, n.6, p.435-440, 2016.

CHANCE, B., DELMAS, R. C., & GARFIELD, J. Reasoning About Sampling Distributions. In Ben-Zvi & J. Garfield (Eds.). **The challenge of developing statistical literacy, reasoning and thinking** (pp. 295-323). The Netherlands: Kluwer Academic Publishers, 2004.

CHANCE, B.; BEN-ZVI, D.; GARFIELD, J.; MEDINA, E. The role of technology in improving Student Learning of Statistics. **Technology Innovations in Statistics Education**, v.1, n.1, p.1-27, 2007.

CHANDRAKANTHA, L. Simulation using excel data tables in teaching introductory statistics. **Journal of Computing Sciences in Colleges**, v.29, n.3, p.29-34, 2014.

DE JONG, T.; VAN JOOLIGEN, W. Scientific discovery learning with computer simulations of conceptual domains. **Review of Educational Research**, v.68, n.2, p.179-201, 1998.

DELMAS, R. C.; GARFIELD, J.; CHANCE, B. L. A Model of Classroom Research in Action: Developing Simulation Activities to Improve Students' Statistical Reasoning. **Journal of Statistics Education**, v.7, n.3, p.1-16, 1999.

DINOV, I. D.; CHRISTOU, N.; SANCHEZ, J. Central limit theorem: New SOCR applet and demonstration activity. **Journal of Statistics Education**, v.16, n.2, p.1-15, 2008.

ERICKSON, T. Using Simulation to Learn about Inference. in A Rossman and B Chance, (eds.) **Proceedings of the Seventh International Conference on Teaching Statistics**. Voorburg, The Netherlands: International Statistical Institute, 2006. < <http://www.ime.usp.br/~abe/ICOTS7/Proceedings/index.html>>. Acessado em 05 de junho de 2019.

FERREIRA, R. S; KATAOKA, V. Y.; KARRER, M. Teaching probability with the support. **Statistics Education Research Journal**, v.13, n.2, p.132-147, 2014.

HAGTVEDT, R.; JONES, G. T.; JONES, K. Pedagogical Simulation of Sampling Distributions and the Central Limit Theorem. **Teaching Statistics**, v.29, n.3, p.94-97, 2007.

_____. Teaching confidence intervals using simulation. **Teaching Statistics**, v.30, n.2, p.53-56, 2008.

HANEY, M. H. A Spreadsheet Simulation to Teach Concepts of Sampling Distributions and the Central Limit Theorem. **Spreadsheets in Education**, 8(3), Article 3, 2015.

HUNT, N. Individualized Statistics Coursework Using Spreadsheets. **Teaching Statistics**, v.29, n.2, p.38-43, 2007.

_____. Using Microsoft Office to Generate Individualized Tasks for Students. *Teaching Statistics*. V.27, n.2, p.45-48, 2005.

ISLAM, M. R. Sample Size and Its Role in Central Limit Theorem (CLT). **Computational and Applied Mathematics Journal**, v.4, n.1, p.1-7, 2008.

KLAHR, D.; NIGAM, M. The equivalence of learning paths in early science instruction. **Psychological Science**, v.15, n.10, p.661-667, 2004.

LANE, D. M. E PERES, S. C. Interactive simulations in the teaching of statistics: Promise and Pitfalls. **Proceedings of the Seventh Annual Meeting of the International Conference on the Teaching of Statistics**, Salvador, Brazil, 2006. Disponível em: <<https://www.ruf.rice.edu/~lane/papers/interactive-simulations.pdf>>. Acessado em 10 novembro de 2018.

MAURER, K.; LOCK, D. Comparison of Learning Outcomes for Simulation-based and Traditional Inference Curricula in a Designed Educational Experiment. **Technology Innovations in Statistics Education**, v.9, n.1, 2016. Disponível em: <https://escholarship.org/uc/item/0wm523b0>. Acessado em 10 novembro de 2018.

MILLS, J. D. Using Computer Simulation Methods to Teach. Statistics: A Review of the Literature. **Journal of Statistics Education**, v.10, n.1, p.1-20, 2002.

NEUMANN, D. L.; NEUMANN, M. M.; HOOD, M. Evaluating computer-based simulations, multimedia and animations that help integrate blended learning with lectures in first year statistics. **Australasian Journal of Educational Technology**, v.27, n.2, p.274-289, 2011.

PACE, L. A.; BARCHARD K. A. Using a spreadsheet programme to teach introductory statistics: reducing anxiety and building conceptual understanding. **Int. J. Innovation and Learning**, v.3, n.3, p.267-283, 2006.

PFANNKUCH, M.; BROWN, C. M. Building on and Challenging Students' Intuitions About Probability: Can We Improve Undergraduate Learning? **Journal of Statistics Education**, v.4, n.1, p.1-15, 1996.

PATRON, H.; SMITH, W. J.; BOLD, D. Demonstrating the central limit theorem in the classroom: an excel exercise. **Journal for Economic Educators**, v.9, n.1, p. 1-15, 2009

RUGGIERI, E. Visualizing the Central Limit Theorem Through Simulation. **PRIMUS**, v.26, n.3, p.229-240, 2016.

SHI, N; HE, X.; TAO, J. Understanding Statistics and Statistics Education: A Chinese Perspective. **Journal of Statistics Education**, v.17, n.3, p.1-9, 2009.

TAYLOR, M. A. Simulating the central limit theorem. **The Stata Journal**, v.18, n.2, p.345-356, 2018.

TISHKOVSKAYA, S.; LANCASTER, G. A. Statistical education in the 21st Century: A review of challenges, teaching innovations and strategies for reform. **Journal of Statistics Education**, v.20, n.1, p.1-56, 2012.

TROSSET M. W. **An Introduction to Statistical Inference and Data Analysis**. Ebook.

2001. Disponível em:

http://inis.jinr.ru/sl/M_Mathematics/MV_Probability/MVas_Applied%20statistics/%D0%A2rosset%20Introduction.pdf. Acessado em 29 dezembro de 2020.

WARNER, C. B.; MEEHAN, A. M. Microsoft Excel As a Tool for Teaching Basic Statistics. **Teaching of Psychology**, v.28, n.4, p.295-298, 2001.

ANEXO I

Questões avaliadas

Questões para as três distribuições

Acione a tecla F9* 10 vezes e observe as formas dos histogramas do lado direito na parte superior (tamanho 300). O que acontece com a forma da distribuição da média amostral?

Acione a tecla F9 10 vezes e observe as formas dos histogramas do lado direito na parte inferior (tamanho 20). O que acontece com a forma da distribuição da média amostral?

Acione a tecla F9 10 vezes - Observando a média das médias dos histogramas do lado direito na parte superior e comparando com a média população, qual sua avaliação?

Acione a tecla F9 10 vezes - Observando a média das médias dos histogramas do lado direito na parte inferior e comparando com a média população, qual sua avaliação?

Acione a tecla F9 10 vezes. Compare os dois histogramas. O que acontece com a variabilidade da distribuição da média amostral à medida que n cresce de 20 para 300?

Questão final

Como você sabe, nos experimentos não vamos retirar 150 amostras, mas uma, duas, três ou um pouco mais. Se você sortearse uma amostra com 300 medidas e calculasse a média, pelos histogramas observados nesta atividade, em que local da distribuição seria mais provável localizá-la?

* A tecla F9 permite que a planilha executasse todos os cálculos e a confecção dos histogramas para cada uma das três distribuições de forma automática – processo de simulação.