

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO MATEMÁTICA

**O TEOREMA CENTRAL DO LIMITE: um estudo ecológico do saber e
do didático**

Chang Kuo Rodrigues

São Paulo

2009

Chang Kuo Rodrigues

**O TEOREMA CENTRAL DO LIMITE: um estudo ecológico do saber e
do didático**

Tese apresentada ao Programa de Pós-Graduação em Educação Matemática da Pontifícia Universidade Católica de São Paulo, como requisito parcial para a obtenção do título de Doutor em Educação Matemática.

Orientadora: Professora Doutora Cileda Queiroz e Silva Coutinho

São Paulo

2009

Chang Kuo Rodrigues
O teorema central do limite:
um estudo ecológico do saber e do didático

Trabalho apresentado à Banca
examinadora do Programa de Pós-
Graduação em Educação Matemática da
Pontifícia Universidade Católica de São
Paulo.

Professora Doutora Cileda Queiroz e Silva Coutinho (Orientadora) (PUC-SP)

Professora Doutora Celi Aparecida Espasandi Lopes (UNICSUL)

Professora Doutora Maria Lúcia L. WODEWOTZKI (Unesp-RioClaro)

Professor Doutor Benedito Antonio da Silva (PUC-SP)

Professora Doutora Maria José Ferreira da Silva (PUC-SP)

São Paulo, ____ de _____ de 2009

Dedico este trabalho às minhas quatro
filhas: Samantha, Amanda, Paula e
Roberta.

AGRADECIMENTOS

À minha orientadora, Professora Doutora Cileda Queiroz e Silva Coutinho, pelas incansáveis vezes de atenção disponíveis a mim, tornando possível a realização deste trabalho.

Aos membros convidados da Banca Examinadora: Professora Doutora Celi, Professora Doutora Maria Lúcia, Professor Doutor Benedito e Professora Doutora Maria José, pela disponibilidade da partilha deste momento.

Ao Colégio Cristo Redentor, que participou com incentivos moral e material.

Aos meus colegas que me acompanharam durante toda essa longa jornada.

Aos meus amigos e amigas que estiveram sempre comigo, seja na presença, seja nas palavras, pois, mesmo sem ter conhecimento, me ajudaram a superar obstáculos e enfrentar novos desafios.

À toda minha família, especialmente ao Tuca, por ter compreendido minhas ausências.

Às minhas filhas e netas pelos momentos em que tiveram que ouvir tantos “nãos” por estar ocupada neste trabalho, mas que sirva de exemplo de que essa mãe e essa avó, jamais foi, é ou será vencedora e, sim, guerreira, mesmo perdendo muitas lutas, mas que a palavra “desistir” não faz parte da vida.

Só há um meio de resgatar a nossa
dignidade, intelectualidade ou a riqueza
que existe dentro de nós: é reconhecer a
superioridade dos humildes.

RESUMO

O presente trabalho refere-se à construção das ideias e dos conceitos matemáticos e/ou estatísticos em torno do Teorema Central do Limite para os Licenciandos de Matemática. O cerne da investigação limita-se à importância do teorema na Inferência Estatística e à sua compreensão pelos futuros profissionais que atuarão na Educação Básica. Nesse sentido, optamos por revisar algumas bibliografias que têm relação com o processo de ensino e de aprendizagem do teorema e enfatizamos sua importância na prática do dia a dia do professor de Matemática. O quadro teórico incide sobre as teorias da Didática da Matemática, particularmente, a Teoria da Transposição Didática (CHEVALLARD, 1985), munido de uma abordagem ecológica sob o ponto de vista do saber e do didático (ARTAUD, 1998). Optamos por procedimentos metodológicos voltados para o *design didático* (ARTIGUE, 2009), de cunho qualitativo e, cujos pressupostos estão aliados à Engenharia Didática (ARTIGUE, 1988). Os sujeitos dessa investigação são os licenciandos que já dispunham de conhecimentos sobre a Estatística Básica e, a partir de uma análise prévia sobre que tipos de conhecimento eles já detinham sobre o tema, apresentamos algumas atividades no contexto de uma situação-problema pertinente ao cotidiano dos professores de Matemática. A análise desses resultados nos propiciou interrelacionar as problemáticas existentes na disciplina de Matemática com alunos da Educação Básica, envolvendo assim, a *literacia* estatística. Após a realização dessas atividades, ocorreu também um diálogo, com discussões acerca do tema, o que nos permitiu analisar como foram construídos as ideias e os conceitos no entorno do Teorema Central do Limite, de modo que sua compreensão fosse o principal alvo para os licenciandos. Além disso, analisamos alguns livros-texto do ensino superior, à luz da Teoria Antropológica do Didático (CHEVALLARD, 1996, 1999), o que também nos indicou que saberes são indispensáveis de modo que o teorema “viva”, já que a abordagem é sob o ponto de vista ecológico do saber e do didático. Por outro lado, detectamos que tipos de limitações, ou restrições, existem nas obras consultadas, interferindo assim, a elaboração das atividades por parte do professor. Portanto, a nossa investigação reitera a importância do ensino e da aprendizagem da Estatística nas diversas aplicações na formação dos futuros

professores de Matemática num mundo ditado pelos avanços tecnológicos, que interferem diretamente na leitura de informações que recebemos a todo instante.

Palavras-chave: Teorema Central do Limite. Ensino de Estatística. Inferência Estatística.

ABSTRACT

This paper refers to the building of mathematical and/or statistical ideas and concepts around Central Limit Theorem for Mathematics graduates. The investigation focuses the importance of the theorem in Statistics Inference and its comprehension by the professionals to be, who will act in Basic Education. Therefore, we chose to research some books related to the teaching and learning process of the theorem and emphasised its importance on the Mathematics teacher daily practice. The theoretical approach is about Mathematics Teaching theories, particularly the Theory of Didactic Transposition (CHEVALLARD, 1985), with an ecological approach under the knowledge and teaching point of view (ARTAUD, 1998). We chose methodological procedures directed to the *didactic design* (ARTIGUE, 2009), with qualitative nature, and whose assumptions are linked to Teaching Engineering (ARTIGUE, 1988). The subjects of this investigation are the graduates who had some knowledge about Basic Statistics and, from a previous analysis about the kind of knowledge they had about the theme, we presented some activities in a problem-situation context connected to the Mathematics teachers' daily practice. The analysis of these results allowed us to relate the existing problems between the subject and the students from Basic Education, which involved statistics *literacy*. After these activities, there was a dialogue, with discussions about the theme, allowing us to analyse how the ideas and concepts around the Central Limit Theorem were built, being its comprehension the main aim for the graduates. Besides that, we analysed some textbooks for higher education, based on the Anthropological Theory of Didactic (CHEVALLARD, 1996, 1999), which also showed us the essential knowledge for the theorem to "live", because the approach is under the knowledge and teaching ecological point of view. On the other hand, we detected what kind of limitations, or restrictions, exist in the books analysed, interfering in the elaboration of the activities by the teacher. Thus, our investigation reaffirms the importance of teaching and learning Statistics in the various applications for the Mathematics teachers to be formation in a world controlled by the technological advances, which interfere directly on the understanding of the information we receive every moment.

Keywords: Central Limit Theorem. Statistics Teaching. Statistics Inference

LISTA DE FIGURAS

FIGURA 1	A Ideia de Uma Distribuição Amostral	69
FIGURA 2	O Gráfico para Distribuição Normal para X e $\bar{X}$	86
FIGURA 3	A Distribuição de Observação	89
FIGURA 4	Distribuição Uniforme com 300 Repetições $\alpha = 2,9$ e $\beta = 5,39$	90
FIGURA 5	Distribuição Logonormal com 300 Repetições $\alpha = 0$ e $\beta = 1$	90
FIGURA 6	Distribuição Bimodal com 300 Repetições	91
FIGURA 7	Distribuição Exponencial com 300 Repetições	91
FIGURA 8	Tabela da Distribuição da Variável Y	93
FIGURA 9	Distribuição de Y e $\bar{Y}$	93
FIGURA 10	A Distribuição de Observação	96
FIGURA 11	A distribuição exata (tracejada) e a aproximação Normal a partir do Teorema Central do Limite (contínua) para o tempo médio necessário para manutenção de um ar-condicionado	97
FIGURA 12	As porções sombreadas são iguais à probabilidade de uma média amostral inferior a 19,8 ou superior a 20,2	100
FIGURA 13	Determinação da área sombreada da distribuição amostral	103
FIGURA 14	Distribuição de Pesos de homens	105
FIGURA 15	Distribuição de Médias Amostrais de 36 homens	106
FIGURA 16	Distribuições: Populacional, da Amostra, Amostral Observada e Amostral Teórica	110

LISTA DE TABELAS

TABELA 1	Atividades nos Livros-Texto	82
TABELA 2	Índice de Autoestima da Situação-Problema Proposta aos Licenciandos	122
TABELA 3	Atividade realizada por BM para o item 7	133

LISTA DE QUADROS

QUADRO 1 Denominação dos Livros-Texto	81
QUADRO 2 Definição do Teorema Central do Limite nos Livros-Texto	108

LISTA DE ESQUEMAS

ESQUEMA 1	Dinâmica da Pesquisa-Ação de Nossa Investigação, com Base em Thiollent (1986)	51
ESQUEMA 2	Esboço de um Ecossistema para o Teorema Central do Limite..	58
ESQUEMA 3	Cadeia Elementar Simples para o Teorema Central do Limite.....	61
ESQUEMA 4	Esquema de Redes Tróficas em um Ecossistema Mostrando a Partilha entre Dois Ecossistemas: o Teorema Central do Limite Enquanto Objeto Matemático e Objeto da Estatística	62
ESQUEMA 5	Ecossistema do Teorema Central do Limite	70
ESQUEMA 6	Representação de um Ecossistema Sob o Ponto de Vista do Didático	74
ESQUEMA 7	Estrutura ecológica para o Teorema Central do Limite no livro-texto A1	84
ESQUEMA 8	Ecossistema da Atividade Proposta por A1 no Contexto Matemático e Estatístico	87
ESQUEMA 9	Ecossistema para o Teorema Central do Limite na Atividade (c).....	101

SUMÁRIO

1	INTRODUÇÃO	17
2	UMA INVESTIGAÇÃO MOTIVADA PELA EXPERIÊNCIA	21
3	REVISÃO DA LITERATURA	27
3.1	O Teorema Central do Limite Sob Ótica da Educação	27
3.2	Um Breve Cenário Histórico do Teorema Central do Limite.....	35
4	PROCEDIMENTOS E REFERENCIAIS METODOLÓGICOS.....	43
4.1	A Análise dos Livros-Texto e a Teoria Antropológica do Didático....	47
4.2	As Atividades Didáticas e os Licenciandos	48
4.3	A Pesquisa e a Prática: o Design Didático.....	52
5	QUADRO TEÓRICO DA INVESTIGAÇÃO	54
5.1	A Problemática Ecológica sob o Ponto de Vista do Saber	55
5.1.1	<i>A Noção de Ecossistema</i>	55
5.1.2	<i>Abordagem Ecológica do Saber: O Teorema Central do Limite</i>	57
5.2	Abordagem Ecológica do Didático.....	71
5.3	Literacia Estatística.....	75
6	A PESQUISA	80
6.1	O Teorema Central do Limite e os Livros-Texto	81
6.1.1	<i>Análise Quantitativa dos Livros-Texto</i>	82
6.1.2	<i>Apresentação e Análise de Atividade no Livro-Texto</i>	83
6.1.2.1	<u>Atividades do Livro-Lexto A1</u>	83
6.1.2.2	<u>Atividades do Livro-Lexto A2</u>	90
6.1.2.3	<u>Atividades do Livro-Lexto A3</u>	92
6.1.2.4	<u>Atividades do Livro-Lexto A4</u>	95
6.1.2.5	<u>Atividades do Livro-Lexto A5</u>	95
6.1.2.6	<u>Atividades do Livro-Lexto A6</u>	98
6.1.2.7	<u>Atividades do Livro-Lexto A7</u>	104
6.1.3	<i>Definição do Teorema Central do Limite nos Livros-Texto</i>	107
6.2	A Análise da Investigação Sob o Ponto de Vista Ecológico do Didático: o Teorema Central do Limite e os Licenciandos	112
6.2.1	<i>As Atividades com os Licenciandos</i>	113

6.2.2	<i>A Primeira Atividade com os Licenciandos</i>	114
6.2.3	<i>Noções Preliminares das Atividades</i>	118
6.2.4	<i>Análise das Atividades 2 e 3</i>	124
6.3	O Diálogo: Intervenções Controladas no Design Didático	136
7	CONSIDERAÇÕES FINAIS	145
	REFERÊNCIAS	148
	APÊNDICE	155
	ANEXOS	180

1 INTRODUÇÃO

O presente trabalho surgiu a partir da motivação e curiosidade por um olhar mais atento sobre o Ensino da Estatística nos cursos superiores, particularmente, no Curso de Licenciatura em Matemática. O foco está direcionado aos futuros professores, porque acreditamos que a mudança aqui proposta se inicia na Educação Básica, segmento em que eles atuarão.

A problemática existente na formação dos professores de Matemática em Estatística, em especial, no estudo da Inferência Estatística, muitas vezes fica a desejar, já que esse ramo da Estatística não é direcionado explicitamente para a Educação Básica. No entanto, a maioria das pesquisas existentes nos meios de comunicação se faz por amostras, que são partes de uma população. Por isso, não podemos ignorar a relevância social que há por trás do estudo da Inferência Estatística.

Diante desse cenário, os sujeitos de nossa pesquisa são os licenciandos em Matemática, tendo em vista que serão os futuros profissionais da Educação. Dessa forma, entendemos que eles devem ter o conhecimento da importância do Teorema Central do Limite. Isso nos reporta a compreender como se estabelecem as ideias matemáticas e estatísticas no entorno do **teorema** de modo que haja a consolidação da aprendizagem por parte deles. Para tal, usaremos uma estrutura ecológica do saber e do didático para o tema de nosso estudo.

Destacamos que o nosso trabalho não está voltado para a demonstração do teorema, mas, sim, essencialmente, para a construção de ideias que surgem a partir de atividades propostas aos licenciandos. Deve-se esclarecer aqui que não se pretende propor o ensino do teorema na Educação Básica, pois entendemos que os professores é quem devem ter esse conhecimento de modo a garantir seus argumentos quando postos em questionamentos na prática. São eles que devem compreender os contextos e as análises de dados, sabendo que um aprofundamento já seria no campo da Inferência Estatística. Em analogia, um professor precisa conhecer estruturas algébricas para compreender a álgebra ensinada na Educação Básica, sem a necessidade de ensinar essas estruturas aos alunos.

Por conseguinte, a problemática de nossa investigação incide sobre a compreensão da importância do Teorema Central do Limite, sob o ponto de vista ecológico do saber e do didático, a partir de questões como: **Por que determinados saberes/conhecimentos deixaram de ser abordados no ensino do Teorema Central do Limite, mas ainda assim ele “sobrevive”? Que saberes são indispensáveis de modo que permite o teorema “viver”? Que relações diretas ou indiretas existem entre o teorema e a construção de saberes estatísticos nas aulas de Estatística?** Essas questões surgiram quando nos deparamos com o trabalho de Artaud (1998), sobre a ecologia das organizações matemáticas e didáticas no estudo da Álgebra, o que nos permitiu fundamentar a nossa investigação no estudo do teorema com os licenciandos da Matemática.

Este trabalho foi dividido em cinco capítulos. No primeiro, expomos as motivações que foram responsáveis pelas primeiras ideias acerca da investigação nesse campo do conhecimento e suas implicações sob as dimensões sociais, educacionais e profissionais.

No segundo capítulo, intitulado Teorema Central do Limite sob a ótica da Educação, tratamos da revisão da literatura, dividida em duas partes: a primeira refere-se a comentários sobre alguns artigos, devidamente selecionados, providos do enfoque de ensino e de aprendizagem do teorema. Optamos pela expressão: ‘processo de ensino e de aprendizagem’, por entender que existe um amplo campo de variáveis associado a essa dinâmica, tendo em vista que a intenção de um sujeito, no caso o(a) professor(a), ao elaborar uma situação de ensino, nem sempre atinge seu objetivo com aquele a quem deseja ensinar, ou seja, apesar de o aluno fazer parte do processo de aprendizagem, não significa necessariamente que ele aprende. Na segunda etapa, optamos por um breve esboço histórico de como o teorema surgiu, partindo daqueles que foram os primeiros interessados, fato esse gerado pela curiosidade em conhecer os fenômenos da natureza.

Vale ressaltar que a nossa investigação não está voltada especificamente para a parte histórica. Por isso, achamos interessante nos limitarmos a uma breve exposição de como as ideias acerca do Teorema Central do Limite surgiram no final do século XVIII, culminando em sua formalização, na estrutura matemática, fato que ocorreu somente na primeira metade do século XX.

Os procedimentos metodológicos da pesquisa, devidamente fundamentados, farão parte do terceiro capítulo, em que apresentaremos a nossa opção pelo

entrelaçamento da teoria e da prática, de modo que efetivamente haja a interação entre elas, formando, assim, um *design* para o nosso trabalho. Essa metodologia segue a proposta de Artigue (2009). Apesar de ser um estudo recente, segue os progressos que vêm ocorrendo nas pesquisas em Educação Matemática, ao tratar do *design didático* aliado à Engenharia Didática (ARTIGUE, 1988) e à Teoria Antropológica do Didático (CHEVALLARD, 1996, 1999). Vale ressaltar que a Teoria Antropológica do Didático, sob a perspectiva ecológica do saber e do didático, foi decisiva na análise dos resultados deste trabalho.

Uma das potencialidades dessa metodologia, o *design didático*, incide sobre a forma como a construção das teorias pode causar impactos na prática, proporcionando-nos uma atitude reflexiva diante da nossa proposta didática. Além disso, identificamos como atores desse trabalho todos aqueles que efetivamente estavam envolvidos na dinâmica do *design*, ou seja, os pesquisadores e os pesquisandos e, a partir disso, nos foi possível, também, aliar alguns fundamentos da pesquisa-ação de Thiollent (1986). A sua pertinência se deve ao fato de que em, certas etapas da pesquisa, um dos pesquisadores atuou também como professor.

A construção do nosso quadro teórico nos motivou buscar a problemática ecológica existente no saber e no didático (ARTAUD, 1998) à luz da Teoria Antropológica do Didático proposta por Chevallard (1985). Fazemos, também, referência à importância da *literacia estatística*. Nesse caso, mesmo reconhecendo que a autora traduz *literacy* para *letramento* na Língua Portuguesa, a nossa opção foi a de manter essa expressão que, segundo Soares (2005, p. 36), “[...] designa o estado ou condição daquele que é *literate*, daquele que não só sabe ler e escrever, mas também faz uso competente e frequente da leitura e escrita”.

No quinto e último capítulo, associamos o nosso quadro teórico com as atividades que foram realizadas com licenciandos de Matemática. Como a nossa investigação é provida de caráter qualitativo, o grupo participante era composto por quatro licenciandos, o que nos permitiu observar e analisar os fenômenos didáticos que constituíram o *corpus* pelo qual responderíamos nossa questão de pesquisa.

Todas as atividades foram anexadas no final do trabalho, incluindo a demonstração do Teorema Central do Limite, mesmo entendendo que não foi foco de nossa investigação. Acreditamos que seja importante conhecer a evolução dos procedimentos específicos no modelo matemático.

Por fim, desejamos que este trabalho se constitua em uma pequena contribuição para a Educação Estatística, com ênfase na Inferência Estatística, a qual, nas palavras de Salsburg (2009, p. 11), configura-se em “um dos usos da estatística mais amplamente apreciados tem sido a pesquisa por amostragem.”

2 UMA INVESTIGAÇÃO MOTIVADA PELA EXPERIÊNCIA

O estudo da Estatística tem sido foco de muitas discussões na Educação Matemática. Esse fato é fortemente evidenciado desde as séries iniciais da Educação Básica, a partir da proposta do currículo de Matemática previsto nos Parâmetros Curriculares Nacionais (1998). Quatro blocos sustentam os PCN de Matemática: Números e Operações, Espaço e Forma, Grandezas e Medidas e Tratamento de Informação e, nesse contexto, o grande desafio é “o de identificar, dentro de cada um desses vastos campos que conceitos, procedimentos e atitudes são socialmente relevantes.” (BRASIL, 1998, p. 49)

Um dos aspectos inovadores desse documento surge com a inclusão do bloco Tratamento de Informação, uma consequência do desenvolvimento social, regida pelos avanços tecnológicos, que gerou demanda para uma mudança curricular. Isso nos reporta a compreender como tratar os dados, que são transformados em informações ao recebermos, seja por meio de estatísticas, tabelas ou gráficos. Então, na Educação Básica, qual será o ponto de vista do professor de Matemática em relação a esse bloco? A partir de trabalhos realizados com professores de Matemática¹, durante catorze anos, nos foi possível constatar vários obstáculos, quando o assunto a ser tratado era Estatística – ao menos no início da implementação dos PCN, no final da década de 90. Um dos agravantes era que, se algum tema tivesse de ser “sacrificado”, seria essa disciplina ou algum tópico relacionado a ela.

Alguns motivos que levaram os professores a optarem por essa atitude podem estar atrelados às suas próprias crenças, adquiridas ao longo de sua formação, pois a própria disciplina Estatística nem sempre foi ou é parte do currículo no curso de Licenciatura. Por exemplo, em alguns cursos de Licenciatura Plena em Matemática no início dos anos 80, como o Centro de Ensino Unificado de Brasília, conhecido na época como CEUB e hoje, como UniCEUB, em sua grade curricular, a

¹ Desde 1995, a autora, Chang Kuo Rodrigues, coordena um grupo de professores de Matemática da Educação Básica de uma escola privada de Juiz de Fora. E, sistematicamente, as reuniões pedagógicas constituíam-se de estudos de grupo voltados para discussão da prática. Quando era necessário, Rodrigues buscava teorias pedagógicas ou específicas, da Matemática, de forma que esses encontros tornaram-se uma pequena contribuição na formação continuada do profissional.

disciplina de Estatística não constava, conforme podemos constatar o histórico curricular no anexo 1.

No trabalho realizado com os professores da Educação Básica descritos acima, durante a discussão dos elementos norteadores presentes nos documentos oficiais, no final da década de 90, constatamos que a maioria do grupo deles, professores² de Matemática de nossa escola, obteve sua formação docente no período que corresponde ao final da década de 70 e início de 80. Dessa forma, podemos dizer que o desconhecimento do saber estatístico foi um dos componentes que gerou motivação, no sentido de promover um círculo de estudos sobre as novas propostas apresentadas pelo MEC (Ministério da Educação), sob um olhar mais atento para o bloco Tratamento de Informações.

Nessa mesma direção, porém atuando em outro nível de estudo, o referido trabalho nos permitiu outros pontos de vista com relação à Estatística pelo fato de que o trabalho também realizou-se em quatro cursos distintos do Ensino Superior: Estatística para os cursos de Ciências Biológicas, Geografia, Tecnologia em Marketing e Ciências com Habilitação em Matemática³.

Dessa forma, a motivação inicial para realização dessa pesquisa partiu de um trabalho que tratou de um mesmo tema sob cinco perspectivas distintas, o que nos permitiu uma visão mais ampla do que vem a ser Estatística no sentido global e específico, devido às particularidades exigidas em cada área. Havia um eixo comum em termos de tópicos, tais quais as medidas de tendência central, de dispersão e análise exploratória de dados, na teoria das probabilidades e na Inferência Estatística.

No entanto, as particularidades do ensino de Estatística incidiam no contexto da necessidade de cada área: no Curso de Ciências Biológicas, o interesse está voltado para questões relativas aos temas sobre saúde, sistemas ecológicos, entre outros; no Curso de Geografia, a Estatística, particularmente a Descritiva, ao trabalhar com recenseamento da população, entre outros, favorece abordagens de

² Esse grupo era composto por 14 de professores de Matemática, de uma escola particular de Juiz de Fora. A formação desse grupo surgiu a partir do evento promovido pelo Sindicato dos Estabelecimentos Particulares de Ensino de Juiz de Fora-MG (SINEPE-JF), convidando alguns autores dos PCN para um encontro com os professores da cidade, no ano de 1997. Posteriormente, sob a coordenação da autora deste trabalho, iniciaram-se encontros semanais para reflexão e discussão sobre o texto.

³ Neste período, além de coordenar o grupo de professores da Educação Básica de uma escola privada, a autora lecionava também em Instituições de Curso Superior como professora da disciplina Estatística.

cunho geopolítico; na Tecnologia em Marketing, são abordados anúncios do dia a dia que aparecem nos meios de comunicação ou a forma como os números definem as tomadas de decisão, como, por exemplo, no momento da escolha de algum produto; no curso de Licenciatura de Matemática, a abordagem deve contemplar dois aspectos: os objetos matemáticos utilizados na Estatística e suas comprovações e, por outro, sua aplicação em situações hipoteticamente reais; e, por fim, na formação continuada dos professores que atuavam na Educação Básica.

Por isso, a relevância acadêmica desta pesquisa justifica-se por dois aspectos. O primeiro está relacionado ao licenciando de Matemática, pois é ele ou ela que estará atuando no ensino da Estatística na Educação Básica. E, além disso, ele(ela) é um(a) adulto(a) que tem suas próprias crenças ao fazer leitura de dados. Holmes (2000) cita dois parágrafos do *National Curriculum in Mathematics* entre as décadas de 80 e 90 sobre o ensino de Estatística:

'775: Estatística não é apenas um jogo de técnicas, mas sim uma atitude mental de tratamento de dados. Em particular, reconhece a incerteza e a variabilidade num conjunto de informações. Torna a pessoa apta a tomar de decisões em situações de incerteza.

'781: *Numeracia* estatística exige um sentido para os números: apreciando os níveis de precisão, o bom senso de fazer estimativas, uma abordagem de senso comum nos dados com suporte em um argumento, a tomada de consciência de uma variedade de interpretações de números e uma judiciosa compreensão dos conceitos amplamente utilizados como médias e porcentagens. Tudo isso são partes do cotidiano. (HOLMES, 2000, p. 51) (Tradução nossa; Grifo nosso).⁴

O segundo é referente à importância que exercerá o papel de professor no 'processo de ensino e de aprendizagem' de Estatística, pois muitas vezes e, na maioria dos Cursos Superiores, fica a cargo do professor de Matemática a responsabilidade de mediar esse tema com os alunos.

Nesse sentido, o cenário no Brasil não é diferente do apresentado pelo *National Curriculum in Mathematics*, como podemos constatar nos Parâmetros Curriculares Nacionais para o Ensino Fundamental (1998), no bloco de conteúdo nomeado de *Tratamento de Informação*, cuja finalidade é abordar os conteúdos da

⁴'775 Statistics is not just a set of techniques; it is an attitude of mind in approaching data. In particular it acknowledges the fact of uncertainty and variability in data collection. It enables people to make decisions in the face of this uncertainty.

'781 Statistical numeracy requires a feel for numbers, an appreciation of levels of accuracy, the making of sensible estimates, a common-sense approach to data in supporting an argument, the awareness of the variety of interpretation of figures and a judicious understanding of widely used concepts such as mean and percentages. All these are part of everyday living. (Texto original)

Estatística, Probabilidade e Combinatória. Notamos no texto desse documento uma vontade explícita de ressaltar a importância de uma leitura real da sociedade. A própria etimologia da palavra *informação* não se limita apenas a sua apropriação cumulativa, posto que o propósito dirige-se para o sujeito que, acima de tudo, tem a possibilidade de mudar a realidade no qual está inserido. Mudar a realidade pode ser um tanto paradoxal, pois questões como ‘*a favor de quem?*’; ‘*para qual objetivo?*’, são básicas e devem ser trabalhadas com afinco. Afinal, teoricamente, todos almejam por uma sociedade mais justa.

Diante desses fatos, o escopo deste trabalho é investigar como os licenciandos de Matemática constroem as ideias e os conceitos que estão no entorno do Teorema Central do Limite. Entendemos as ‘ideias’ como conceitos intuitivos que surgem a partir da vivência e crença do sujeito e por ‘conceito’, tal como Artigue (1990) o define quando estuda a evolução do tratamento de um determinado saber, partindo de um contexto que diz respeito a este saber, mesmo na elaboração de situações artificiais, porém hipoteticamente reais, que podem ocorrer no cotidiano da sala de aula.

Reforçando o fato de que a Estatística tem seu papel fundamental na leitura de mundo e, portanto, engendra facilmente situações-problema reais para a sala de aula, faz-se imprescindível preocupar-se com a forma como essa área do conhecimento irá ser abordada pelos futuros professores. E, nessa direção, Contreras (2002) reforça que:

[...], a responsabilidade profissional dos professores é estarem suficientemente conectados com suas comunidades para entender suas demandas e compreender os interesses de seus alunos, mas distantes o suficiente para poderem cultivar nos seus alunos o distanciamento crítico necessário que lhes permita reconsiderar estes interesses e demandas frente a outros com os quais entram em conflito. (CONTRERAS, 2002, p. 203)

Esse distanciamento entre os sujeitos, professor e alunos, estabelece um compromisso com valores educativos, isto é, as interpretações de dados estatísticos não são impostos, mas, sim, administrados por cada um, com autonomia, para reflexão e discussão. Diante desse contexto, na formação do professor pressupõe-se o entrelaçar-se entre a prática e a teoria, administrando os conflitos, de modo que a criticidade seja alvo de reflexão e, sobretudo, em ação: a literacia estatística. Isso justifica os sujeitos de nossa investigação que são os licenciandos em Matemática.

Ademais, temos ainda outra relação a analisar: o sujeito professor e a área do conhecimento a ser abordada, a Estatística, pois Besson (1995) nos alerta que:

[...] se as estatísticas fossem apenas um *reflexo* (concepção fotográfica), disporiam de um critério de verdade (a exatidão). Não é este o caso: elas não provêm da denotação, mas da conotação, pois é o contexto, o contorno, que determina seu sentido. De fato, o estatístico não escolhe seus índices. A realidade lhe aparece então pré-modelada pelas categorias já existentes na representação ou na prática individual, social, administrativa. (BESSON, 1995, p. 52)

A concepção do profissional e sua criticidade tornar-se-ão de extrema importância na abordagem da Estatística: os obstáculos não se restringem apenas ao modo de mediar a construção do conhecimento estatístico com os alunos. A reciprocidade de uma realidade a ser reconstruída requer a superação de conflitos diante da diversidade cultural de cada um. Entende-se que a análise exploratória de dados é uma forma de se olhar para um conjunto de dados de forma crítica, pois não se limita a algoritmos e procedimentos, indo até a construção de significados dos resultados produzidos dentro de um determinado contexto.

Entretanto, muitas vezes os resultados analisados, para um subconjunto de dados, são inferidos para o conjunto total (população). Isso reforça a importância do Teorema Central do Limite para alunos da Licenciatura em Matemática, pois contribuirá na construção da rede de conhecimentos que favorecem o exercício da consciência crítica naqueles que estarão atuando com a Educação Básica.

O Teorema Central do Limite garante que a série constituída pelo conjunto das médias amostrais converge para a média da população, desde que o tamanho das amostras, n , seja significativamente grande para que tal fato aconteça. A relevância dos conhecimentos prévios sobre os conceitos de variáveis aleatórias discretas e contínuas, média, variabilidade, noções sobre os erros estatísticos, é de natureza extremamente importante para a compreensão desse teorema. Todas essas noções deveriam fazer parte do conteúdo a ser desenvolvido na disciplina Estatística dos cursos de Licenciatura em Matemática. No entanto, nos cursos com os quais temos convivido em nossa prática docente, a abordagem, quando era feita, o era de forma tecnicista, ou seja, limitando-se à aplicação de fórmulas.

Assim, a partir de uma motivação proveniente da prática aliada à força tecnológica do mundo contemporâneo, passamos a repensar sobre como podemos interferir na formação dos futuros professores. Mais precisamente, como se

constroem os conhecimentos, no nosso caso, o Teorema Central do Limite, com aqueles que irão atuar na Educação Básica?

O nosso trabalho segue essa direção e, para tentar responder a esta e a outras questões já citadas, inicialmente exporemos alguns trabalhos que foram realizados nesse campo, bem como apresentaremos um breve cenário histórico do teorema, como será vista no próximo capítulo.

3 REVISÃO DA LITERATURA

Neste capítulo apresentaremos uma revisão da literatura que empreendemos com o objetivo de averiguar como o nosso tema de investigação se coloca frente a outras pesquisas relacionadas.

A revisão bibliográfica a seguir foi dividida em duas partes. Na primeira, buscamos responder às seguintes questões: no campo da Educação Matemática, quais trabalhos abordam o Teorema Central do Limite? Quais deles discutem a abordagem do teorema em cursos de Licenciatura em Matemática? Quais trabalhos surgem o teorema associado ao processo de ensino e de aprendizagem?

Na segunda parte, relatamos uma breve leitura histórica da origem e do desenvolvimento do referido teorema, desde as suas primeiras ideias até à sua formulação teórica rigorosa, baseado nos trabalhos de Paulauskas (2006), Mether (2003), Fischer (2000), Cam (1986), Gnedenko e Kolmogorov (1954), além de alguns historiadores matemáticos como Cajori (2007), Eves (1997), Struik (1987) e Rodriguez (1989). Entendemos que a leitura histórica nos proporcionará importantes elementos para a compreensão do Teorema Central do Limite.

3.1 O Teorema Central do Limite Sob Ótica da Educação

Partimos da constatação de que o ensino de estatística tem sido um dos temas de pesquisas na Educação Matemática nos últimos dez anos, o que fez emergir uma nova linha de pesquisa, a Educação Estatística. As publicações⁵ mais recentes sobre o ensino do Teorema Central do Limite podem ser encontradas em revistas e jornais eletrônicos como *Journal of Statistics Education* e *Statistics Education Research Journal* (SERJ), entre outros. Outras contribuições de natureza investigativa são apresentadas em eventos internacionais e nacionais, como, por exemplo, *International Conference on Teaching Statistics* (ICOTS), *Congress of the European Society for Research in Mathematics Education* (CERME), *Congreso*

⁵ Procuramos citar os eventos que possuem Grupos de Trabalho (GT) em Educação Estatística.

Iberoamericano de Educación Matemática (CIBEM), *International Congress on Mathematical Education* (ICME), *Conferência Interamericana de Educação Matemática* (CIAEM), *Seminário Internacional de Pesquisas em Educação Matemática* (SIPEM), *Simpósio Internacional de Pesquisas em Educação Matemática* (SIPEMAT), *Encontro Nacional de Educação Matemática* (ENEM), entre outros.

Iniciamos nossa revisão com o trabalho de Glencross (1986), publicado nos anais do ICOTS-2, cujo objetivo foi apresentar a importância do Teorema Central do Limite no ensino da Estatística e discriminar algumas atividades que podem favorecer a aprendizagem do mesmo. A metodologia utilizada por ele é a pesquisa de fonte bibliográfica com ênfase na epistemologia do teorema.

Esse artigo pode ser considerado o marco inicial do estudo do Teorema Central do Limite na Educação Estatística, pois nele encontramos algumas questões didáticas que foram expostas de forma a chamar atenção para alguns princípios de aprendizagem no seu estudo. Segundo o autor:

Para estudantes de um curso introdutório de estatística, incluindo inferência estatística, o Teorema Central do Limite é explicitamente uma situação *sine qua non*. Contudo, embora o teorema possa ser visto como base fundamental da inferência estatística, experiências registraram que muitos estudantes, durante vários anos, tiveram muitas dificuldades em compreendê-lo e, além disso, o conceito geral de uma distribuição de amostras teóricas pode ser demasiadamente abstrato para ser amplamente apreciada num primeiro contato. Mesmo alunos com bases matemáticas bem consolidadas, aqueles que apreciam a demonstração do teorema, pode vir a ter dificuldades de compreensão em seu significado. (GLENCROSS, 1986, p. 92) (Tradução nossa)⁶

Nesse contexto, o autor busca entender as dificuldades presentes no cenário em que ocorre o ensino do Teorema Central do Limite para estudantes universitários a partir da teoria de Skemp (1971). Essa teoria é um modelo para aprendizagem no desenvolvimento das teorias matemáticas, no qual se estabelecem com dois princípios:

⁶ "For students taking an introductory statistics course which includes inferential statistics, the Central Limit Theorem is clearly a *sine qua non*. However, although this theorem may be regarded as the cornerstone of statistical inference, it has been the writer's experience over a number of years that many students find it difficult to understand and that the general concept of a theoretical sampling distribution may be too abstract to be fully appreciated when first met. Even students with a good mathematical background, who can appreciate a proof of the theorem, may have difficulty understanding its significance." (Texto original)

- O princípio da variabilidade perceptiva, posteriormente chamado de “*multiple embodiment*”⁷, que, sob o nosso contexto, significa que “abstrair uma estrutura matemática, efetivamente, encontra inúmeras situações diferentes para perceber suas propriedades puramente estruturais” (SKEMP, 1971, p. 32);
- O segundo é o princípio da variabilidade matemática baseado em Dienes (1963, p. 156), para quem “[...] como todo conceito matemático envolve variáveis essenciais, todas essas variáveis matemáticas precisam ser transformadas se o objetivo é alcançar a generalidade do conceito matemático”⁸. Apesar de o autor tratar da construção de conceito, devemos considerar as variáveis aleatórias no caso do teorema em estudo, considerando-se o nosso contexto.

Sob esse aspecto, Glencross (1986), tal como Skemp (1971), defende a idéia de que quem mais precisa conhecer cada um desses princípios é o próprio professor. E, nesse sentido, o autor estende esse princípio, para os professores, em três: deve estar provido de coleções de exemplos adequados para o conceito; deve garantir que o conceito seja conhecido em inúmeras diferentes situações; deve alegar que todas as variáveis envolvidas no conceito são aleatórias.

Glencross (1986) esclarece que, se o professor faz uso desses três elementos básicos, quando explicitados, servem como ponto de partida para nortear as atividades voltadas para aprendizagem dos alunos.

Um outro aspecto do artigo a ser destacado é quando o autor afirma que os membros do *Joint Education Committee of the Royal Statistical Society* e o *Institute of Statisticians* têm notado críticas dos estatísticos em relação ao ensino nesse nível, pois, em particular, muitas vezes, os próprios professores desse segmento ignoram as situações práticas em prol do excesso de manipulação formal dos objetos matemáticos.

As atividades propostas são de diferentes tipos de distribuições para a abordagem do Teorema Central do Limite, que faz uso de tecnologias para gerar

⁷ Optamos por não traduzir essa expressão por entender que na Psicologia da Educação Matemática tem sido um termo bastante usado entre os pesquisadores. Mas, a título de entendimento, mesmo de maneira superficial, poderia ser traduzida como “personificações múltiplas”.

⁸ Tradução nossa. No original: “As every mathematical concept involves essential variable, all these mathematical variables need to be varied if full generality of the mathematical concept is to be achieved.” (DIENES, 1963, p. 156).

números aleatórios. O cerne de seu trabalho está na aplicação do teorema que, sob o ponto de vista didático, é a primeira etapa para compreendê-lo. Por exemplo, situações-problema que podem ocorrer no dia a dia, constatados ao serem realizados pelos alunos, antes da formalização do conceito do teorema.

O segundo trabalho que encontramos sobre o tema foi publicado nos anais do ICOTS-7 por Alvorado e Batanero (2006). Os autores propõem analisar a atividade do professor quando está planejando a aula ao tratar-se do conceito do Teorema Central do Limite para engenheiros; abordar diferentes situações em que o teorema pode ser apresentado como proposta didática para seu ensino; apresentar sua abordagem histórica no sentido de favorecer a aprendizagem dos alunos; conhecer as formas de apresentação verbal, simbólica e gráfica do teorema a partir de análise de livros-texto de Estatística para o curso de Engenharia.

A metodologia dessa pesquisa deu-se a partir de uma proposta didática em que incorporou simulações com materiais manipuláveis e o uso do software *@Risk*, um programa da Microsoft Excel. E o trabalho, de um modo geral, propôs um método didático para melhorar o ensino de Estatística para engenheiros.

Para os autores, o planejamento do ensino está baseado no conhecimento matemático sobre o teorema. A partir do modelo teórico de Godino (2002) para análise de livros-texto usados no curso introdutório de Estatística para Engenharia, eles buscaram diagnosticar elementos que permitiram a abordagem do Teorema Central do Limite, tal como o conjunto de problemas que deram origem ao teorema; as representações verbal, simbólica e gráfica usadas pelos alunos na resolução desses problemas e nos diferentes objetos matemáticos envolvidos; os procedimentos usados por diferentes autores para resolver os problemas que envolvem o teorema; as propriedades dos objetos matemáticos usados nesses procedimentos; e, por fim, nos argumentos usados para justificar as soluções dos problemas. Foram analisados 16 livros-texto indicados para o curso de Engenharia, na Universidade de Granada.

Apesar da pesquisa ainda estar em vias de experimentação, os autores puderam constatar nesse artigo que, mesmo que o Teorema Central do Limite seja reconhecido como uma ferramenta fundamental para Inferência Estatística, as leituras prévias dos estudantes e o tempo disponível de ensino não são suficientes para prestar-lhe devida atenção. Em geral, há dificuldade de compreensão quando é apresentada uma demonstração formal de algumas formulações do teorema até

mesmo por alunos da pós-graduação em curso. E, além disso, verificaram que vale a pena dispor de algumas aulas para investigar o significado do teorema por meio de programas gráficos computacionais dinâmicos.

A investigação de Lundsford, Rowell e Goodson-Espy (2006) envolveu, como sujeitos de pesquisa dois grupos distintos, um composto por estudantes da disciplina de Introdução à Probabilidade e outro da Introdução à Estatística, com o objetivo de averiguar certos tipos de “erros” que podem ocorrer durante o aprendizado da distribuição de média amostral e da aplicação do Teorema Central do Limite.

Esse trabalho refere-se a uma pesquisa realizada em salas de aula com alunos, do curso de Matemática de pós-cálculo, de duas disciplinas: *Math300* Introdução à Probabilidade e *Math400*, Introdução à Estatística Matemática. A maioria dos alunos de *Math300* é da área das ciências da computação e de *Math400*, da área de Matemática (bacharelado ou licenciatura). Os autores analisaram o desenvolvimento de conceitos sobre distribuição de média de amostras e do Teorema Central do Limite, a partir dos mesmos instrumentos para os dois grupos, por simulações computacionais. Todos os alunos submeteram-se às fases de pré-teste e pós-teste, o que possibilitou o confronto entre as análises dessas fases.

Algumas conclusões a que os autores chegaram com relação às habilidades específicas de interpretação gráfica, comparação, raciocínio sobre distribuições amostrais, foram: as apresentações gráficas de conceitos pela simulação computacional não são suficientes para desenvolver as habilidades; os conhecimentos prévios sobre o centro, a forma e a dispersão das variáveis numa distribuição amostral não implica, necessariamente, que os alunos tenham habilidades de aplicá-los para resolver os problemas computacionais; houve fortes indícios, durante as atividades, de que a maioria dos estudantes não compreenderam completamente o Teorema Central do Limite, pois não reconheceram o quão rápido uma distribuição amostral torna-se unimodal à medida que aumenta o tamanho n de amostras; muitos alunos tiveram dificuldades para reconhecer variabilidade graficamente e computacionalmente, e isso pode ser atribuído ao fato de não serem capazes de identificar graficamente o efeito do desvio padrão sobre a forma da curva normal, ou de distinguir entre os conceitos de variabilidade e de frequência; conceitos como distribuições de probabilidade,

especialmente distribuições amostrais e o Teorema Central do Limite, foram difíceis de entendimento para todos os alunos.

Porém, em geral, os alunos da turma *Math300* apresentaram respostas positivas ao realizar as atividades e simulações, reconhecendo a contribuição do ambiente computacional para o processo de aprendizagem.

A partir dos procedimentos metodológicos da pesquisa-ação, os resultados observados proporcionaram, aos autores, sugestões de algumas ações que poderiam contribuir para aprendizagem desses temas. Eles propuseram explorar melhor as habilidades de raciocínio gráfico por meio de atividades e exercícios que permitiram aos alunos explorar e estimar parâmetros, como a forma, média e dispersão das distribuições em situações-problemas; explorar melhor a compreensão de variáveis aleatórias e suas distribuições pelo uso da simulação como uma ferramenta de atividades e exercícios durante todo o período do curso; e, particularmente, continuar a utilização das atividades de Distribuições Amostrais (*Sampling Distributions*) no sentido de clarear os conceitos entre o valor esperado e variância da distribuição amostral para médias amostrais no Teorema Central do Limite.

Por fim, Lundsford, Rowell e Goodson-Espy (2006) enfatizam a importância de promover pesquisas em sala de aula, no intuito de melhorar o ensino e também desenvolver técnicas de avaliação com o propósito de obter informações quantitativas no que concerne à compreensão de conceitos por parte dos alunos.

Além desses três artigos, existem outros trabalhos nos quais o teorema não é o objeto matemático principal, mas há consenso entre os autores de que o teorema é um dos mais importantes para Inferência Estatística. Como exemplo, o artigo de David (2003), cuja publicação foi significativa para os nossos estudos, pois sua abordagem também faz o uso da história da Estatística na sala de aula, como forma de motivar os alunos, na perspectiva da didática da matemática, particularmente, a educação estatística, mesmo o teorema não sendo abordado de forma explícita.

A tese de doutorado de Méndez (1991) estuda os sistemas de crenças dos alunos veteranos e principiantes do curso de Doutorado em Estatística e Economia, sobre os aspectos fundamentais do Teorema Central do Limite. Ele classificou os erros mais comuns e observou o alcance dessas crenças, por parte dos alunos novatos, refletidas pelas representações dos veteranos. Isso comprova a

complexidade na construção de significados para conceitos quando passa por crenças de uns para outros.

A partir da análise de dez livros-texto, Méndez (1991) constatou quatro premissas fundamentais para a compreensão do Teorema Central do Limite:

- i) A média da distribuição amostral é igual a média da população, na medida em que o tamanho da amostra aumenta, tendendo ao infinito.
- ii) A variância da distribuição amostral é menor que a da população.
- iii) A forma da distribuição amostral tende a ser *normal* na medida em que aumenta o tamanho da amostra, isto é, aproximadamente normal, independentemente da forma da distribuição da população.
- iv) A forma da distribuição amostral cresce em altura e decresce em dispersão conforme se aumenta o tamanho da amostra.

A partir das premissas acima, Méndez (1991) trabalhou com textos escritos pelos alunos veteranos, verificando seus conhecimentos sobre a representação de um conjunto de conhecimentos implícitos acerca do Teorema Central do Limite. E, por meio do mapa conceitual, definiu dois níveis de compreensão. No primeiro nível, reconheceu as habilidades e os conhecimentos exigidos para resolver as atividades propostas pelos livros-texto. O segundo nível diz respeito a outras áreas do conhecimento que, geralmente, não estão nos livros-texto.

Os dois níveis de compreensão, oriundos do resultado da análise de duas tarefas realizadas na primeira fase da pesquisa, suscitaram o raciocínio em relação aos conceitos que exigem diferentes níveis de interpretação. A primeira tarefa constituiu-se de um teste contendo questões clássicas de múltiplas escolhas extraídas de livros-texto, previamente analisadas pelo investigador; a segunda tarefa revelou várias situações diferentes, porque direcionava recolher as declarações e procedimentos que usaram nas terminologias específicas adquiridas por cada um dos sujeitos participantes, providos de suas crenças.

Para a segunda fase, realizaram entrevistas com os alunos que haviam participado da primeira, o que permitiu dividir em três grupos de sujeitos, os principiantes que tinham algum estudo prévio de Estatística; os que não tinham e estudantes de doutorado em Estatística e Economia. O objetivo geral da investigação foi observar como eles compreendem ou não o Teorema Central do

Limite e, em particular, observar suas compreensões a partir das quatro premissas anteriormente citadas.

Em síntese, Méndez (1991) concluiu que os estudantes de doutorado mostraram uma boa compreensão do teorema e de seus elementos implícitos, mas o discurso foi excessivamente formal. Já para os demais grupos, faltavam-lhes habilidade em expressar-se intuitivamente, pois suas explicações eram formais e, a partir disso, constatarem compreensão superficial sobre o teorema. Em todos os grupos, a maioria dos participantes usaram os dados sem considerar a população na qual foram extraídas as amostras e, também, não levaram em conta o tamanho das amostras.

Tomando como base essas conclusões, o autor recomenda que um curso introdutório de Estatística deve levar em conta a natureza dos recursos de aprendizagem para conceitos e procedimentos que se deseja ensinar. Méndez (1991) ainda recomenda o uso de dados que propiciam aos alunos observar os aspectos principais do Teorema Central do Limite; a utilização de materiais concretos, como, por exemplo, simulações por meio de lançamento de dados para criar representações nos processos de extração das amostrais e a distribuição amostral das médias, bem como organizar simulações quando o tamanho da amostra for suficientemente grande. Destaca, ainda, a importância de que o professor seja consciente dos diversos níveis de compreensão que se distinguem entre os alunos, propondo que a compreensão intuitiva deve anteceder o pensamento formalizado do teorema. Por fim, faz uma crítica aos livros-texto analisados, pois lhe foi possível constatar uma supervalorização quanto aos aspectos quantitativos em detrimento dos qualitativos. Consequentemente, os alunos acabam por usar linguagem formal sem compreender efetivamente o conceito do teorema.

Um outro trabalho que cita a importância do Teorema Central do Limite, embora não seja o tema central da pesquisa, é a tese de Tauber (2001), porque ela dá ênfase ao teorema a partir da importância da distribuição normal na Estatística:

A distribuição normal desempenha um papel muito particular no estudo da inferência, porque **o teorema central do limite** permite a aplicação dos métodos de inferência, sem preocupar-se pela forma concreta da distribuição da população, sempre que o tamanho da amostra for suficientemente grande. (TAUBER, 2001, p. 14) (Grifo nosso)

Os resultados obtidos nessas leituras nos permitiram reforçar a problemática do nosso trabalho. Contudo, parece não haver ainda uma abordagem do teorema sob a perspectiva ecológica na Didática da Matemática, nos dando, assim, o poder de atribuir caráter inédito para a nossa pesquisa.

3.2 Um Breve Cenário Histórico do Teorema Central do Limite

Cajori (2007), Struik (1992), Eves (1997), Rodriguez (1989), Mether (2003) e Fischer (2000) são alguns historiadores que fizeram do século XVIII a era da Matemática Ilustrada, devido à influência do modelo vigente: o Iluminismo. Nesse contexto, deu-se o desenvolvimento das ideias que fundamentaram a formulação do Teorema Central do Limite.

A evolução desse teorema está estreitamente relacionada ao desenvolvimento das ideias da probabilidade, como observa Coutinho (1996):

A noção de acaso data da História Antiga, tendo sua origem ligada aos jogos de azar, notadamente na civilização egípcia, primeira dinastia, 3500 a.C., certamente com um aspecto lúdico. O desenvolvimento, porém, das idéias que formam a base do desenvolvimento da probabilidade ocorreu bem mais tarde, com Jérôme Cardan (*De Ludo Aleae*), Galileu (*Sulla Scoperta dei Dadi*) e Fra Luca dal Borgo, que em sua obra publicada em 1494 e intitulada *Summa de Arithmetica, Geometria, Proportioni et Proportionalita*, enuncia o problema mais tarde resolvido por Blaise Pascal (1623 – 1662) e Pierre de Fermat (1601 – 1665), a quem podemos, de certa forma, atribuir a origem da concepção de Probabilidade. (COUTINHO, 1996, p. 12)

Dessa forma, podemos destacar Jacob Bernoulli (1654-1705) que iniciou o enfoque frequentista de probabilidade, isto é, aproximou a probabilidade de um evento observando a frequência de ocorrência quando um experimento é repetido um grande número de vezes. Mas foi Laplace (1749-1827) que, em suas investigações sobre Astronomia, “[...] provou que as velocidades médias ou as distâncias médias dos planetas são invariáveis, ou meramente sujeitas a pequenas mudanças periódicas” (CAJORI, 2007, p. 347). A partir daí, ele passou a interpretar as pequenas mudanças periódicas como pequenos “erros” que deram origem às distribuições de probabilidades de variáveis aleatórias independentes.

O método desenvolvido por Laplace, conforme se encontra no apêndice A, deu origem às primeiras ideias acerca do Teorema Central do Limite e esse feito o levou a ser considerado o seu “criador”. Um outro trabalho importante realizado por ele, principalmente para Estatística, foi a obra *l'Essai Philosophique sur les Probabilités*, publicada em 1814, em que fez a primeira tentativa de axiomatização da Teoria das Probabilidades e, ainda, explorou a natureza da teoria probabilística do acaso (COUTINHO, 2001). Nesse mesmo trabalho, Laplace usou a propriedade relativa à soma de variáveis aleatórias como sendo base para a demonstração⁹ do Teorema Central do Limite e, além disso, trabalhou com a função característica, conforme apêndice B, numa distribuição de probabilidades. Apesar do caráter dedutivo, Laplace não tinha intenção de estudar o teorema em si, mas utilizá-lo apenas como técnica para a teoria das probabilidades. Apesar do caráter dedutivo, Laplace não tinha intenção de estudar o teorema em si, mas utilizá-lo apenas como técnica, ou seja, um procedimento para o cálculo de probabilidades e consolidar sua teoria.

Nos estudos de Blaiotta e Delieutraz (2004, p. 2), é atribuída a Polya a utilização do termo “central” nesse teorema, pois significa ‘de importância central’ na teoria das probabilidades, tendo em vista que o Teorema Central do Limite é, muitas vezes, implicitamente, aplicado em situações reais. Por exemplo, como bem descreveu Laplace, ao considerar o erro total como a soma de numerosos erros essencialmente muito pequenos, devido a causas independentes, essas distribuições de erros são normais. Isso ocorre em muitas situações realizadas, como na Astronomia, nas pequenas variações de temperatura, nas correntes irregulares de ar, entre outros.

Fischer (2000) afirma que, de todas as contribuições para a compreensão da demonstração do teorema, no século XIX, a de Poisson (1781-1840) pode ser considerada aquela que mais desenvolveu conceitos que cercam a formulação do teorema. Poisson publicou dois artigos, um em 1824 e outro em 1829, levantando questões para serem discutidas sobre o teorema. Sua ideia era a de que todo procedimento do mundo físico fosse governado por leis matemáticas distintas. Nesse contexto, ele tentou apresentar uma análise matemática um pouco mais

⁹ Segundo Balacheff (2004), a demonstração é a pedra angular do pensamento matemático, do raciocínio dedutivo, o qual tem sua base teórica no processo de provar por meio de códigos e de formalidades.

detalhada do que aquela apresentada por Laplace. Ele considerou dois aspectos: apresentou uma demonstração para uma variável contínua, a partir das ideias iniciais sobre variáveis aleatórias; e discutiu a validade do Teorema Central do Limite.

Esse tratamento inicial, dado por Poisson ao teorema, fundamenta-se a partir da condição de que as variáveis devem ser identicamente distribuídas: primeiro pela soma delas e, em seguida, por uma combinação linear de seus elementos. Dessa forma, generaliza a demonstração da soma de variáveis aleatórias para diferentes distribuições, conforme se pode comprovar no apêndice D.

As transições entre os séculos XVIII e XIX, e os séculos XIX e XX, foram marcadas pelos esforços e dedicação por parte dos matemáticos no sentido de construir fundamentação lógico-dedutiva para as questões “abertas” que, inclusive, ainda existem na atualidade, como, por exemplo, a *Hipótese de Riemann* que, segundo Devlin (2004, p. 15), “é o único problema de Hilbert de 1900 que continua sem solução”. Segundo Boyer (1974):

No Congresso de Paris de 1900, Hilbert, renomado professor em Göttingen, apresentou uma exposição em que tentou, com base nas tendências da pesquisa matemática no fim do glorioso século dezanove, predizer a direção de progressos futuros. Isso ele fez propondo vinte e três problemas que ele acreditava estariam ou deveriam estar entre os que ocupariam a atenção dos matemáticos no século vinte. (BOYER, 1974, p. 443)

Esses problemas propostos por Hilbert contribuíram para que o rigor matemático ocupasse o centro da atenção daqueles que estavam no meio acadêmico. Para Eves (1997, p. 463), foi nesse contexto que surgiu a distinção entre matemática “pura” e “aplicada”. A primeira foi destinada aos especialistas cujos interesses estavam voltados para os objetos matemáticos em si; já a segunda priorizou o estudo de suas aplicações. Esse argumento é questionado por Bruter (1998) quando diz:

[...] no plano do conhecimento matemático, o matemático aplicado nem sempre produz resultados significativos, o mesmo acontecendo, é claro, com o matemático puro, quer por não ter tido ainda a sorte de ser tocado pela graça, quer porque lhe falta a prática profunda da disciplina na qual opera o seu modelo, e que lhe permitiria vislumbrar propriedades interessantes, originais: tendo adivinhado a sua presença, fá-las-ia surgir do modelo, descobrindo assim talvez propriedades matemática novas. (BRUTER, 1998, p. 20)

No entanto, a teoria das probabilidades, no século XIX, foi considerada mais como ‘senso comum’ do que uma teoria advinda da Matemática. Não tardaria para que alguns matemáticos investissem em demonstrar o teorema, conforme era exigido pela comunidade, tais como Bessel (1784-1846), Dirichlet (1805-1859), Cauchy e Ellis (1814-1890). Esses estudiosos trabalharam, por meio de várias tentativas, a primeira versão do Teorema Central do Limite que foi formulada por Laplace (*vide* apêndice E).

No entanto, até então, segundo Hald (1998, p. 402), as tentativas não foram satisfatórias sob três aspectos: (i) o teorema não havia sido mostrado para distribuições infinitas; (ii) não havia condições explícitas, em termos de *momentos* (apêndice C), sobre as quais o teorema é consolidado; (iii) não havia conhecimento na época sobre a razão de convergência.

Nas palavras de Lakatos (1978, p. 177), nesse período, “eles não sabiam que após o descobrimento de um contra-exemplo eles não tinham que analisar sua prova cuidadosamente e tentar encontrar o lema oculto.”

Outros grandes personagens da história da Matemática também contribuíram diretamente para o desenvolvimento do teorema e, dentre eles, podemos destacar a família Bernoulli, Bayes (1702-1761), DeMoivre (1667-1754) e Taylor (1685-1731)¹⁰.

Mas os três itens apontados por Hald (1998) foram resolvidos por matemáticos russos, entre 1870 e 1910. Destacam-se Chebyshev (1821-1894), Markov (1856-1922) e Liapounov (1857-1918).

A publicação do artigo de Chebyshev é, tradicionalmente, considerada o início das demonstrações com rigor para o teorema. Apesar de a demonstração estar incompleta, Chebyshev usou o ‘método dos momentos’, conforme apêndice C. Mais tarde, seu trabalho recebeu a contribuição de Markov, que também trabalhou arduamente para obter a generalização do método de momentos após a prova de Liapounov. Ele finalmente foi bem sucedido em 1913, quando apresentou um artigo que continha uma demonstração do Teorema Central do Limite. Mais detalhes encontram-se no apêndice F.

Liapounov e Markov foram alunos de Chebyshev. Liapounov queria introduzir provas rigorosas para teoria das probabilidades e foi bem sucedido em seu intento. Ele não recorreu ao ‘método de momentos’, mas seguiu a idéia de Laplace, fazendo

¹⁰ Não temos a intenção de elencar todos os matemáticos que contribuíram para a História da Estatística, mas apenas citar alguns.

o uso das funções características. Liapounov publicou, em 1901, uma demonstração que é considerada a “primeira” constituída de rigor para o teorema, porém ainda incompleta.

De acordo com Methner (2003), Lindeberg (1876-1932) chega a iniciar a demonstração, finalizando com as condições *necessária* e *suficiente* de Lévy (1838-1910) e Feller (1906-1970) para o Teorema Central do Limite. Contudo, o teorema recebe uma demonstração elementar por meio da publicação de Lindeberg em 1922, cujo argumento era simples e aplicável para os valores do plano euclidiano (CAM, 1986), pois considerava as variáveis x_i , como variáveis aleatórias independentes com expectância zero e variância σ_i^2 igual a um, isto é, segundo Gnedenko e Kolmogorov (1954, p. 90), “seja s_n o desvio padrão da soma S ,

$s_n^2 = \sum \sigma_i^2$. Se $\frac{1}{s_n^2} \sum E \left\{ x_i^2 I_{\left[\frac{x_i}{s_n} > \varepsilon \right]} \right\} \rightarrow 0$, então, $\frac{S}{s_n} \rightarrow N(0,1)$ ”, em geral, conhecida

como a “condição de Lindeberg”. Entretanto, essa condição permaneceu inalterada até a primeira metade do século XX, isto é, até o período anterior à publicação de Trotter sobre o teorema, em 1959. Para Paulauskas (2006), o trabalho de Trotter apresentava duas vantagens: foi escrito em inglês e a demonstração apresentada foi muito clara. Isso chamou atenção dos pesquisadores — o que fez ressurgir a condição de Lindeberg —, cujos interesses estavam voltados para os teoremas sobre limites nos espaços de dimensões infinitas.

Atualmente a condição de Lindeberg ainda é usada na maioria dos casos de convergência para uma distribuição normal e, também, para variáveis aleatórias que não são distribuídas identicamente. As variáveis aleatórias identicamente distribuídas também são conhecidas como uniformemente distribuídas que, segundo Meyer (1983):

- (a) Uma variável aleatória uniformemente distribuída tem uma função densidade de probabilidade que é *constante* sobre o intervalo de confiança. A fim de satisfazer à condição. $\int_{-\infty}^{+\infty} f(x)dx = 1$, essa constante deve ser igual ao inverso do comprimento do intervalo.
- (b) Uma variável aleatória uniformemente distribuída representa o análogo contínuo dos resultados igualmente prováveis, no seguinte sentido: Para qualquer subintervalo $[c, d]$, onde $a \leq c \leq d \leq b$, $P(c \leq X \leq d)$ é a mesma para todos os subintervalos que tenham o mesmo comprimento [...], isto é, depende unicamente do comprimento do intervalo e não da posição desse intervalo.
- (c) Agora podemos tornar mais precisa a noção intuitiva de *escolher ao acaso um ponto P*, em um intervalo $[a, b]$. Por isto simplesmente

queremos dizer que a coordenada x do ponto escolhido, digamos X , é uniformemente distribuída sobre $[a, b]$. (MEYER, 1983, p. 90)

A condição de Lindeberg fundamenta-se em dois pontos: (i) pode ser aplicada, em geral, em diversos contextos; (ii) considera a razão de convergência como o menor valor.

As evidências confirmaram que sua demonstração foi provida de rigor matemático, entretanto, sob condições suficientes. Faltavam-lhe ainda as condições necessárias para consolidar a demonstração do Teorema Central do Limite. Como Poisson mostrou em 1824, a aproximação de uma distribuição assimétrica para normal nem sempre é de variáveis independentes. Essa lacuna foi parcialmente preenchida por Lévy e Feller em 1935 e 1937, respectivamente.

Lévy demonstrou a condição de Lindeberg em 1925, aplicando funções características. Para Cam (1986), ele transformou o trabalho de Lindeberg em simples e sofisticado para a época. Entre 1925 e 1930, publicou vários artigos sobre o teorema, dando ênfase às funções características em suas demonstrações. Porém, após 1930, ele evitou o uso dessas funções e, no seu artigo de 1935, não fez nenhum uso das funções características no todo. Esse mesmo trabalho de Lévy foi apresentado somente poucos meses depois de Feller, e, apesar de terem tratado da mesma questão em ambos os artigos, eles negaram qualquer contato anterior sobre o assunto. Lévy provou vários itens relacionados ao Teorema Central do Limite: (i) anunciou as condições *necessária* e *suficiente* para convergência de somas normalizadas de variáveis aleatórias independentes e identicamente distribuídas para uma distribuição normal; (ii) apresentou condições *necessária* e *suficiente* para o caso geral de somatórios independentes; (iii) tentou dar condições *necessária* e *suficiente* para variáveis dependentes.

Havia, no entanto, alguns problemas com as demonstrações de Lévy. As condições *necessária* e *suficiente* para o caso de distribuições marginais de probabilidade não eram satisfatórias o bastante e não se fez um teste padrão com rigor. É chamada de distribuição de probabilidade marginal para o caso de variáveis discretas e, se contínuo, é chamada de função densidade de probabilidade marginal. Para Meyer (1983, p. 116), “a cada variável aleatória bidimensional (X, Y) associa-se a duas variáveis aleatórias unidimensionais, a saber, X e Y , individualmente. Isto é, pode-se estar interessado na distribuição de probabilidade de X ou na distribuição de probabilidade de Y ”.

Lévy provou as condições necessária e suficiente para casos gerais de variáveis aleatórias independentes corretamente, mas sua demonstração dependia do lema fundamental, que ainda não havia sido provado. E esse lema é, segundo Mether (2003, p. 21), “Se a soma $S = X + Y$ de duas variáveis aleatórias independentes (X e Y) tem uma distribuição normal, então X e Y são distribuições normais”.

Toda prova de Lévy dependia desse lema, no entanto não foi satisfatória na época em que teve a oportunidade de ser apresentada. No ano seguinte, 1936, Cramér provou o lema (como um teorema). Com ajuda desse mesmo teorema, a utilização de somas normalizadas poderia ser apresentada para validar os teoremas de Lévy e Feller, pois, assim, seriam aplicáveis para casos gerais. Tanto Feller quanto Lévy retornaram e aperfeiçoaram seus trabalhos em 1937, após os resultados apresentados por Cramér. Cam (1986, p. 90) afirma que “o Teorema Central do Limite foi então provado nas condições necessária e suficiente”. Assim, foi satisfeita a demonstração do teorema tal qual se conhece nos dias atuais (*vide* apêndice G).

Em suma, a história do Teorema Central do Limite iniciou a partir da discussão sobre a necessidade de aproximação das somas dos erros de uma distribuição, sendo Laplace o precursor desse processo, no século XVIII, que só se concluiu em meados do século XX, ou seja, precisou-se de mais de 200 anos para a realização da demonstração do teorema. Esse longo percurso mostra a complexidade do tema e a necessidade de desenvolvimento de um conhecimento matemático preciso para sua correta e completa demonstração.

Portanto, conforme exposto, na primeira parte nos foi possível constatar que os estudos que têm como foco o Teorema Central do Limite ainda são bastante escassos, principalmente quando comparados com outros temas da Educação Estatística, o que confere relevância de nossa investigação.

Já para a segunda parte da revisão da literatura, uma abordagem histórica, mesmo que superficial, nos reporta a uma visão de como o teorema surgiu, os obstáculos pelos quais os matemáticos se depararam e, por fim, a justificativa sobre a importância dele como objeto de estudo de nossa pesquisa. Além disso, no contexto do processo de ensino e de aprendizagem, Kline (1976) confirma que:

[...] Podemos dar aos estudantes as abordagens certas, e eles as compreenderão. Pode-se contestar esse argumento dizendo que os maiores

matemáticos procuraram realmente construir os fundamentos lógicos para as várias questões, mas malograram durante séculos. O malogro deles deve servir como prova de que as abordagens lógicas não são fáceis de apreender. Pode-se resumir a história e evitar muitos dos esforços desperdiçados e armadilhas, mas não se pode eliminá-las. (KLINE, 1976, p. 60)

E, assim, justificamos nossa opção por não ter detalhado a demonstração do Teorema Central do Limite, uma vez que a delimitação do nosso tema se refere à abordagem didática. Porém, os diversos passos para a construção dessa demonstração podem ser encontrados nos apêndices de A a G.

4 PROCEDIMENTOS E REFERENCIAIS METODOLÓGICOS

A partir da revisão da literatura, averiguamos algumas lacunas existentes sobre o Teorema Central do Limite quanto à abordagem que lhe é dada, atualmente, nos cursos de Licenciatura. Podemos apontar como um dos motivos a falta de ênfase, ao menos nos livros-texto, quanto à demonstração do teorema, assim como a ausência de citações a uma abordagem histórica sobre o seu surgimento, mesmo que os autores afirmem sua importância ao tratarem de Inferência Estatística. Além disso, as aplicações são diretas, em situações-problema que nem sempre estão direcionadas para o teorema em si, mas, sim, subjacente aos problemas.

Para realizar a nossa investigação, optamos pela abordagem de Artaud (1998) quanto à problemática ecológica do objeto, o que nos permitiu questionar o real, isto é, “o que existe? Por que existe?” Ou então, o que não existe? Poderia existir? Sobre quais condições?, que, em síntese, no contexto do Teorema Central do Limite, constitui-se a problemática de nossa investigação:

Por que determinados saberes/conhecimentos deixaram de ser abordados no ensino do Teorema Central do Limite, mas ainda assim ele “sobrevive”? Que saberes são indispensáveis para o teorema “viver”? Que relações diretas ou indiretas existem entre o teorema e a construção de saberes estatísticos nas aulas de Estatística?

O aprofundamento sobre a relevância ecológica atribuída à problemática desse estudo será abordada no capítulo relativo ao quadro teórico. Por enquanto, nos restringimos a expor os procedimentos metodológicos que nortearam a nossa investigação.

Para Artigue (1995), a engenharia didática é um método de investigação que se caracteriza, fundamentalmente, por um esquema experimental, cuja base está nas realizações didáticas que ocorrem na sala de aula, envolvendo, assim, da concepção à análise de sequências didáticas. Mas, com as avaliações que ocorrem internacionalmente com alunos, como, por exemplo, o programa PISA (*Programme for International Students Assessment*), sob responsabilidade da OECD (*Organisations for Economic Co-operation and Development*), a tendência das

pesquisas didáticas nos reporta a seguir o desenvolvimento da própria Matemática na nossa sociedade.

Pela “medida” do letramento matemático dos estudantes e das competências que devem proporcionar, estas avaliações questionam de fora a eficiência dos nossos sistemas educacionais tornando mais visível no cenário internacional os sucessos, mas também as limitações e falhas, e conduz para a questão de pesquisa educacional sobre seu potencial para informar e guiar decisões e normas curriculares.¹¹(ARTIGUE, 2009, p.1-2)

Assim, podemos dizer que o *design didático* nos permite reconhecer as intervenções controladas, exercendo, segundo a mesma autora, um papel essencial no sentido de como a teoria e a prática podem ser efetivamente úteis. O termo “letramento foi escolhido para refletir a amplitude dos conhecimentos, habilidades e competências que estão sendo avaliados” (BRASIL, 2009). E o letramento em Matemática é avaliado sob três dimensões: conceitual, procedimental e interdisciplinar.

Sob a dimensão conceitual, o letramento matemático recebe uma abordagem mais ampla, destacando-se, em primeiro lugar, os conceitos relativos à estimativa, mudança e crescimento; espaço e forma; raciocínio lógico; incerteza e dependências e relações; em segundo, conceitos matemáticos contidos no currículo, como, por exemplo, as relações numéricas, álgebra e geometria.

A dimensão procedimental, em linhas gerais, refere-se às competências e às habilidades que o estudante deve desenvolver de modo que saiba resolver problemas a partir da escolha de estratégias, que são divididas em três classes: a primeira diz respeito às operações simples; a segunda, ao reconhecimento em estabelecer as interrelações entre os diferentes modelos matemáticos para resolução de problemas; e, por fim, a terceira classe dispõe da prática do raciocínio matemático em estabelecer generalizações e exercer a descoberta.

A terceira dimensão do letramento matemático, a interdisciplinar, consiste na resolução de problemas em todas as vertentes da ciência ou do cotidiano, entendendo que a Matemática não é uma disciplina isolada das demais, isto é, em oposição à visão fragmentada do conhecimento.

¹¹ Texto original: “Through the “measure” of students’ mathematical literacy and competences that they pretend to provide, these evaluations question from the outside the efficiency of our educational systems, make more visible on the international scene successes but also limitations and failures, and lead to question educational research about its potential for informing and guiding curricular decisions and policies.” (Trad. COUTINHO, Cileda de Queiroz e Silva; SILVA, Maria José Ferreira da)

A pertinência do letramento matemático na nossa investigação é tocante a dois aspectos, a saber: aos licenciandos em Matemática e à *literacia* estatística. Em decorrência disso, optamos pela metodologia do *design didático*, pois Artigue (2009) acredita que a associação entre essa metodologia, a Teoria das Situações Didática (BROUSSEAU, 1996) e a Teoria Antropológica do Didático (CHEVALLARD, 1996, 1999) será extremamente apropriada, pois pode dar suporte reflexiva ao *design didático*

Com relação à TSD (Teoria das Situações Didáticas), a engenharia didática tem seus termos bem definidos quanto à intervenção controlada, pois é baseada em sua teoria e, para Artigue (2009), vale apontar algumas características da TSD, pois afeta o ponto de vista do *design*. Em primeiro lugar, o objeto da TSD é a situação em si e, quanto à aprendizagem, depende da abordagem em que se dão as situações que ocorrem na sala de aula; em segundo, há ênfase na epistemologia do conhecimento, o que significa uma análise mais específica das situações, tendo em vista que o aspecto positivo é o de despertar o conhecimento matemático visado; terceiro, a influência do *milleu* na situação e na interação dos estudantes como forma de “assegurar uma adaptação adidática produtiva” (ARTIGUE, 2009, p. 5); em quarto, o conhecimento matemático se distingue na *ação*, na *formulação* e na *validação*; e, em quinto lugar, o papel do professor como ator, desde a organização das relações didáticas e adidáticas até o encaminhamento do processo de *devolução* e *institucionalização*. Para Brousseau (1996), é de responsabilidade do professor porque:

O matemático não comunica seus resultados tal como os obteve, mas os reorganiza, lhes dá forma mais geral possível; realiza uma “didática prática” que consiste em dar ao saber uma forma comunicável, descontextualizada, despersonalizada, fora de um contexto temporal.

O professor realiza primeiro o trabalho inverso ao do cientista, uma recontextualização do saber: procura situações que dêem sentido aos conhecimentos que devem ser ensinados. Porém, se a fase de personificação funcionou bem, quando o aluno respondeu às situações propostas não sabia que o que “produziu”, é um conhecimento que poderá utilizar em outras ocasiões. Para transformar suas respostas e seus conhecimentos em saber deverá, com a ajuda do professor, re-despersonalizar e re-descontextualizar o saber que produziu, para poder reconhecer no que fez algo que tenha caráter universal, um conhecimento cultural reutilizável. (BROUSSEAU, 1996, p. 48)

No entanto, Artigue (2009) afirma que há limitação no papel do professor, mesmo durante sua participação no processo de devolução e institucionalização,

considerando-se que, se essa estrutura pode ser associada ao *design* e à engenharia didática, pode enriquecer as sucessivas fases de análise *a priori* e *a posteriori*. A partir do *design didático*, pode-se garantir a internalização do processo em estudo, sem excluir as ferramentas metodológicas.

Um outro aspecto a ser considerado é o da Teoria Antropológica do Didático e o *design didático*. Segundo Artigue (2009, p. 8), “a escala de níveis de co-determinação didática certamente pode ajudar a fazer um design didático sensível aos diferentes tipos de limitações a que as ações didáticas são submetidas, do nível da civilização ao nível do sujeito”¹². Assim, como um dos procedimentos metodológicos de nossa investigação foi a análise das atividades que aplicam o Teorema Central do Limite em alguns livros-texto do Ensino Superior, conciliar a TAD e o *design didático* nos permitiu entrelaçar a pesquisa e a prática. O fator positivo do *design didático* é que pode tornar as questões ecológicas visíveis e possíveis de serem organizadas para seu estudo.

Os procedimentos metodológicos foram divididos em três partes. A primeira é a análise dos livros-texto à luz da Teoria Antropológica do Didático. A segunda parte diz respeito às atividades realizadas com os quatro licenciandos, seguindo um estudo qualitativo sobre como eles construíram as ideias em torno do teorema de modo que conseguissem atingir a consolidação do saber enquanto objeto de estudo. Para tal, foi apresentada uma série de atividades, cada qual constituindo-se numa sequência didática de forma a dar contorno ao ecossistema do saber e do didático, conforme será abordado no quadro teórico de nosso trabalho.

Na terceira parte, optamos por interagir com os alunos a partir de um diálogo sobre as atividades realizadas, cujo instrumento utilizado foi gravação oral e, posteriormente, transcritas. Nesse momento, a pesquisadora atuou como professora, recorrendo, assim, a alguns fundamentos da pesquisa e da ação, seguindo a proposta da intervenção controlada ou não, no contexto do *design didático* de Artigue (2009). Isso confirmou uma investigação qualitativa que, durante os diálogos, nos permitiu averiguar a *literacia* (ou letramento em) estatística. Vejamos cada fase dos procedimentos metodológicos que nortearam a pesquisa.

¹² Texto original: “The scale of levels of didactic codetermination certainly can help to make didactical design sensitive to the different kinds of constraints to which didactical action is submitted, from the civilization level to the subject level.” (Trad. COUTINHO, Cileda de Queiroz e Silva; SILVA, Maria José Ferreira da)

4.1 A Análise dos Livros-Texto e a Teoria Antropológica do Didático

Nos estudos de Almouloud (2007, p. 113), “a problemática antropológica amplia o campo de análise e permite abordar os problemas que se criam entre os diferentes objetos do saber a ensinar.” Nessa perspectiva, tentaremos buscar uma estrutura hierárquica, estabelecendo as interrelações entre os objetos, do ponto de vista antropológico, de modo que nos permita identificar e analisar as estruturas ecológicas dos objetos relacionados com o Teorema Central do Limite. E, para tal, iniciamos com as justificativas que nos levaram a escolher determinados livros-texto, para, em seguida, apresentar uma atividade de cada um deles e, finalmente, expor como o teorema é definido por eles.

Entendemos que uma das atividades do professor é a de elaborar suas aulas a partir de bibliografias que o levam a refletir sobre as etapas didáticas com as quais irá se deparar. Portanto, a nossa escolha pelos livros-texto deve-se a alguns critérios:

- Levar em consideração a experiência dos autores no ramo da Estatística.
- Pertencer, seja como bibliografia básica, seja como complementar, a algumas ementas de cursos de Licenciatura em Matemática.
- Ter edições mais recentes.

A partir da Teoria da Transposição Didática (CHEVALLARD, 1985) e suas vertentes, faremos uma abordagem antropológica para as atividades realizadas, que servirão como modelo para análise dos livros-texto e do desenvolvimento das atividades, reforçando a questão ecológica.

Retomando o que Chevallard (1996) afirma, a teoria antropológica do conhecimento ou antropologia cognitiva provém das primeiras formulações sobre a teoria da transposição didática, a partir de três elementos: o objeto, as pessoas e as instituições. O autor alega que o objeto de estudo tem seu destaque no sentido de que ocupa uma posição privilegiada, e isso lhe é atribuído por ser um material de base para investigação. Assim, a praxeologia (ou organização) da ecologia do objeto matemático de estudo, o Teorema Central do Limite, é de duas espécies: a

matemática e a didática, o que reforça a estrutura do primeiro procedimento metodológico.

Segundo Chevallard (1996), a teoria (Θ) é o nível superior de justificativa-explicação-produção e nem sempre está presente numa atividade. No entanto, podemos dizer que essa atividade, por si, encarrega-se por pertencer a uma organização praxeológica pontual e global, ou seja, quando a praxeologia relaciona uma única atividade T ao complexo: tarefa, técnica, tecnologia e teoria (T, τ, θ, Θ), refere-se ao pontual. E, por outro lado, amplia a teoria para outras organizações praxeológicas regionais, cujo conjunto se constitui no complexo global.

A técnica não se restringe apenas a um modo de ser realizado. Em contrapartida, alguns conceitos devem ser do conhecimento do estudante para que se possa aplicar a técnica. A tecnologia (θ) existe para justificar e explicar as técnicas (τ), e a teoria (Θ) é a justificativa da justificativa, isto é, a tecnologia da tecnologia.

O desenvolvimento das praxeologias proporcionou o *design didático* numa forma própria e única para cada etapa, isto é, sucessivamente, pois nos permitiu a análise pontual para a local e a regional.

4.2 As Atividades Didáticas e os Licenciandos

A segunda parte dos procedimentos metodológicos constitui-se de uma série de atividades que permitiram aos quatro licenciandos construir a definição do Teorema Central do Limite, para, conseqüentemente, poderem atingir à compreensão do conhecimento. Para elaboração das atividades, utilizamos um exemplo, cujos dados coletados foram realizados por uma pesquisa¹³ que mensurou o índice de autoestima de alunos do Ensino Fundamental, particularmente do 6º ano.

¹³ A pesquisa foi de Iniciação Científica e contou com a participação dos alunos da Licenciatura em Matemática, Andréa Stambassi Souza e Vanderson Damasceno Ribeiro, sob a orientação da autora deste trabalho. Como os dados coletados não foram analisados estatisticamente, pareceu oportuno lhes dar um tratamento estatístico apropriado. Isso fez com que contemplassem as duas dimensões desta pesquisa: o **Teorema Central do Limite** e outros objetos estatísticos numa **situação-problema** em que representa um tema pertinente e real para alunos da Licenciatura em Matemática.

Além de desenvolver, analisar e avaliar uma proposta de ensino centrada nos principais conceitos ligados ao teorema, buscamos identificar e compreender argumentos e procedimentos utilizados por alunos do curso de Licenciatura em Matemática. Essa situação pode ser evidenciada na descrição de quando eles trabalharam num cenário de aprendizagem privilegiado por situações-problema, em que os objetos estatísticos puderam surgir a partir de ideias advindas das concepções espontâneas, seguindo definição de Artigue (1990). Isso quer dizer que a intenção foi a de fazer com que diversas ideias de objetos matemáticos e/ou estatísticos pudessem emergir antes que sua definição fosse oficialmente apresentada.

Para que isso ocorresse, caracterizamos o nosso trabalho, propondo uma investigação qualitativa, e contamos, inicialmente, com apenas quatro licenciandos do Curso de Licenciatura em Matemática, de um Centro de Ensino da cidade de Juiz de Fora, Minas Gerais. Os participantes da pesquisa já se encontravam no último ano do curso e, diante disso, a nossa investigação ocorreu durante as aulas de Estatística, dirigida especialmente para eles, o que nos proporcionou observações mais pontuais durante as discussões sobre essa disciplina. Assim, pudemos exercer a pesquisa e a prática conforme a proposta do *design didático*. Além disso, nos foi possível também extrair alguns fundamentos teóricos da metodologia da pesquisa-ação, cujos atores são todos os que participaram da pesquisa. Segundo Thiollent (1986),

[...] a pesquisa-ação pode ser vista como modo de conceber e de organizar uma pesquisa social de finalidade prática e que esteja de acordo com as exigências próprias da ação e da participação dos atores da situação observada. Neste processo, a metodologia desempenha um papel de “bússola” na atividade dos pesquisadores, esclarecendo cada uma das suas decisões por meio de alguns princípios de cientificidade. (THIOLLENT, 1986, p. 26)

No desenvolvimento dessa metodologia, os pesquisadores recorrem a métodos e técnicas de grupos para lidar com a dimensão coletiva e interativa da investigação, bem como técnicas de registro, de processamento e de exposição de resultados. Por isso, um outro instrumento também utilizado foi o de discussão sobre as atividades realizadas e devidamente registradas por meio de gravação de vozes, que representará a terceira parte dos procedimentos metodológicos. Assim, garantimos o anonimato, sem expor os licenciandos em nenhuma situação de

constrangimento pessoal. Por meio de um termo de compromisso¹⁴, os licenciandos aceitaram participar da pesquisa voluntariamente, e o seu desligamento independeu da época em que se encontrava a investigação, seguindo sempre o desejo de participação de cada um.

Segundo Thiollent (1986), os objetivos de conhecimento potencialmente alcançáveis em pesquisa-ação são:

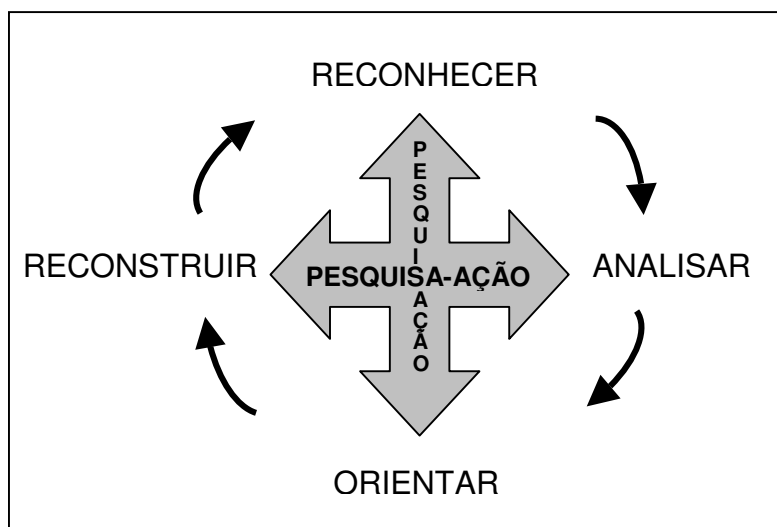
- a) A coleta de informação original acerca de situações de atores em movimento.
- b) A concretização de conhecimentos teóricos, obtida de modo dialogado na relação entre pesquisadores e membros representativos das situações ou problemas investigados.
- c) A comparação das representações próprias aos vários interlocutores, com aspecto de cotejo entre saber formal e informal acerca da resolução das categorias de problemas.
- d) A produção de guias ou de regras práticas para resolver os problemas e planejar as correspondentes ações.
- e) Os ensinamentos positivos ou negativos quanto à conduta da ação e condições de êxito.
- f) Possíveis generalizações estabelecidas a partir de várias pesquisas semelhantes e com o aprimoramento da experiência dos pesquisadores.

(THIOLLENT, 1986, p. 41)

A partir do *design didático*, cuja proposta incide sobre as relações produtivas entre pesquisa e prática, podemos sintetizar os seis itens expostos por Thiollent (1996) nas ações como: reconhecer o objeto de estudo mediante os participantes (a) e (b); analisar os impactos oriundos das respostas intuitivas e formais (c); orientar como forma de devolução para reflexão e discussão dos resultados de êxito ou “fracasso”¹⁵ (d) e (e); e, por fim, reconstruir, a partir dos aspectos de êxitos e “fracassos” (f). Entendemos que essa dinâmica representa um ciclo, nem sempre hierárquico, mas que, didaticamente, parece ser, do ponto de vista de interação, o pesquisador com os pesquisandos. Assim podemos sintetizar esses procedimentos conforme o esquema 1:

¹⁴ Anexo 8.

¹⁵ A palavra “fracasso” não tem conotação negativa, por isso está colocada entre aspas, pois entendemos que seria no caso de respostas de não êxito.



**Esquema 1- Dinâmica da pesquisa-ação de nossa investigação, com base em Thiollent (1986).
Fonte: Dados da pesquisa.**

A fase de *reconhecer* pode ser considerada a exploratória, pois consiste em estabelecer um primeiro diagnóstico da situação, dos conhecimentos prévios que os estudantes têm sobre o Teorema Central do Limite, sem interferência do pesquisador. O sentido de *analisar* está associado ao fato de comparar uma resposta com outra e observar como as ideias surgem e são concernentes aos objetos estatísticos e, eventualmente, associá-los aos objetos matemáticos. Em seguida, vem a fase de *orientar*, que diz respeito às interferências que podem ocorrer durante o diálogo com os estudantes, em se tratando das etapas em que foi possível estruturar as ideias para dar um contorno ao teorema. E, finalmente, *reconstruir* significa que, apesar de haver uma definição própria para o teorema, entendemos, tal como Chevallard e Joshua (1991), que o saber é reconhecido para ser produzido. A reconstrução se refere à produção do conceito que os licenciandos tiveram quando lançaram suas ideias acerca do teorema. Esse ciclo pode ser retomado a partir de respostas que porventura possam levá-los ao insucesso.

Consequentemente, pode haver a reconstrução da definição e, assim, iniciar um outro nível, obedecendo às mesmas etapas, mas, sobretudo, mais consistente que a anterior para a consolidação do saber.

4.3 A Pesquisa e a Prática: o Design Didático

Partindo do princípio da participação, verifica-se em que condições os atores, pesquisador e licenciandos, deverão atuar na investigação. Assim, Thiollent (1986) sugere os seguintes passos para concretização da investigação:

- a) Análise e delimitação da situação inicial;
 - b) Delineamento da situação final, em função de critérios de desejabilidade e de factabilidade;
 - c) Identificação de todos os problemas a serem resolvidos para permitir a passagem de (a) e (b);
 - d) Planejamento das ações correspondentes;
 - e) Execução e avaliação das ações.
- (THIOLLENT, 1986, p. 53)

As intervenções controladas ocorreram também durante as atividades, mas, efetivamente, aconteceram claramente durante o diálogo entre a pesquisadora e os licenciandos. Nesse contexto, a *literacia* estatística ficou em evidência a partir do *pensamento estatístico* (WILD; PFANNKUCH, 1999), cujo termo pressupõe leituras no sentido vago e intuitivo, as quais, em grande parte, não são examinadas acirradamente. Pode-se dizer que o pensamento estatístico é uma forma de *senso comum*, isto é, passível de entendimento a partir do momento em que é visível. Em contrapartida, sua ausência é, ao menos, aparentemente óbvia para uns, porque, de acordo com Wild e Pfannkuch (1999), tem sido muito mais produto da experiência vivida e intuição do que produto instrucional formal transmitida pelas gerações passadas. Portanto, as situações com que nos deparamos nos proporcionou momentos de reflexão e discussão, o que interferiu, mesmo que indiretamente, numa orientação em torno das ideias estatísticas que surgiram nas atividades propostas.

Desse modo, quando afirmamos que o objetivo de nosso estudo foi o de desenvolver, analisar, avaliar e reconstruir para uma proposta de ensino centrada nas principais ideias e conceitos ligados ao Teorema Central do Limite, significa dizer que tivemos um olhar voltado para ações dos alunos à medida que eles expressavam os objetos matemáticos implícitos nos objetos estatísticos. Buscamos, portanto, a análise e a compreensão dos procedimentos e argumentos que cada um se fez valer em cada situação. Podemos também dizer que, nesse ponto, se

estabeleceu uma relação produtiva entre a pesquisa e a prática pela reflexão sobre todas as propostas apresentadas na nossa investigação.

No próximo capítulo exploraremos mais detalhadamente o quadro teórico que norteou todo o nosso trabalho.

5 QUADRO TEÓRICO DA INVESTIGAÇÃO

Neste capítulo vamos apresentar as teorias da Didática da Matemática que utilizamos no desenvolvimento desta pesquisa. Para que efetivamente haja o engendramento entre o objeto estatístico e matemático de nosso estudo, no caso o Teorema Central do Limite, e a compreensão de como as ideias podem surgir durante uma sequência didática sobre o tema, optamos por uma abordagem sob o ponto de vista ecológico do saber e do didático.

Diante desse contexto, dividimos o quadro teórico segundo três pontos de vista. O primeiro trata-se do objeto do saber, o Teorema Central do Limite, numa abordagem ecológica; o segundo, especificamente, o ponto de vista do didático; e, o terceiro, no que se refere à aplicabilidade do teorema no nosso entorno, isto é, as implicações do teorema na *literacia* estatística.

Na primeira parte, visamos o objeto, o teorema em questão, sob o ponto de vista ecológico do saber, isto é, quais são os elementos matemáticos e estatísticos indispensáveis para a compreensão do Teorema Central do Limite? Quais os dispensáveis que permitiram sua “sobrevivência” até nos dias de hoje?

E, em seguida, na segunda parte, sob o ponto de vista ecológico do didático, tratamos do processo de ensino e de aprendizagem, já que podemos afirmar que a dificuldade, em certas ocasiões, é potencializada pelo fato de o professor de Matemática, responsável pelo ensino desse conteúdo na Educação Básica e no Ensino Superior, segundo Shulman (1986), precisa ter mais conhecimentos específico e didático do conteúdo, justificando, assim, o nosso trabalho com os licenciandos em Matemática. Partimos da hipótese de que a compreensão do Teorema Central do Limite, na transição entre a Descritiva e a Inferencial, deve estabelecer os níveis de ensino, segundo a Teoria da Transposição Didática.

E, na terceira parte, trataremos de expor algumas ideias a respeito da *literacia* estatística no intuito de poder apresentar uma relação direta ou indireta entre o teorema e a proposta nessa vertente.

5.1 A Problemática Ecológica sob o Ponto de Vista do Saber

Sob o ponto de vista ecológica do saber, entendemos que a problemática deve envolver uma abordagem antropológica do objeto, no caso, o Teorema Central do Limite, nos possibilitando retomar a nossa questão de cunho ecológico de nossa investigação:

Por que determinados saberes/conhecimentos deixaram de ser abordados no ensino do Teorema Central do Limite, mas ainda assim ele “sobrevive”? Que saberes são indispensáveis para teorema “viver”? Que relações diretas ou indiretas existem entre o teorema e a construção de saberes estatísticos nas aulas de Estatística?

Para aprofundarmos um pouco mais a relevância ecológica atribuída à problemática desse estudo, faremos um breve esboço sobre o que vem a ser uma estrutura ecológica a partir de um ecossistema.

5.1.1 A Noção de Ecossistema

A criação da palavra ecologia veio de Haeckel, em 1866. Ele definiu “ecologia” como a ciência que engloba todas as relações dos organismos com o mundo exterior que os envolve, incluindo suas condições de existência, isto é, as relações extrínsecas e intrínsecas dos organismos com o meio, formando uma estrutura ecológica chamada de ecossistema.

Segundo Dajoz (2005, p. 244), ecossistema é “um conjunto de elementos em interação uns com os outros, formando um todo coerente e ordenado. É um sistema hierarquizado no qual os próprios elementos constitutivos são subsistemas estruturados.”

Para Dajoz (2005), o termo biocenose é atribuído ao conjunto de espécies que ocupam e partilham um mesmo meio, bem delimitado, no qual foram reunidas de maneira fortuita. Esses conjuntos de espécies coexistem regularmente e de forma

previsível. São também as mesmas que podem manter, entre si, relações que resultam em uma verdadeira coevolução. Em outras palavras, as diversas espécies não são independentes: os elementos do ecossistema são interdependentes, porque mantêm entre si relações múltiplas, coerentes e ordenadas, formando, assim, um conjunto, em geral, estável e autônomo.

O trabalho de Artaud (1998) identifica quatro tipos de ecossistema para o saber matemático:

[...] ecossistema do saber, no qual se produz a matemática; ecossistema didático escolar, no qual se estuda a matemática; ecossistema profissional, onde utilizam a matemática para concretizar algumas tarefas; ecossistema noosferiano, enfim, em que a matemática é manipulada para fins de transposição. (ARTAUD, 1998, p. 4)

Por conseguinte, nessa estrutura, reconhecemos os ecossistemas que abarcam o Teorema Central do Limite: (i) quanto às condições e motivações que permitiram a produção do teorema; (ii) quanto ao estudo do teorema no sistema escolar; (iii) quanto ao estudo do teorema para a construção da *literacia* estatística no cotidiano profissional e pessoal; (iv) quanto ao estudo do teorema no processo de ensino e de aprendizagem.

O item (i) refere-se ao da revisão da literatura, em que foi possível traçar um breve esboço histórico e epistemológico sobre as condições do desenvolvimento do Teorema Central do Limite, desde as primeiras ideias até a forma como é conhecida nos dias de hoje. Os itens (ii), (iii) e (iv) fazem parte dos resultados da pesquisa, portanto, objetos de estudo no capítulo ulterior.

Contudo, para se ter uma ideia da estrutura ecológica do teorema como objeto do saber, entendemos que um ecossistema tem caráter paradoxal devido à dependência e independência em relação a outros ecossistemas. A dependência nos proporciona uma visão global do nosso estudo, e a independência nos permite distinguir o surgimento de objetos do saber a partir das diferentes tarefas que requerem a instrumentalização dos mesmos. Assim, para o Teorema Central do Limite, podemos dizer que a dependência global pode ser representada pela importância que a maioria dos autores dos livros-texto atribuem ao teorema, pela aplicação na Inferência Estatística e pelos objetos matemáticos e estatísticos que são indispensáveis para seu estudo – como, por exemplo, a definição da curva

normal. Para a independência do ecossistema do teorema, exemplificamos a partir de sua abordagem histórica e epistemológica, tal qual como foi apresentada.

Vale ressaltar que os objetos do saber e os objetos didáticos estão associados, desde que as organizações do saber iniciaram suas existências para pessoas ou instituições, a partir do processo de estudo.

5.1.2 Abordagem Ecológica do Saber: O Teorema Central do Limite

Para iniciarmos a apresentação do ecossistema do teorema, faz-se necessário explicitar a sua definição. Segundo Guimarães e Cabral (1997), o Teorema Central do Limite,

[...], no domínio da Estatística, constitui um dos desenvolvimentos teóricos mais notáveis, com inúmeras aplicações — permite, em particular, fazer progressos significativos na caracterização de distribuições por amostragem. De uma forma extremamente simplificada, o teorema pode ser enunciado nos seguintes termos:

$$S = \sum_{n=1}^N X_n$$

Sejam $X_1, \dots, X_N$ variáveis aleatórias independentes com a mesma distribuição, que se admite ter variância finita (quase todas as distribuições com interesse prático têm variância finita, pelo que esta condição não é particularmente restritiva). Qualquer que seja a forma da distribuição destas variáveis, se o valor N (sic) for suficientemente grande, a variável soma segue aproximadamente uma distribuição Normal.

Esta distribuição é inteiramente especificada através do valor esperado e da variância de S , que são dados por

$$\begin{aligned}\mu &= N \cdot \mu_x \text{ (sic)} \\ \sigma_s^2 &= N \cdot \sigma_x^2 \text{ (sic)}\end{aligned}$$

onde μ_x e σ_x^2 , representam o valor esperado e a variância das variáveis X_n , (GUIMARÃES; CABRAL, 1997, p. 240)

Para a definição que os autores apresentam, vale ressaltar que consideramos, em geral, N , (maiúsculo), para representar o tamanho da população, e o n minúsculo para amostras, fato este não exposto por eles. Assim, o resultado dessa definição pode ser interpretado da seguinte forma: para qualquer população com variância finita, a distribuição da média amostral calculada com base numa amostra simples tende para uma distribuição normal à medida que a dimensão da amostra cresce. Como a média amostral é dada pelo produto entre a variável soma

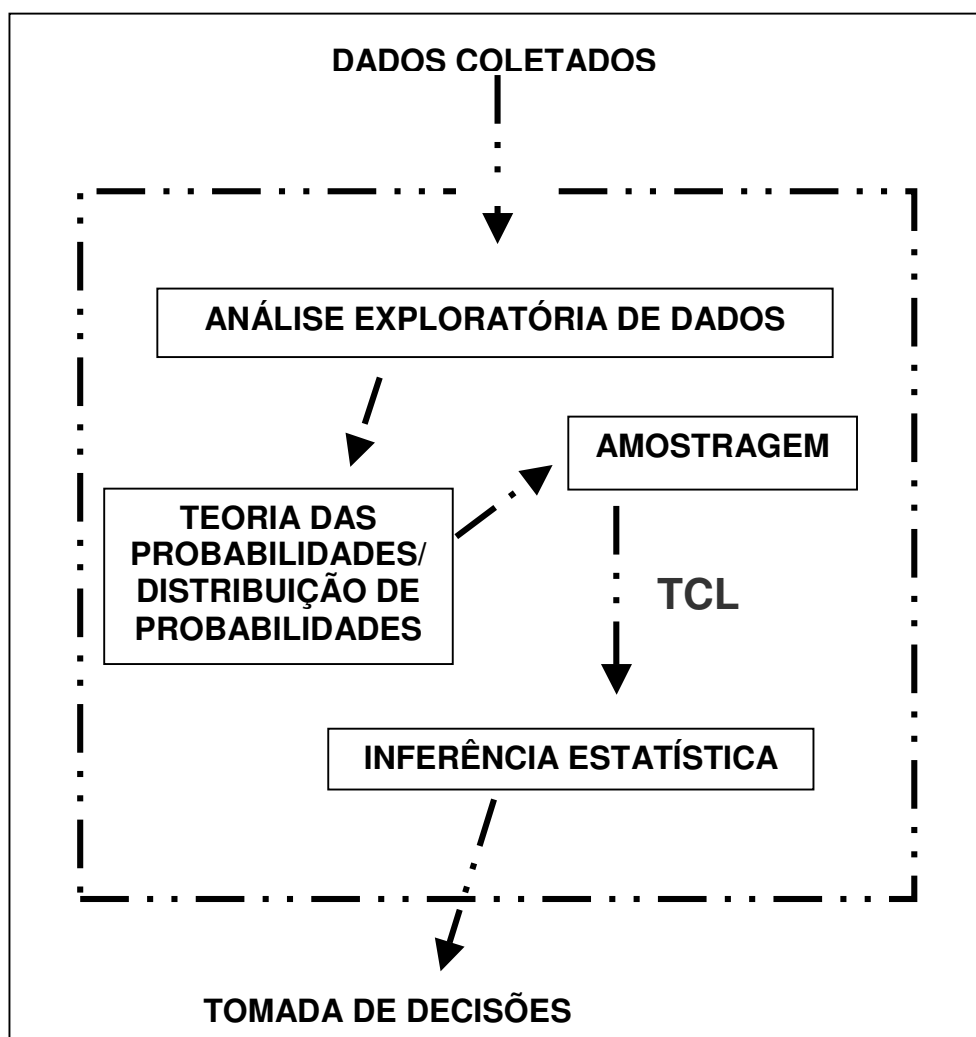
(S) pelo coeficiente $k = \frac{1}{N}$, então a distribuição da soma se aproxima de uma distribuição normal. E, para que tal fato aconteça, é preciso que satisfaça a duas condições:

1ª) Se S é variável aleatória, então $k.S$ também é variável aleatória;

2ª) Se $S \sim N(\mu, \sigma^2)$, então $k.S \sim N(k\mu, k^2\sigma^2)$.

Satisfeitas as duas condições acima, tem-se que o mesmo sucederá à distribuição da média amostral.

A seguir, apresentamos um esquema que pode representar um primeiro esboço de um ecossistema associado ao Teorema Central do Limite:



Esquema 2: Esboço de um ecossistema para o Teorema Central do Limite.
Fonte: Dados da pesquisa.

O esquema 2 retrata um primeiro esboço de ecossistema para o Teorema Central do Limite a partir dos dados coletados. É a parte inicial do tratamento estatístico a ser realizado para seguir uma estrutura ecológica. A análise exploratória dos dados, neste caso, os procedimentos, envolve as medidas de posição central e de variabilidade, culminando nas técnicas estatísticas que podem permitir a construção de efeitos conclusivos por meio das representações gráficas. Em seguida, os dados incidem nos fundamentos da Teoria das Probabilidades e, quando na forma de distribuição de probabilidades, seguem para as técnicas da Amostragem. Essas técnicas permitem aplicação do Teorema Central do Limite, constituindo-se o fundamento para Inferência Estatística.

Podemos dizer que a *biocenose* do teorema é o conjunto de saberes indispensáveis que o faz existir na estrutura do conhecimento estatístico. **Mas quais são esses saberes? Quais os objetos matemáticos indispensáveis para compreensão do Teorema Central do Limite? E os "dispensáveis", isto é, os que não estão aparentes, de modo que, assim mesmo, o teorema sobreviva no meio biótico?**

Artaud (1998) parte do questionamento ecológico da teoria da transposição didática. O termo 'transposição didática' é atribuída a Chevallard (1991, p. 39) quando afirma que "um conteúdo do saber, que é destinado ao saber a ser ensinado, sofre um conjunto de alterações no sentido de adaptar com mais eficiência seu lugar entre os objetos da educação. Esse 'trabalho' que acontece com o saber a ser ensinado é chamado de *transposição didática*."¹⁶

Nesse sentido, os objetos matemáticos utilizados no teorema estão contidos na organização matemática, que depende de três condições fundamentais para sua existência. Na primeira condição, o objeto matemático ensinado deve ser compatível com seu meio social, neste caso, a Matemática que os licenciandos têm durante o curso.

Na segunda, as ideias matemáticas que estão em torno do teorema devem respeitar a sequência didática em relação ao tempo didático, isto é, sob o ponto de vista do ensino. Chevallard, Bosch e Gastón (2001, p. 122) propõem "[...] a reconstrução das obras matemáticas selecionadas no currículo *como obras que*

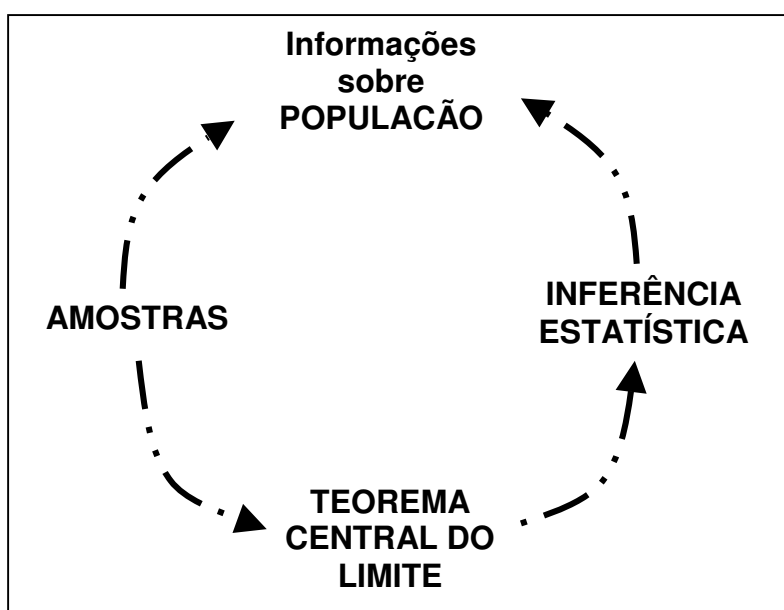
¹⁶ "Un contenu de savoir ayant été désigné comme savoir à enseigner subit dès lors un ensemble de transformations adaptatives qui vont le rendre apte à prendre place parmi les *objets d'enseignement*. Le 'travail' qui d'un objet de savoir à enseigner fait un objet d'enseignement est appelé *la transposition didactique*." (CHEVALLARD, 1991, p. 39) (Tradução nossa)

devem ser estudadas, e não só ensinadas.”, respeitando assim a sequenciação e a temporalização didática. De fato, esse estudo se desenvolve, tendo como pano de fundo os cursos de Licenciatura em Matemática, porque se pressupõe que os alunos já devem ter cursado disciplinas que dizem respeito aos objetos matemáticos necessários para o estudo do Teorema Central do Limite, tais como noção de limite, derivada, integral e variável aleatória. São satisfeitas, assim, as duas primeiras condições.

A terceira condição, que será explorada ainda neste capítulo, diz respeito às relações institucionais entre a posição do professor e a posição do aluno. Em termos de Chevallard (1996), a palavra *objeto* é entendida em seu sentido amplo, isto é, todos os elementos da didática podem ser chamados de *objetos*, da mesma forma como atribui o significado a “instituições”. Então, essas relações institucionais estabelecem-se a partir do *meio*, pois será ele que permitirá “uma série de questões ‘ecológicas’ para que uma situação didática possa, não apenas existir, mas continuar a existir: para que possa *funcionar*.” (CHEVALLARD, 1996, p. 134).

A questão ecológica inicial do Teorema Central do Limite diz respeito à identificação dos objetos indispensáveis de forma que o teorema permaneça vivo. Para que isso ocorra, como alega Dajoz (2005), faz-se necessário conhecer a estrutura trófica¹⁷ das biocenoses, isto é, os elementos que se interrelacionam para nutrir os ‘tecidos’ do objeto ecológico. A formação dessa estrutura é derivada do resultado de uma cadeia alimentar. Por exemplo, uma “cadeia alimentar” simples para o teorema poderia ser representado assim:

¹⁷ Segundo Dajoz (2005), a estrutura trófica constitui-se nos elementos que estão presentes na cadeia alimentar e estabelece as relações de natureza alimentar que se mantêm entre si, ou seja, nos elementos imprescindíveis que mantêm existência ecológica do objeto de estudo.

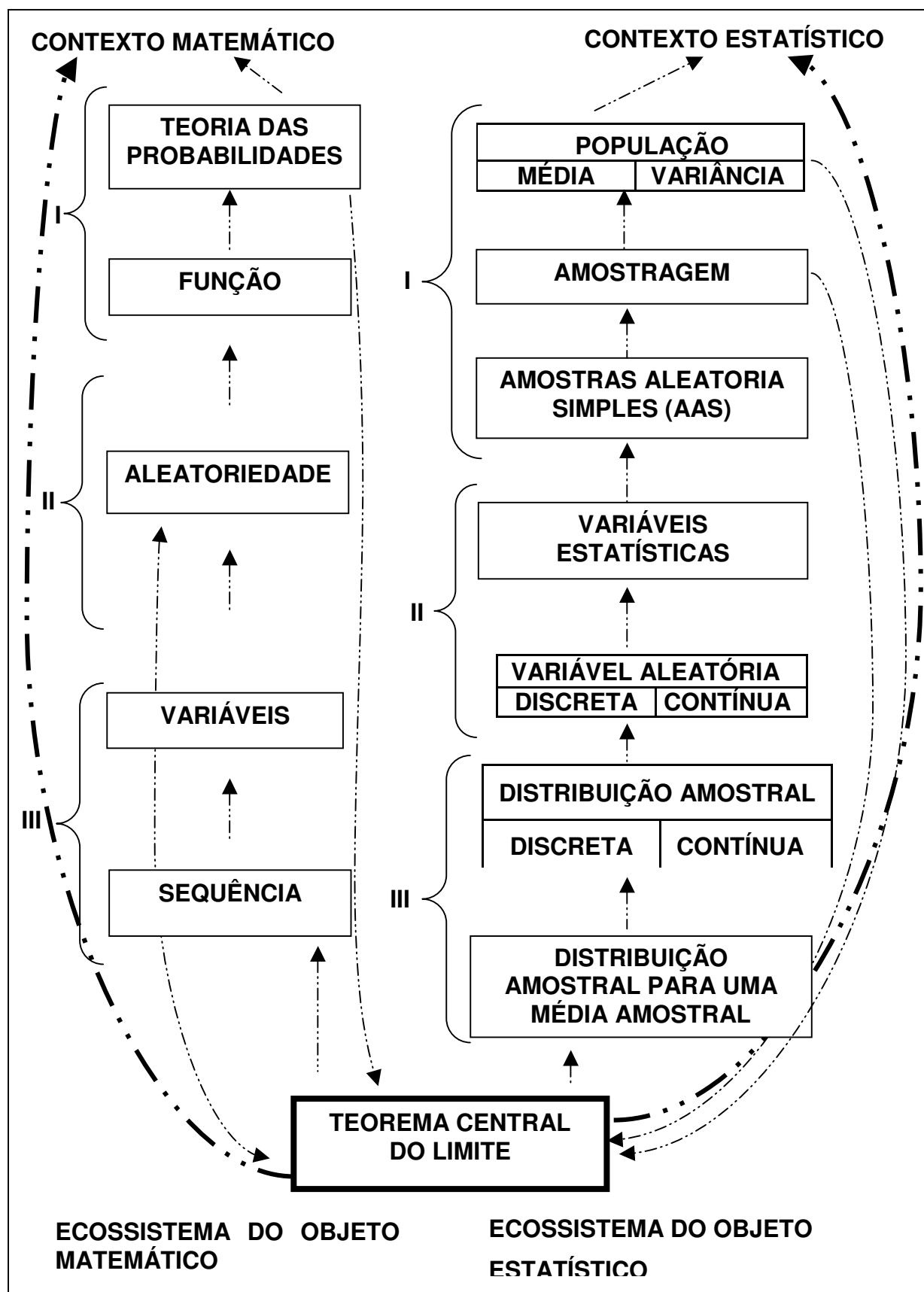


Esquema 3 - Cadeia elementar simples para o Teorema Central do Limite.
Fonte: Dados da pesquisa.

O esquema 3 apresenta uma cadeia elementar simples e nos conduz a duas ideias essenciais do teorema, a Amostragem e a Lei dos Grandes Números, já que “a inferência estatística usa dados amostrais para tirar conclusões sobre a população inteira” (MOORE, 2005, p. 216). Por exemplo, ao estimar a renda média de famílias que moram num determinado bairro da cidade de Juiz de Fora-MG, quanto maior for o número de observações da amostra aleatória, mais preciso será o procedimento estatístico. Entendemos então que nesse exemplo foi aplicada a Lei dos Grandes Números.

Apesar da simplicidade dessa cadeia, nem todas as pessoas sabem que, por trás das informações, como, por exemplo, o caso da renda média das famílias, existe uma teia de complexidade que envolve objetos e modelos matemáticos de tal forma que a validação da afirmativa tem respaldo científico.

Nessa direção, uma forma de representar a estrutura ecológica do teorema é fazê-lo pertencer, simultaneamente, a duas cadeias alimentares, formando, assim, redes tróficas de razoável complexidade. A importância dessa estrutura complexa é o resultado que se conhece sobre o teorema, e é a partir daí que, supostamente, o Teorema Central do Limite atinge um nível estável na estrutura ecológica do saber. Vejamos um esquema de duas redes tróficas no ecossistema do Teorema Central do Limite, na partilha entre os ecossistemas do modelo matemático e da estatística, conforme o esquema 4.



Esquema 4 - Esquema de redes tróficas em um ecossistema mostrando a partilha entre dois ecossistemas: o Teorema Central do Limite enquanto objeto matemático e objeto da estatística.

Fonte: Dados da pesquisa.

Os itens (I), (II) e (III) correspondem respectivamente aos níveis dos “consumidores”: primário, secundário e terciário. Esse esquema apresenta o Teorema Central do Limite “alimentando” e “sendo alimentado” por dois ecossistemas distintos: matemático e o estatístico.

As biocenoses dos ecossistemas apresentadas reiteram as interdependências entre os elementos que fazem parte da estrutura. Essa analogia com as redes tróficas permite buscar fatores que determinam estabilidade das comunidades ecológicas do saber. No nosso caso, elas determinam a estabilidade das comunidades ecológicas do Teorema Central do Limite. Segundo Dajoz (2005), existe uma forte tendência de proporcionalidade direta entre o que tange à complexidade dessas comunidades e à estabilidade das mesmas, e “muitos ecólogos buscam os fatores que determinam a estabilidade das comunidades, e procuram saber particularmente se as comunidades complexas são mais estáveis do que as comunidades simples” (DAJOZ, 2005, p. 251).

Diante disso, criamos a hipótese de que o teorema sobrevive pela aplicabilidade tanto como objeto matemático quanto estatístico.

Segundo o esquema 3, a interrelação entre dois ecossistemas pode ser declarada quando autores, como Wonnacott e Wonnacott (1991), apresentam o Teorema Central do Limite, dividindo-o em três pontos-chave: (i) é um teorema, deve ser demonstrado a partir de duas hipóteses: a primeira é quanto à independência das variáveis aleatórias e a segunda é que estas variáveis tenham variância finita; (ii) é um teorema limite, portanto significa que sua conclusão foi cuidadosamente estabelecida na forma de um limite: a probabilidade que uma realização da variável

centrada reduzida¹⁸, $z = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}$, – isto é, a variável reduzida, z , é obtida pela média

das amostras ($\bar{x}$), pela média da população (μ), pelo desvio padrão da população (σ) e, por fim, pelo tamanho da amostra selecionado (n) – que pertença a um dado intervalo que converge para o limite da probabilidade, em que a variável normal centrada reduzida z pertença a esse intervalo; e, finalmente, o terceiro: (iii) é um

¹⁸ Segundo Magalhães e Lima (2002, p. 184), “[...] as probabilidades para o modelo Normal são calculadas com o auxílio de tabelas. Para se evitar a multiplicação desnecessárias de tabelas para cada par de valores (μ , σ^2), utiliza-se uma transformação que conduz sempre ao cálculo de probabilidades com uma variável de parâmetros (0, 1), isto é, média 0 e variância 1.” E essa variável é a centrada reduzida, valor de z .

teorema central, no sentido que descreve como a média amostral, $\bar{x}$, que se concentra em torno de seu valor central, μ , a média da população.

Para conhecermos a estrutura trófica do Teorema Central do Limite, alguns objetos matemáticos e estatísticos são imprescindíveis, pois estabelecem relações entre si na cadeia alimentar para manter ‘vivo’ o teorema. A primeira hipótese se refere à independência das variáveis aleatórias. Vejamos quais são os elementos que “alimentam” essa primeira hipótese.

- Amostras Aleatórias Simples

O ponto de partida para a exposição de amostras aleatórias simples é a distinção entre população e amostra. Entende-se por população um grupo inteiro de indivíduos sobre os quais desejamos obter informações; amostra é a parte dessa população que efetivamente examinamos com objetivo de reunir informações.

A extração de uma determinada amostra requer procedimentos apropriados de tal forma que permita reproduzir as mesmas características da população e, para isso, existem métodos de amostragem. Eles podem ser aleatórios ou não.

O pesquisador tem como principal preocupação verificar se os elementos da amostra são suficientemente representativos de toda a população, de modo a permitir generalizações precisas sobre aquela população. A fim de fazer essas inferências, o pesquisador escolhe um método adequado de amostragem que permita que cada elemento da população tenha a mesma chance de ser incluído na amostra. Se todo elemento da população tem igual chance de escolha, estamos pondo em prática um método de amostragem *aleatório*, caso contrário, o método de amostragem é *não aleatório*. (LEVIN; FOX, 2004, p. 178)

Uma amostra aleatória simples pode ser com reposição, caso uma unidade possa ser sorteada mais de uma vez, e sem reposição, se a unidade sorteada for removida da população. Para autores como Bussab e Morettin (2003, p. 262), do ponto de vista da quantidade de informação contida na amostra, o procedimento sem reposição é mais adequado. Contudo, a amostragem com reposição pode simplificar o tratamento teórico, pois a *independência* entre as unidades selecionadas facilita o desenvolvimento das propriedades dos estimadores que serão considerados.

Nos casos para os quais a população seja muito grande e diversificada, existem procedimentos para construção de amostras (não aleatórias), tais como amostras intencionais, por cotas ou por julgamento.

A extração de elementos para amostras aleatórias simples exige técnicas específicas de forma que contemplem a metodologia da pesquisa em questão. Levin e Fox (2004, p. 180) afirmam que “todos os métodos de amostragem aleatória não são mais do que variações do processo de amostragem aleatória simples”.

- O tamanho das amostras

Para o Teorema Central do Limite ainda tem a questão do tamanho da amostra. Mas quão grande deve ser uma amostra para que o teorema seja aplicado?

Alguns autores, como Stevenson (2001) e Mann (2006), adotam amostras a partir de 30 (elementos) e, utilizando o Teorema Central do Limite, inferem que a média das amostras tende a ser a da população. Mas, por outro lado, existem outras obras em que esse número é reduzido para 25, o tamanho mínimo de amostras, por exemplo, a obra de Witte e Witte (2005).

Diante dessa indefinição, qual deve ser o tamanho da amostra para que se aplique o teorema? Nesse contexto, optamos por citar Triola (1999):

Se a população original é em si normalmente distribuída, então as médias de amostras de qualquer tamanho serão normalmente distribuídas. Se a população original não é em si normalmente distribuída, então dizemos que a média de amostras de tamanhos $n > 30$ tem uma distribuição que é aproximada por uma distribuição normal. A condição de que o tamanho da amostra seja $n > 30$ é comumente usada como uma diretriz, mas não é possível identificar um tamanho amostral mínimo específico que seja suficiente para todos os casos. O tamanho amostral mínimo, na verdade, depende de como a distribuição populacional se afasta de uma distribuição normal. Tamanhos amostrais de 15 a 30 são adequados se a população parece ter uma distribuição que não se afasta muito da normal, mas algumas outras populações têm distribuições que são extremamente distantes da normal, e então tamanhos amostrais de 50, ou mesmo 100, podem ser necessários. Utilizamos o critério simplificado de $n > 30$ como justificativa para tratar a distribuição das médias amostrais como uma distribuição normal. (TRIOLA, 1999, p. 246)

De acordo com a citação, o tamanho das amostras estabelece uma relação direta com a população, por isso o caráter relativista em contraposição a um critério simplificado como por conveniência para $n > 30$. Nem sempre é possível conhecer a população a não ser que sua natureza seja apresentada e, diante disso, deve-se levar em consideração que não há um número ideal para o tamanho da amostra, mas, sim, um modo simplificado, contemplando as propostas de atividades de ensino que aparecem nos textos. Portanto, parece senso comum, ou por

conveniência, a adoção de amostras de 30 ou mais médias amostrais para aplicação do teorema.

Um outro aspecto a ser considerado é o fator de proporcionalidade de uma amostra em relação à população. Como nos casos acima, há também o senso comum, ou por conveniência, sobre o tamanho da proporção de uma determinada população quando se deseja examinar um item. Portanto, surge a necessidade de uma referência para delimitar um percentual mínimo de modo que a amostra seja significativa. Geralmente é considerada uma amostra superior ou igual a 5% da população como aquela que é significativa. Segundo Stevenson (2001):

Se o tamanho da amostra é pequeno em relação ao da população, a não-reposição do item examinado terá efeito desprezível nas probabilidades dos itens restantes, e a amostragem sem reposição não causará dificuldades sérias. Por outro lado, amostras relativamente grandes tendem a distorcer as probabilidades dos itens restantes no caso de amostragem sem reposição. Uma regra prática geralmente aceita é fazer a reposição quando o tamanho da amostra excede 5% do tamanho da população. (STEVENSON, 2001, p. 159)

Assim, o tamanho das amostras constitui também elemento indispensável na estrutura ecológica do Teorema Central do Limite.

- Variáveis Estatísticas

Variável estatística é a característica que se quer observar em um conjunto de dados, averiguando se é qualitativo ou quantitativo. Para Graham (2006):

[...] Palavras como “quantos” e “menos” referem-se às medidas discretas, separada, itens contáveis e, por outro lado, “quanto” e “menos” referem-se a algo que não pode ser contado, tais como quantidade de água, fatia de torta e, assim por diante. Para distinguir esses termos, utiliza-se na estatística “discreta” e “contínua”, respectivamente. (GRAHAM, 2006, p. 10)¹⁹
(Tradução nossa)

Por conseguinte, tratamos de variáveis estatísticas discreta e contínua. Na língua portuguesa, os vocábulos “fewer” e “less” significam “menos”. Porém, na Língua Inglesa, “fewer” é uma forma usada para substantivos contáveis e “less” para substantivos incontáveis, o que justifica a relação entre os elementos representados pela variável discreta e pela variável contínua, conforme exposto no texto.

¹⁹ “[...]. Words like “how many” and “fewer” refer to measures of discrete, separate, countable items, whereas “how much” and “less” refer to something that cannot be counted out, such as amount of water, size of slice of a pie, and so on. The terms used in statistics to make this distinction are “discrete” and “continuous”, respectively.” (GRAHAM, 2006, p. 10)

- Variáveis Aleatórias

Triola (1999, p. 93) define “uma variável aleatória é uma variável (geralmente representada por X) que tem um valor numérico único (determinado aleatoriamente) para cada resultado de um experimento”. Por exemplo, consideremos o experimento de observar a pontuação dos alunos, por turma, na 5ª Série, ou 6º ano, do Ensino Fundamental, obtida a partir de instrumento de mensuração do índice de autoestima. Podemos associar, a cada turma, o valor X que indica a média aritmética dos pontos dos alunos daquela turma.

Dessa forma, se a escolha da turma em questão for feita por um sorteio aleatório (experimento aleatório), então se pode dizer que X é uma variável aleatória.

Meyer (1983) apresenta variáveis aleatórias a partir de situações de experimentação, como, por exemplo, “atribuir o valor *um* às peças perfeitas e o valor *zero* às defeituosas” de um determinado objeto. Como definição, o autor apresenta: “seja ε um experimento e S um espaço amostral associado ao experimento. Uma função X , que associe a cada elemento $s \in S$ um número real, $X(s)$, é denominada variável aleatória”. (MEYER, 1983, p. 66)

O modelo matemático que descreve uma variável aleatória é de uma *função*, pois associa cada um dos eventos resultantes de uma experiência aleatória a um número real, seja esse número definido em um conjunto enumerável ou não enumerável. Observa-se que as duas definições apresentadas são equivalentes.

- Distribuição de Probabilidades

De acordo com Levin e Fox (2004, p. 145), “uma distribuição de probabilidades é diretamente análoga a uma distribuição de frequências, com a única diferença de que se baseia na teoria (teoria das probabilidades) e não no que se observa no mundo real (dados empíricos)”. Matematicamente, podemos descrever uma distribuição de probabilidades como uma função P , denominada função de probabilidade, para os casos de variável aleatória discreta ou contínua:

- $P(X = x_i) = p(x_i) = p_i$, $i = 1, 2, \dots$ para o caso da variável discreta;
- f é uma função contínua de probabilidade ou função densidade de probabilidade para uma variável aleatória contínua X . Satisfazem-se duas condições: (i) $f(x) \geq 0$, para todo $x \in (-\infty, \infty)$; e, (ii) $\int_{-\infty}^{+\infty} f(x)dx = 1$.

- Distribuições Amostrais

Para Moore (2005, p. 219), “a distribuição amostral de uma estatística²⁰ é uma distribuição dos valores assumidos pela estatística em todas as amostras possíveis de mesmo tamanho de uma mesma população”. Por exemplo, podemos simular a medida das alturas de 1000 estudantes de escola pública, na faixa etária de 9 a 11 anos, e extrairmos, aleatoriamente, várias amostras de 10 crianças, obtendo, assim, uma distribuição de amostras dessa população. Daí, podemos dizer que as amostras foram extraídas a partir de um número fixo de ensaios e as informações obtidas sobre a média das alturas das amostras será uma aproximação da distribuição amostral, ou uma estatística da amostra apresentada. Um dos modelos para obtenção dessas distribuições pode ser representado pela Teoria das Probabilidades.

- Distribuição Amostral de uma Média Amostral

A distribuição amostral de uma média amostral retrata os diversos valores que essa mesma média pode assumir de uma única população. E, além disso, Mann (2006, p. 295) garante que também ocorre da mesma forma para a probabilidade de cada uma delas, isto é, das médias amostrais. As implicações dessa definição são fundamentais para Inferência Estatística:

- A média de estatística $\bar{X}$ é sempre igual à média μ da população. Ou seja, a distribuição amostral de $\bar{X}$ tem seu centro em μ . [...]
- [...] Se observações individuais têm desvio padrão σ , então, as médias amostrais $\bar{X}$ de amostras de tamanho N têm desvio padrão $\sigma/\sqrt{N}$. [...] (MOORE, 2005, p. 221)

Consequentemente, a distribuição de médias amostrais é menos dispersa, uma vez que $\frac{\sigma}{\sqrt{N}} \leq \sigma$, pois $N \geq 1$. Se a distribuição da população segue uma distribuição normal, então ocorrerá o mesmo para a média amostral, conforme figura 1:

²⁰ Chama-se ‘estatística’ uma característica da amostra. Por exemplo, a média da amostra, a variância da amostra, o menor valor da amostra, entre outros.

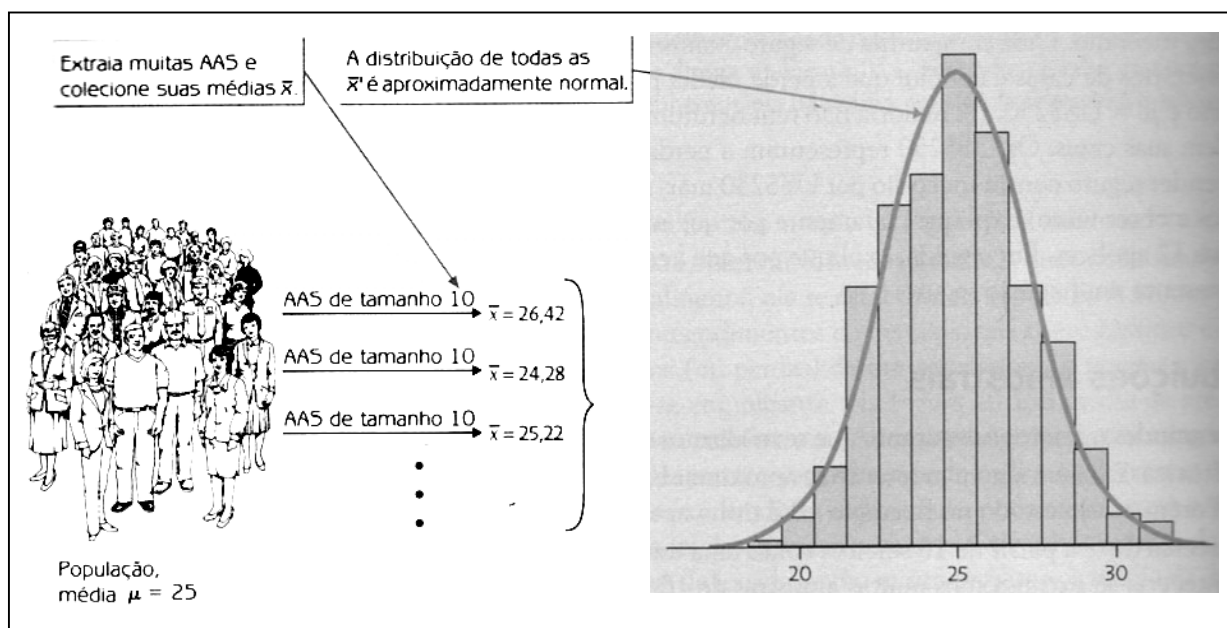
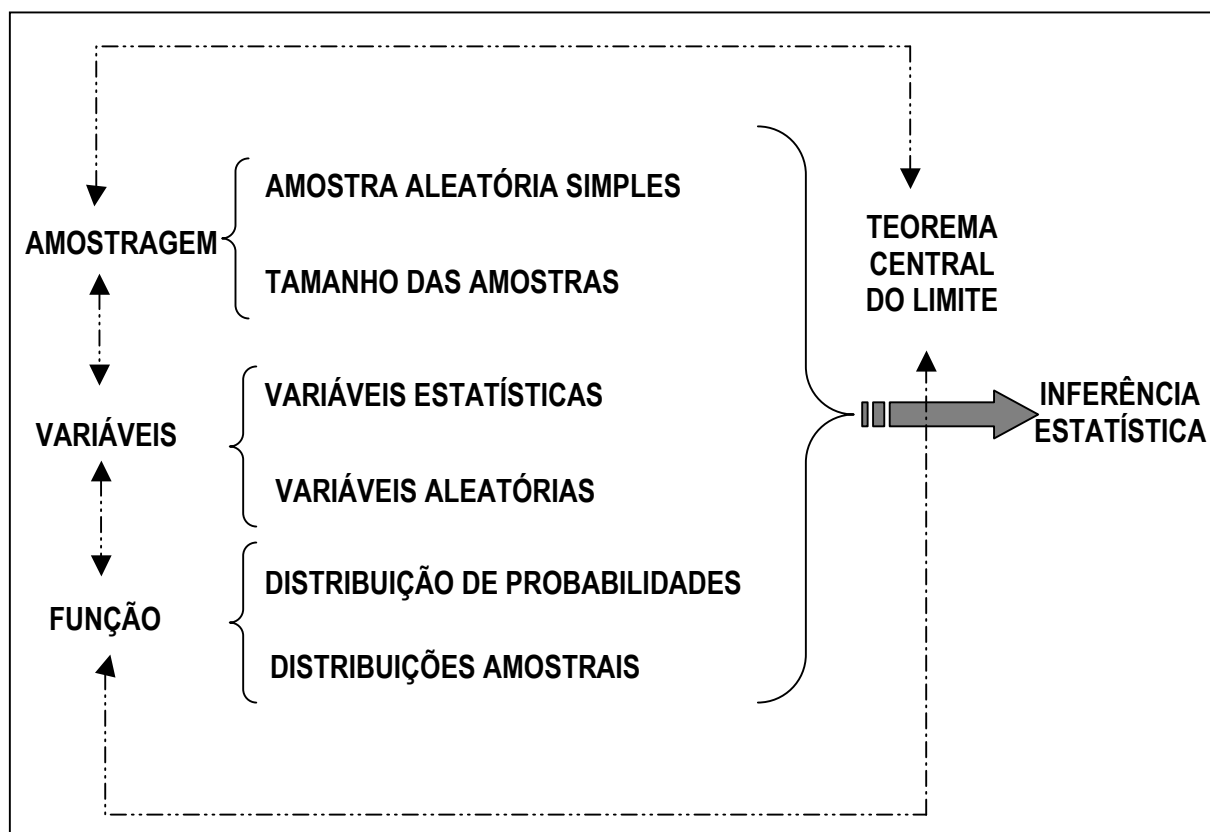


Figura 1 – A ideia de uma distribuição amostral.
Fonte: MOORE, 2005, p. 220

Observe o exemplo apresentado por Moore (2005, p. 219), figura 1, na mensuração dos limiares de odor de DMS de adultos individuais, em que os valores seguem a distribuição normal com média $\mu = 25$ microgramas por litro e desvio padrão $\sigma = 7$ microgramas por litro. Ao selecionar amostras de tamanho 10 dessa população, os valores da média amostral $\bar{X}$ encontrados variam para as amostras, pelo fato de que o desvio-padrão da amostra também diminui, contribuindo, assim, para que os valores estejam mais em torno da média amostral. Esse resultado torna os valores das médias das amostras bem próximos do parâmetro verdadeiro da média da população, μ .

O ecossistema do Teorema Central do Limite foi construído a partir desses objetos estatísticos abordados, o que permite a sua “sobrevivência” na Inferência Estatística. A cadeia alimentar para o teorema, em síntese, pode ser representado pelo esquema 5:



Esquema 5: Ecossistema do Teorema Central do Limite
Fonte: Dados da pesquisa

Observando o esquema 5, percebemos que o Teorema Central é “alimentado” pelos elementos externos que são constituídos por elementos endógenos, indispensáveis para que o teorema sobreviva neste *habitat* e chegue à Inferência Estatística.

Borovcnik (2005) declara que uma forma dos erros de hipóteses elementares da física manifesta-se por um exemplo concreto do Teorema Central do Limite. Isso porque o teorema estabelece a restrição de que uma distribuição de variáveis aleatórias independentes tende a uma distribuição normal, conforme já citado anteriormente. Em termos de uma distribuição normal, e quando a variação decresce à medida que aumenta o número de provas (por meio da raiz quadrada do número dessas provas), a aleatoriedade serve como um dos principais elementos de vários procedimentos estatísticos.

Diante desses resultados, a definição do Teorema Central do Limite, formado a partir de uma estrutura de rede trófica que alimenta dois ecossistemas do teorema, pode ser classificada como o matemático e o estatístico, representando, portanto,

uma abordagem sob o ponto de vista ecológico do saber. A seguir, apresentaremos o quadro teórico que servirá como fundamento para a análise dos resultados de nossa investigação, cuja vertente também é de cunho ecológico, mas com ênfase no didático.

5.2 Abordagem Ecológica do Didático

O nosso estudo visa diminuir o distanciamento entre o saber e a didática, valorizando o processo de ensino e de aprendizagem, por meio de questionamento ecológico que permite a aproximação do didata. Retoma, dessa forma, a abordagem de Artaud (1998) quanto à problemática ecológica do objeto sobre a forma de questionar o real, isto é, “o que existe? Por que existe?” Ou então, o que não existe? Poderia existir? Sobre quais condições? Por outro lado, a partir de um conjunto de condições, quais objetos são obrigados a viver ou, em contrapartida, quais são impedidos de viver nessas condições? Aparentemente essas questões parecem ser triviais, entretanto, no âmbito do didata, pode permitir o domínio da realidade dos objetos que desejamos estudar e, no caso, de todos os *objetos* que se interrelacionam no processo de estudo, como os alunos, os professores, as intuições e o saber a ser ensinado.

Existe também, assim como a abordagem ecológica do saber, uma organização própria para o estudo ecológico dos objetos didáticos, que, amiúde, podem ser os objetivos, métodos, estratégias de avaliação, isto é, os elementos que constituem os processos de ensino e de aprendizagem do teorema.

Dessa forma, fundamentamo-nos na teoria da Transposição Didática (CHEVALLARD, 1985, 1991), porque podemos identificar os conjuntos de condições que permitem aos matemáticos existirem no sistema didático.

A idéia da *transposição didática* foi introduzida no início da década de 80, por Chevallard (1985). Nessa época, ela foi entendida intuitivamente “como o trabalho de adaptação, transformação do saber em objeto de ensino, em função do lugar, do público e das finalidades didáticas a que se propõe” (D’AMORE, 2007, p. 224).

O funcionamento didático depende de como o professor irá conduzir cada fase da aula, porque a recontextualização do saber propicia novas criações didáticas, conforme afirmam Joshua e Dupin (1993):

[...] as transformações que excedem o único conhecimento do texto para cobrir o aumento das características de uma situação didática é inevitavelmente uma certa medida. Desta forma, as transformações são viáveis? Na verdade, a questão já é a única transformação do texto de conhecimento. De fato, as restrições sobre a aplicação específica que conduziram à identificação dos itens de educação, uma vez feito, eles podem realmente "viver" de forma sustentável no sistema educativo concreto. Em particular, é necessário que a integração possa ser feita com os antigos elementos que não foram modificados.

O estudo da transposição didática estende então na viabilidade de um determinado objeto de ensino. Não há dúvida de que um conhecimento detalhado dessas condições seria de grande ajuda para definir o leque de possibilidades no ensino da ciência, e as esperanças para evitar a decepção e perda de energia correspondente. Um verdadeiro conhecimento ecológico é necessária [...]. (JOSHUA; DUPIN, 1993, p. 202-203) (Tradução nossa)²¹

Seguindo essa abordagem, o que de fato faz o professor, o saber a ser ensinado e o aluno “sobreviver” no âmbito acadêmico? Quais as relações institucionais que permitem essa “sobrevivência”? Como impedir a perda de “energias” que existem nas relações estabelecidas entre os sujeitos no *habitat*? Assim, a transposição didática consistiria, sob o ponto de vista do professor, em construir suas próprias aulas, as quais são apresentadas pelas orientações curriculares, tais como os PCN, pelos programas das disciplinas (saber a ensinar), para adaptá-los à própria classe, aos alunos e efetivar os objetivos buscados. Em outras palavras, a transposição didática consiste em extrair um elemento de saber do seu contexto (universitário, social etc.) para recontextualizá-lo no ambiente sempre singular, sempre único, da própria classe (D’AMORE, 2007, p. 226).

²¹ “[...] des transformations qui dépassent la seule structuration du texte des savoirs pour s’étendre à plusieurs des caractéristiques d’une situation didactique sont inévitablement d’une certaine ampleur. Sont-elles alors viables? A vrai dire, la question se pose déjà pour les modifications du seul texte du savoir. En effet, les contraintes qui pèsent sur la transposition ne se limitent pas à celles qui concernent la proposition d’objets à enseigner. L’élaboration d’une transposition particulière conduisant à la détermination d’objets d’enseignements une fois réalisée, encore faut-il que ces derniers puissent réellement “vivre”, de façon durable, dans le système d’enseignement concret. Il est nécessaire en particulier que l’intégration puisse se faire avec les éléments anciens qui n’ont pas été modifiés.

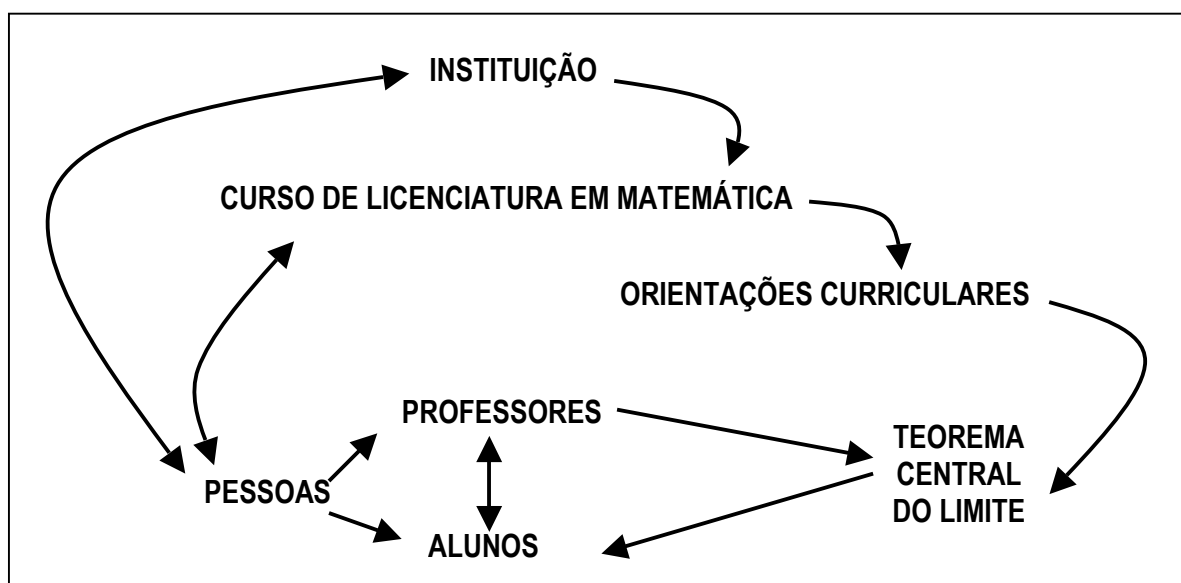
L’étude de la transposition didactique s’étend alors à celle des conditions de la viabilité de tel ou tel objet d’enseignement. Nul doute qu’une connaissance précise de ces conditions serait d’un grand secours pour délimiter le champ des possibles en didactique des sciences, et pour éviter les espoirs déçus et la perte d’énergie correspondante. Une véritable écologie des savoirs s’avère nécessaire, [...]. (JOSHUA; DUPIN, 1993, p. 202-203)

Essa recontextualização do saber, num ambiente próprio, estabelece um conjunto de elementos didáticos que Chevallard (1985) admite que toda atividade humana pode ser descrita como um modelo. Ele considera que o matemático e o didático são inseparáveis, inaugurando, assim, a Teoria Antropológica do Didático .

Na abordagem antropológica, Chevallard (1996, p. 127) afirma que “para começar, são necessários três elementos primitivos (outros virão acrescentar-se-lhes subsequentemente): *os objetos O, as pessoas X, as instituições I.*” Em termos gerais, no sentido *semântico*, o autor considera *objetos* todos os elementos de uma situação didática e, em particular, as pessoas *X* e as instituições *I* também os são. Neste contexto, o Teorema Central do Limite é um objeto (matemático e estatístico), a Instituição, o curso de Licenciatura em Matemática e as pessoas, os alunos e os professores.

[...] *Conhecer um objecto O*, no sentido da teoria apresentada (e não no sentido das diversas instituições que ela deve permitir-nos estudar) é – tanto para uma pessoa como para uma instituição – *ter uma relação com O*. A pessoa *X* (ou a instituição *I*) conhece *O* se existir $R(X,O)$ (respectivamente, $R_I(O)$). Podemos dizer que um objecto existe se for conhecido por pelo menos uma pessoa ou uma instituição (poderá mesmo existir apenas – o que constitui um caso limite – para essa pessoa ou para essa instituição). Um objecto só existe porque é *objecto de conhecimento*. (CHEVALLARD, 1996, p. 128)

Nesses termos, podemos dizer que o autor acaba por traçar um quadro conceitual antropológico do conhecimento ou uma antropologia cognitiva. E, a partir dessas noções, podemos pensar o *objeto* de maneira mais precisa, o real, isto é, o Teorema Central do Limite. A estrutura ecológica do didático pode ser representada pelo esquema 5:



Esquema 6: Representação de um ecossistema sob o ponto de vista do didático.
Fonte: Dados da pesquisa

Observando o esquema 6, podemos dizer que é um ecossistema básico para as relações existentes entre a instituição (I), as pessoas (X_1 e X_2) e o objeto do conhecimento (O), o Teorema Central do Limite. A instituição e as pessoas estabelecem relações interdependentes, pois, por parte do professor (X_1), ele é um componente da instituição; já o aluno (X_2), mesmo mantendo essa característica de interdependência, fez uma ‘escolha’ para que isso ocorresse.

Por outro lado, no âmbito da instituição, existe o curso de Licenciatura em Matemática (I_1) que segue as orientações curriculares (I_2), e tem-se um dos programas da disciplina (I_3), que é intrínseco a ele, pois há o objeto do conhecimento: o Teorema Central do Limite (O). O objeto do conhecimento (O) ‘sofre’ transformações a partir das circunstâncias em que o X_1 elabora suas atividades, seguindo as relações estabelecidas entre as instituições, o objeto do saber (O) e as condições em que se encontra o aluno, X_2 .

Observamos que há interdependência entre os dois ecossistemas — do saber e do didático —, e a “sobrevivência” do teorema está dependente das transformações, ou melhor, das retroalimentações que ocorrem durante todo o processo didático.

Contudo, para ‘fechar’ ao menos essa abordagem, existe um elemento imprescindível nos dois ecossistemas, que é a interrelação entre os sujeitos e o objeto do saber. Então, partimos dos seguintes questionamentos: como os sujeitos

entendem o teorema na aplicação de problemas de situações, hipoteticamente, reais? Que tipo de interpretação eles atribuem aos dados quando os objetos estatísticos partem da construção do conhecimento para a compreensão do teorema? Quais são os argumentos usados pelos estudantes ao avaliar criticamente os dados? Essas são algumas questões que podem ser respondidas ao tratarmos da *literacia* estatística.

5.3 Literacia Estatística

Retomando o que foi exposto no início dessa investigação, optamos pela expressão '*literacia*' ao invés de "letramento" (SOARES, 2005), por entender que estaremos mais próximos da comunidade da Educação Estatística, principalmente no que diz respeito aos últimos eventos internacionais, como, por exemplo, o ICOTS-7, em que esse termo foi significativamente mencionado.

Gal (2002) chama atenção da necessidade de estarmos concentrados na questão da *literacia*, em particular, dos adultos. O autor revela o primeiro passo para entendermos o que vem a ser *literacia* estatística por meio de questões em que ela pode ser definida. A proposta de seu trabalho está aquém do que ocorre na sala de aula, conforme ele mesmo afirma:

[...], o termo '*literacia estatística*' refere-se amplamente a dois componentes correlatos, primeiramente (a) habilidades das pessoas em *interpretar e avaliar criticamente* informação estatística, argumentos de dados relacionados, ou fenômenos estocásticos, os quais podem ser encontrados em vários contextos; e quando destaca (b) habilidades em discutir ou comunicar suas reações para cada informação estatística, tais maneiras compreendem o significado dessas informações, suas opiniões sobre a implicação das mesmas, ou seus pontos de vista no que concerne à aceitação das conclusões apresentadas. (GAL, 2002, p. 2-3) (Tradução nossa).²²

²² "[...], the term "statistical literacy" refers broadly to two interrelated components, primarily (a) people's ability to *interpret and critically evaluate* statistical information, data-related arguments, or stochastic phenomena, which they may encounter in diverse contexts, and when relevant (b) their ability to *discuss or communicate* their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of given conclusions."

A dinâmica da literacia estatística envolve, claramente, dois elementos cognitivos fundamentais para compreensão de como é sua estrutura, sua natureza: o pensamento e o raciocínio estatístico. Porém, Silva (2007, p. 25-26) alerta que, para “além dos elementos cognitivos, há os elementos de disposição: postura crítica, atitudes e crenças”, tão importantes quanto outros. A reunião desses componentes, mais o modelo ativo de questões contextualizadas, levam o indivíduo a desenvolver a *literacia* estatística.

Segundo o mesmo autor (GAL, 2002), para compreender e interpretar dados, também é necessário que o indivíduo se aproprie de habilidades específicas, tais como as de *literacia*, do conhecimento matemático e o conhecimento do contexto. Com relação às habilidades da literacia, ele revê as bases do conhecimento estatístico que se manifestam pelas mensagens produzidas nos textos escritos ou orais, nas informações gráficas ou nas tabelas. Os produtores dessas mensagens possuem características próprias, podendo induzir o leitor ou o ouvinte a inferir sob apenas um ponto de vista. Daí a necessidade de leitura crítica por parte do(a) leitor(a), pois ele, ou ela, não deve se restringir sob um ponto de vista, mas, sim, pensar as diferentes formas que uma determinada informação pode conter, principalmente, nas ‘entrelinhas’ dos dados ou dos textos, ou das tabelas, ou dos gráficos.

Portanto, é relevante o conhecimento básico sobre os objetos estatísticos pertencentes a esse ramo da Ciência, pois é composto por conceitos e procedimentos elementares, além das técnicas matemáticas. Scheaffer, Watkins e Landwehr (1998) pesquisaram diversos currículos de Estatística na Educação Básica e descobriram quais são os principais aspectos de seu ensino que deveriam ser abordados na vida acadêmica. Eles apontaram para os seguintes tópicos dessa área, a serem trabalhados nesse segmento de ensino, o Ensino Superior: senso numérico; compreensão de variáveis; interpretação de tabelas e gráficos; aspectos de planejamento de pesquisa ou experimento; processos de análise de dados; relação entre probabilidade e estatística; raciocínio inferencial.

Esses itens servem de tentativa para acelerar o processo de desenvolvimento da habilidade em *literacia* estatística e dizem respeito à formação do indivíduo para a leitura de mundo, que está impregnado de informações estatísticas.

Garfield e Ben-Zvir (2007) afirmam que o raciocínio estatístico é o modo pelo qual as pessoas raciocinam as ideias estatísticas, produzindo significado para elas.

Essa habilidade pode estabelecer conexões entre um conceito e outro, como, por exemplo, a ideia de centro e dispersão, ou pode também combinar ideias sobre dados e chances, conforme alegam os mesmos autores.

Para Moore (2005, p. xxiii), “a estatística utiliza dados para se adquirir *insight* e para se chegar a conclusões”. Afirma ainda que os “dados são números dentro de um contexto”. Nesse sentido, apresenta o seguinte exemplo:

O número 4,8, por exemplo, não carrega nenhuma informação sozinho. Mas se ouvirmos que o filho recém-nascido de uma amiga pesava 4,8 quilogramas ao nascer, vamos parabenizá-la pelo tamanho saudável do bebê. O contexto situa o nosso conhecimento prévio e nos permite fazer julgamentos. Sabemos que um bebê que pesa 4,8 quilogramas é bem grande e que é improvável que um bebê humano pese 4,8 gramas ou 4,8 toneladas. O contexto faz com que o número seja informativo. (MOORE, 2005, p. xxiii).

Ou seja, o contexto é um componente fundamental do pensamento estatístico. Mas, além dos contextos nos quais os dados estatísticos estão inseridos, Wild e Pfannkuch (1999) defendem o fato de que esse tipo de pensamento é similar ao de um indivíduo envolvido numa indagação ou em processo de pesquisa. Assim, em situações de pesquisa, quando o sujeito se envolve em algum tipo de questionamento – neste caso, necessariamente exige um contexto –, estará pensando estatisticamente.

Os mesmos autores ainda alegam que “[...] uma solução para um problema real baseado no conhecimento requer melhor compreensão de como se trabalha um sistema e, talvez também, como reagirá às mudanças, às concepções iniciais, aos posicionamentos de seu entorno”. (WILD; PFANNKUCH, 1999, p. 225) (Tradução nossa)²³. Isso significa que é necessário conhecer a realidade que nos cerca e, daí, participar também como sujeitos de transformação, porque, diante do posicionamento de cada um, há o desenvolvimento da *literacia* estatística.

Garfield e Ben-Zvi (2007) sintetizam a distinção entre *literacia*, raciocínio e pensamento estatístico dizendo que, apesar de pertencerem a uma mesma área, existe certa hierarquia que se constitui a partir da *literacia* como base para o raciocínio e o pensamento estatístico. E, além disso, eles sublinham que o raciocínio estatístico significa compreender, ser capaz de explicar processos estatísticos e

²³ “[...]. A knowledge-based solution to the real problem requires better understanding of how a system works and perhaps also how it will react to changes to input streams, settings or environment.” (WILD; PFANNKUCH, 1999, p. 225)

interpretar os resultados, isto é, seria uma forma de como desencadear os modelos mentais que estão estruturados para cada situação. Já o pensamento seria a forma como irá estruturar a situação em si, o que inclui o como e o porquê usar, particularmente, um método, uma medida, um esboço ou modelo estatístico. E, por fim, a *literacia* constitui-se na união entre o raciocínio e o pensamento estatístico aliados ao fato de que o sujeito precisa de pensar, raciocionar criticamente para agir e, talvez, transformar a realidade concreta de seu entorno.

As três dimensões até aqui apresentadas, a abordagem ecológica do saber, a ecologia do didático e a *literacia* estatística, nos propiciaram estabelecer o nosso quadro teórico com a finalidade de fundamentar essa investigação.

E, diante da problemática ecológica proposta por Artaud (1998), neste caso, o Teorema Central do Limite sob o ponto de vista ecológico do saber e do didático, partimos do pressuposto de que seus fundamentos nos permitirão subsídios para analisarmos as atividades propostas aos alunos. O primeiro é o saber em si; saber a ser ensinado, que sofre transformações por parte do professor em situações, como o estudo do teorema em diferentes livros-texto e sua escolha, tendo em mente os sujeitos que estarão predispostos a ‘receber’ tal saber, no caso, o teorema. Uma de suas tarefas é o de transformar o saber científico em saber a ser ensinado, mas, até aí, nada se pode afirmar que este saber será aprendido pelo aluno. A estrutura ecológica que abarca o teorema nos dá a ideia de como o teorema ‘sobrevive’ nos dias de hoje, particularmente, nos cursos de Licenciatura em Matemática e, de certa forma, constatamos que a demonstração não é privilegiada pela maioria dos livros-texto consultados.

O segundo quadro teórico, a Teoria Antropológica do Didático, nos reporta a entender a estrutura interna que nos permite conhecer o ‘como’ ocorre a sequenciação e a temporalização didática, constituindo-se também de um ecossistema do didático. E, por fim, o terceiro, a *literacia* estatística, nos apresenta o contorno didático para análise de nossa investigação, pois, quando ela recebe um tratamento de interrelação com os dois primeiros quadros teóricos, reforça a relevância de nossa investigação no âmbito individual e social, consolidando, assim, a metodologia de nossa pesquisa, o *design didático*.

O próximo capítulo será a pesquisa em si, devidamente fundamentada pelo quadro teórico exposto, além da apresentação de algumas atividades realizadas pelos licenciandos que, possivelmente, atuarão como professores da Educação

Básica no futuro não muito distante. Deixamos claro que, em momento algum, o nosso trabalho esteve voltado para uma pesquisa quantitativa, mesmo precisando de usar meios que requerem dados quantitativos para compreensão dos objetos estatísticos diretamente associados ao Teorema Central do Limite. Portanto, vale reforçar que o cunho de nossa investigação é qualitativa.

6 A PESQUISA

Vamos retomar a nossa questão de pesquisa:

Por que determinados saberes/conhecimentos deixaram de ser abordados no ensino do Teorema Central do Limite, mas ainda assim ele “sobrevive”? Que saberes são indispensáveis para o teorema “viver”? Que relações diretas ou indiretas existem entre o teorema e a construção de saberes estatísticos nas aulas de Estatística?

Para tentar responder a essas questões, seguindo a metodologia adotada, dividimos este capítulo em três partes, com a opção de atender às três questões expostas. A primeira refere-se ao estudo dos livros-texto analisados, somente nos itens que dizem respeito ao teorema, apresentação e análise praxeológica de alguns problemas propostos neles e a apresentação de como o teorema é definido em cada livro. Partindo dessa dinâmica, parece-nos bastante oportuno abordarmos, ao menos em parte, **os saberes/conhecimentos que deixaram de ser abordados no ensino do Teorema Central do Limite, mas que, ainda assim, lhe permitem sobreviver**”, já que é fundamental para Inferência Estatística.

A segunda parte corresponde às atividades realizadas com os licenciandos, expondo a forma como foi conduzida e os diferentes tipos de abordagem que cada um dos quatro sujeitos da pesquisa apresentou em suas respostas. Essas atividades foram apresentadas de tal forma que fossem sustentadas pelo nosso quadro teórico. À luz do *design didático*, reconhecemos os tipos **saberes que são indispensáveis para o teorema “viver”**, pois os estudantes atingiram a compreensão do teorema e de sua definição na construção de ideias e conceitos estatísticos.

E, por fim, a terceira parte constituiu-se de um diálogo realizado entre o professor-pesquisador e os licenciandos, cujos procedimentos metodológicos seguiram as intervenções controladas e devidamente refletidas entre o grupo. Dessa forma, nos foi possível, efetivamente, atender a terceira questão ecológica de nossa investigação: **estabelecer relações diretas ou indiretas que existem entre o teorema e a construção de saberes estatísticos nas aulas de Estatística**.

6.1 O Teorema Central do Limite e os Livros-Texto

A relevância da análise dos livros-texto quanto à abordagem do Teorema Central do Limite é apontado por Almouloud (2007), que afirma:

A praxeologia associada a um saber é a junção de dois blocos: saber-fazer (técnico/prático) e saber (tecnológico/teórico), cuja ecologia refere-se às condições de sua construção e vida nas instituições de ensino que a produzem, utilizam ou transpõem. (ALMOULOU, 2007, p. 123)

Chevallard (1999) define as praxeologias (ou organizações) em duas espécies: as matemáticas e as didáticas. A primeira refere-se ao saber matemático na realidade matemática, podendo ser construída para ser desenvolvida na sala de aula; já a segunda diz respeito ao modo como se realiza essa construção. Portanto, baseados nessa justificativa, optamos por analisar alguns livros-texto, seguindo os pressupostos já apresentados no capítulo referente aos procedimentos metodológicos.

Denominamos A1 referente aos autores 1, A2, à 2 e assim por diante até a 7ª obra, conforme o quadro abaixo:

DENOMINAÇÃO	AUTORES
A1	Bussab e Morettin (2003)
A2	Farias, Soares e César (2003)
A3	Guimarães e Cabral (1997)
A4	Levin e Fox (2004)
A5	Moore (2005)
A6	Stevenson (2001)
A7	Triola (1999)

Quadro 1: Denominação dos Livros-Texto.

Fonte: Dados da pesquisa.

Para apresentação desse estudo, vale, preliminarmente, esclarecermos alguns itens sobre o Teorema Central do Limite, pois, nos livros-texto consultados, deparamo-nos com a seguinte questão: dizemos “Teorema do Limite Central” ou “Teorema Central do Limite”? Ambas as expressões são conhecidas, ao menos na Língua Portuguesa, mas as obras divergem quanto a esse tratamento. Farias,

Soares e César (2003, p. 140) consideram que “alguns textos referem-se erroneamente ao Teorema do Limite Central, pois o que é central é o teorema, e não o limite”. Assim, optamos por Teorema Central do Limite ao invés de Teorema do Limite Central. Outro item a ser destacado é com relação à estrutura dessas obras, já que nenhuma delas demonstra o teorema efetivamente, com exceção de A3, que apresenta algumas demonstrações mais próximas do que vem a ser o Teorema Central do Limite. Além disso, a maioria se limita primeiramente a defini-lo para, em seguida, apresentar alguns exemplos e, por fim, algumas aplicações.

6.1.1 Análise Quantitativa dos Livros-Texto

Anterior à análise das atividades dos livros-texto, vale apresentar a quantidade de atividades existentes em cada livro e as características dessas atividades. Assim, nos foi possível constatar atividades com contexto matemático, exigindo aplicação de técnicas, e outros com contexto do cotidiano, mesmo que hipoteticamente reais. Além disso, incluímos nesse rol alguns exemplos e exercícios sobre distribuição amostral de médias e de proporção, tendo em vista que abordam também o teorema. Não foram incluídos exercícios complementares, pois continham atividades mais genéricas, cuja aplicação não estava diretamente associada ao teorema. Embora isso possa consistir em uma articulação implícita, optamos por não abordá-la porque não é o foco de nosso estudo.

TABELA 1
Atividades nos Livros-Texto

Denominação	Quantidade de atividades de aplicação do Teorema Central Do Limite		Total
	Contexto matemático	Contexto cotidiano	
A1	2	8	10
A2	0	0	0
A3	4	3	7
A4	0	0	0
A5	0	5	5
A6	10	12	22
A7	0	23	23

Fonte: Dados da pesquisa.

A partir da tabela 1, podemos observar que A2 e A4 não apresentam atividades específicas para o Teorema Central do Limite, apesar de apresentarem de forma implícita sua aplicação. Eles distinguem-se pelo fato de A2 apresentar definição e a representação gráfica de como “ocorre” o teorema, isto é, conforme se aumenta o tamanho das amostras de médias amostrais, mas não constatamos nenhum exemplo específico de aplicação para o teorema. Em contrapartida, A4 não menciona sequer o teorema, porém deixa implícita a utilização do teorema em outros problemas, como no tópico “estimação” ao tratar-se de “intervalos de confiança”. Porém, como o nosso interesse está diretamente relacionado ao teorema, optamos por não apresentar esse tipo de atividade, embora tais noções também façam parte de um ecossistema que contém o Teorema Central do Limite.

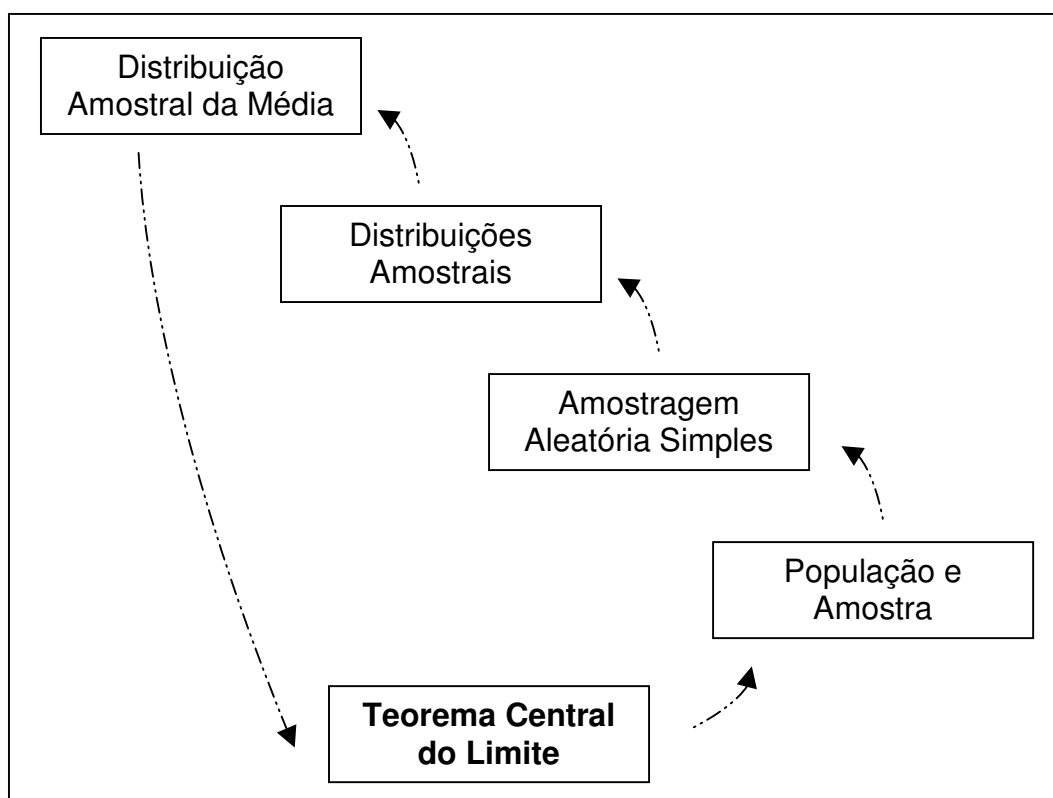
6.1.2 Apresentação e Análise de Atividade no Livro-Texto

De acordo com a Teoria Antropológica do Didático (CHEVALLARD, 1999), as noções de *tarefa*, *técnica*, *tecnologia* e *teoria* permitem que as práticas sejam modeladas de forma a dar contorno às atividades em geral. Segundo Silva (2007, p. 68), “para a análise de uma *praxis*, o saber-fazer, a teoria revela o estudo das tarefas propostas pelos livros bem como as técnicas apresentadas para a solução das tarefas”, e, posteriormente, obtém-se o discurso teórico-tecnológico, justificando a *praxis*, ao analisar o *logos*. Essa dinâmica, associada ao enfoque ecológico, será a base de nossa reflexão das atividades selecionadas.

Ao selecionar as atividades, optamos por escolher duas atividades dos livros-texto que contenham a aplicação do teorema, sendo um com contexto matemático e outro, contexto cotidiano.

6.1.2.1 Atividades do livro-texto A1

A estrutura ecológica de A1, para o Teorema Central do Limite, parte do ramo da Inferência Estatística. Podemos descrevê-la da seguinte forma:



Esquema 7 –Estrutura ecológica para o Teorema Central do Limite no livro-texto A1.
 Fonte: Dados da pesquisa.

Em linhas gerais, o livro-texto A1 parte da definição de população e amostra, seguindo para amostragens aleatórias simples. Depois, aborda as distribuições amostrais para, enfim, tratar das distribuições amostrais para médias. Além disso, expõe graficamente a variação das distribuições para médias de amostras de números diferentes. Vejamos as duas atividades apresentadas por A1:

- **Contexto Matemático**

A atividade no contexto matemático encontra-se na parte de “problemas” (BUSSAB; MORETTIN, 2003, p. 274), conforme segue o enunciado:

Uma v.a. X tem distribuição normal, com média 100 e desvio padrão 10.

(a) Qual a $P(90 < X < 110)$?

(b) Se $\bar{X}$ for a média de uma amostra de 16 elementos retirados dessa população, calcule $P(90 < \bar{X} < 110)$?

(c) Represente, num único gráfico, as distribuições de X e $\bar{X}$.

(d) Que tamanho deveria ter a amostra para que $P(90 < \bar{X} < 110) = 0,95$?

Na Teoria Antropológica do Didático, chamamos cada item dessa atividade de tarefas, pois pressupõe um verbo, uma ação a ser realizada, portanto a tarefa (T) (a) é: pede-se 'qual a $P(90 < X < 110)$ ', ou seja, os estudantes que estão familiarizados com as simbologias e o tema interpretariam que é para calcular (T) a probabilidade de que as variáveis se encontrem entre 90 e 110, de uma distribuição normal, quando a média da população é 100 e o desvio padrão, 10, cuja representação é $N(100, 100)$, isto é, média 100 e variância 100. A utilização das técnicas (τ) para execução da tarefa (T) pressupõe que sejam encontrados os valores reduzidos, z , padrão, já que a distribuição é normal:

$$z_1 = \frac{x_i - \mu}{\sigma} = \frac{90 - 100}{10} = -1 \quad (1)$$

$$z_2 = \frac{x_i - \mu}{\sigma} = \frac{110 - 100}{10} = 1 \quad (2)$$

Isso se justifica porque uma distribuição normal é similar à distribuição normal reduzida, pois a distribuição real pode ser reduzida para o caso de $Z \sim N(0, 1)$, isto é, média zero e variância 1; por isso há necessidade de calcular os valores de z (escores reduzidos). Essa parte corresponde à tecnologia (θ), já que estamos justificando a técnica. Por isso,

$$P(90 < X < 110) = P(-1 < z < 1) = 0,6826, \quad (3)$$

conforme a tabela da distribuição normal padrão. A resposta para a tarefa (T) do item (a) é 0,6826. Para justificar a tecnologia, culminaria na teoria (Θ), conforme toda a descrição feita até então. Porém, vale ressaltar que essa tarefa constitui-se de

contexto matemático, sem aplicação no cotidiano, é uma atividade pontual, sem a possibilidade de estender para uma análise regional.

No item (b) a tarefa é similar ao do item (a), porém cria-se uma situação em que a média, 100, não é mais de uma população, percebe-se uma sequência de tarefas que, gradativamente, vão construindo a ideia do Teorema Central do Limite. Portanto, resolvendo o item (b), tem-se que o desvio padrão da amostra de 16 elementos tende a ser menor que o da população:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{16}} = 2,5 \quad (4)$$

E os escores reduzidos serão:

$$z_1 = \frac{x_i - \bar{x}}{\sigma_{\bar{x}}} = \frac{90 - 100}{2,5} = -4 \text{ e } z_2 = \frac{x_i - \bar{x}}{\sigma_{\bar{x}}} = \frac{110 - 100}{2,5} = 4 \quad (5)$$

Os escores são aparentemente números que tendem à probabilidade compreendida entre -4 e 4 , é muito próximo de 1,00. A praxeologia dessa atividade é similar ao do (a).

Para realização do item (c), devemos partir da representação gráfica. Assim, podemos construir para X e $\bar{X}$:

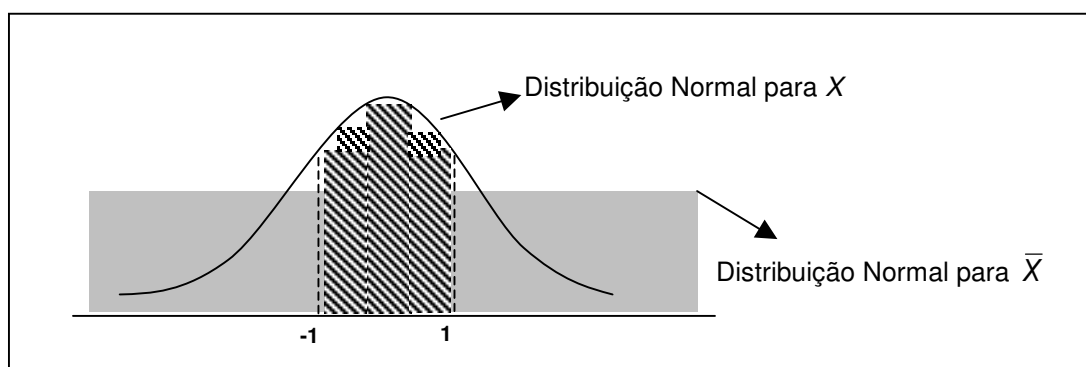


Figura 2 – O gráfico para distribuição normal para X e $\bar{X}$.
Fonte: Dados da pesquisa.

Conforme podemos observar, o gráfico para a distribuição normal tem a probabilidade 0,6826, cuja área está representada pelo hachurado. Para a distribuição da amostra de 16 elementos da população, a área do retângulo não se constitui como curva, destacando que sua probabilidade é 1,00.

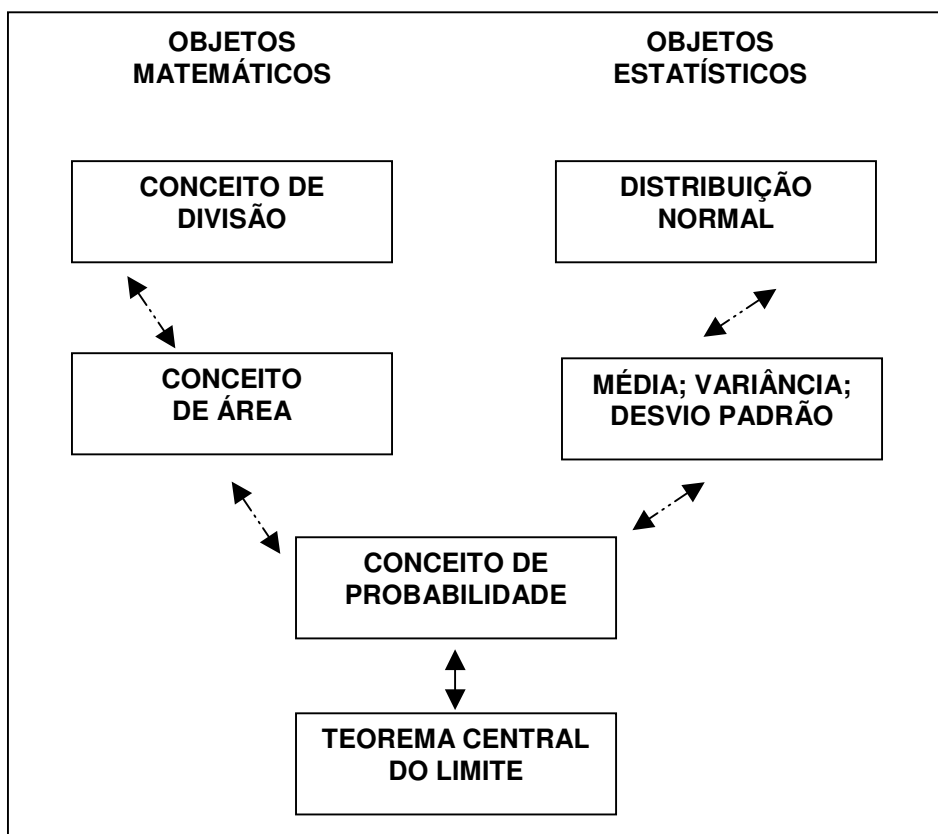
Para o item (d), a sequência das tarefas exige o raciocínio matemático reverso. Isto é, dada a probabilidade, pede-se para encontrar (tarefa) o tamanho da

população. Seja $P(90 < \bar{X} < 110) = 0,95$, sendo que a probabilidade é de uma amostra. Assim, podemos utilizar a técnica de encontrar os valores reduzidos, ou os escores reduzidos z , quando a probabilidade é 0,95. Nesse caso, o $z = \pm 1,96$, a partir da tabela da Distribuição Normal Padrão. E, para calcular o tamanho da amostra, tem-se que:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \therefore -1,96 = \frac{90 - 100}{\frac{10}{\sqrt{n}}} \therefore n = 4.$$

(6)

Como a aplicação para a verificação do Teorema Central do Limite pode se efetivar como também pode causar confrontos entre a aplicação e os conceitos espontâneos que poderiam surgir por parte dos alunos, a sequência de tarefas proposta privilegia, efetivamente, os conceitos e ideias mais técnicas abstratas para o teorema. No entanto, reconhecemos a aplicação do contexto matemático, exigindo habilidades que requerem conceitos específicos já adquiridos. Numa abordagem ecológica, podemos elaborar um ecossistema que requer objetos matemáticos e estatísticos:



Esquema 8 – Ecossistema da atividade proposta por A1 no contexto matemático e estatístico.

Fonte: Dados da pesquisa.

Na abordagem ecológica do saber, constatamos que o teorema atende a dois ecossistemas simultaneamente, observando os saberes imprescindíveis para que o saber “sobreviva”.

- **Contexto Cotidiano**

O enunciado para um contexto cotidiano em que se aplica o Teorema Central do Limite nos permitiu escolher o seguinte:

A capacidade máxima de um elevador é de 500 kg. Se a distribuição X dos pesos dos usuários for suposta $N(70, 100)$: (a) Qual a probabilidade de sete passageiros ultrapassarem esse limite? (b) E seis passageiros? (FARIAS; SOARES; CÉSAR, 2003, p. 275)

Para resolver essa tarefa (T), que também é a de calcular a probabilidade, no caso do item (a), as técnicas (τ), que também podem se chamar os modos de *fazer*, requerem os mesmos procedimentos que a atividade anterior, porém traz um contexto consigo, que é a *capacidade máxima que um elevador pode suportar*. No entanto, para o item (a), deseja-se verificar a probabilidade de sete passageiros ultrapassarem o limite de 500 kg. Portanto, deve-se calcular a probabilidade para $P(X \geq 500)$.

Vejamos a atividade acima: o contexto apresenta um elevador que suporta no máximo 500 kg. Para interpretar a distribuição da variável aleatória X , dos pesos dos usuários, tem-se que os pesos dos indivíduos são independentes. Além disso, a representação $N(70, 100)$ deve ser algo já conhecido pelo estudante, já que se estima que a média da amostra é 70 e a variância 100 de uma distribuição normal.

Então, a técnica utilizada seria, em primeiro lugar, encontrar o “peso” total dos 7 indivíduos e, como a média amostral de um indivíduo é 70 kg, tem-se que a média amostral para os sete indivíduos seria aproximadamente de $\bar{X} = 7 \times 70 = 490$. Deseja-se, portanto, calcular a área que corresponde $P(X > 500)$. E, para encontrar o valor reduzido de $X = 500$, tem-se que:

$$z = \frac{500 - 490}{10} = 1.$$

(7)

A área normal reduzida equivalente a probabilidade superior a 500 kg será:

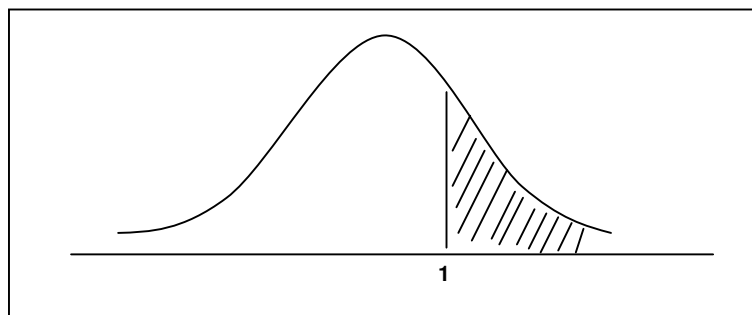


Figura 3 - A distribuição de observação.
Fonte: Dados da pesquisa.

Assim, o modelo real pode ser reduzida para o modelo padrão:

$$P(X > 500) = P(z > 1) = 0,5 - 0,3413 = 0,1587 \text{ ou } 15,87\%.$$

(8)

Portanto, encontra-se o valor aproximado do escore reduzido para 7 pessoas, cuja probabilidade é de 15,87%, respondendo ao item (a).

Essa atividade envolve os dois tipos de organização praxeológica: a local, que é específica na aplicação da técnica para realização da tarefa, justificado e explicado por uma tecnologia, cuja justificação é o próprio teorema; e global, pois envolve outros conceitos matemático e estatístico (como, por exemplo, o conceito de divisão, média, variância, desvio padrão, função da curva normal e a teoria das probabilidades). O ecossistema que pode representar essa atividade é bastante similar ao anterior, porém destacamos o contexto pelo qual foi apresentado. O que significa que, além das técnicas que envolvem objetos matemáticos e estatísticos, existe também a parte reflexiva sobre o resultado em si. Isto é, a probabilidade, 15,87%, de sete pessoas ultrapassarem o limite de peso do elevador é significativa, já que há variação entre os pesos (massas) das pessoas.

Mesmo uma situação hipotética, mas que poderia ser real, permite ao aluno imaginar que decisão tomaria se deparasse com esse tipo problema ou até como analogia para outras situações. Isso torna mais interessante a aplicação do objeto do saber: o Teorema Central do Limite.

6.1.2.2 Atividades do livro-texto A2

Nessa obra, os autores apresentam o comportamento das curvas normais nos gráficos, de acordo com o número de observações de amostrais extraídas de uma população, sem mencionar contexto cotidiano para tal, conforme as figuras abaixo:

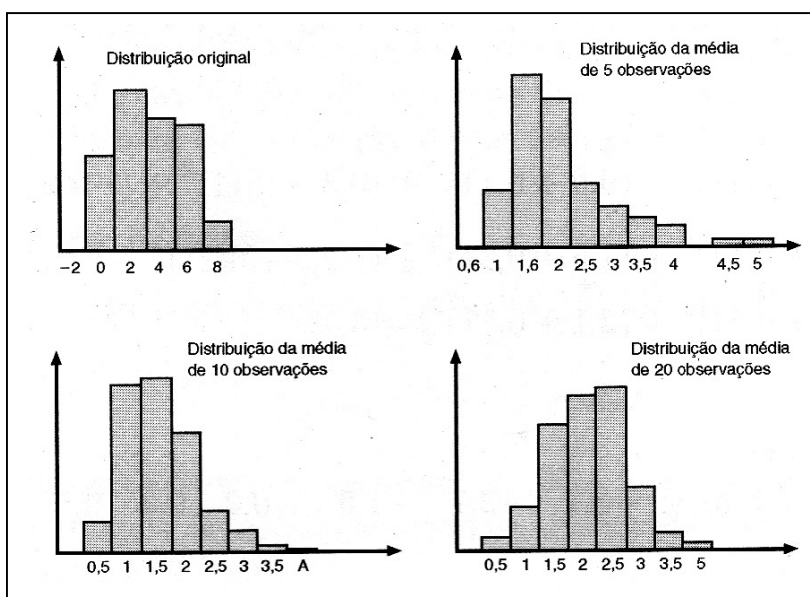


Figura 4 – Distribuição uniforme com 300 repetições $\alpha = 2,9$ e $\beta = 5,39$
Fonte: FARIAS; SOARES; CÉSAR, 2003, p. 140.

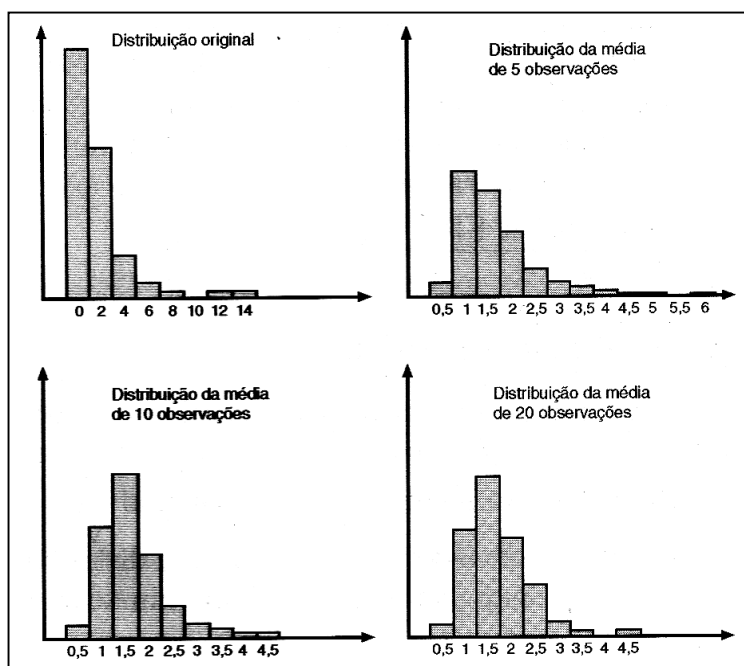


Figura 5 – Distribuição logonormal com 300 repetições $\alpha = 0$ e $\beta = 1$
Fonte: FARIAS; SOARES; CÉSAR, 2003, p. 140.

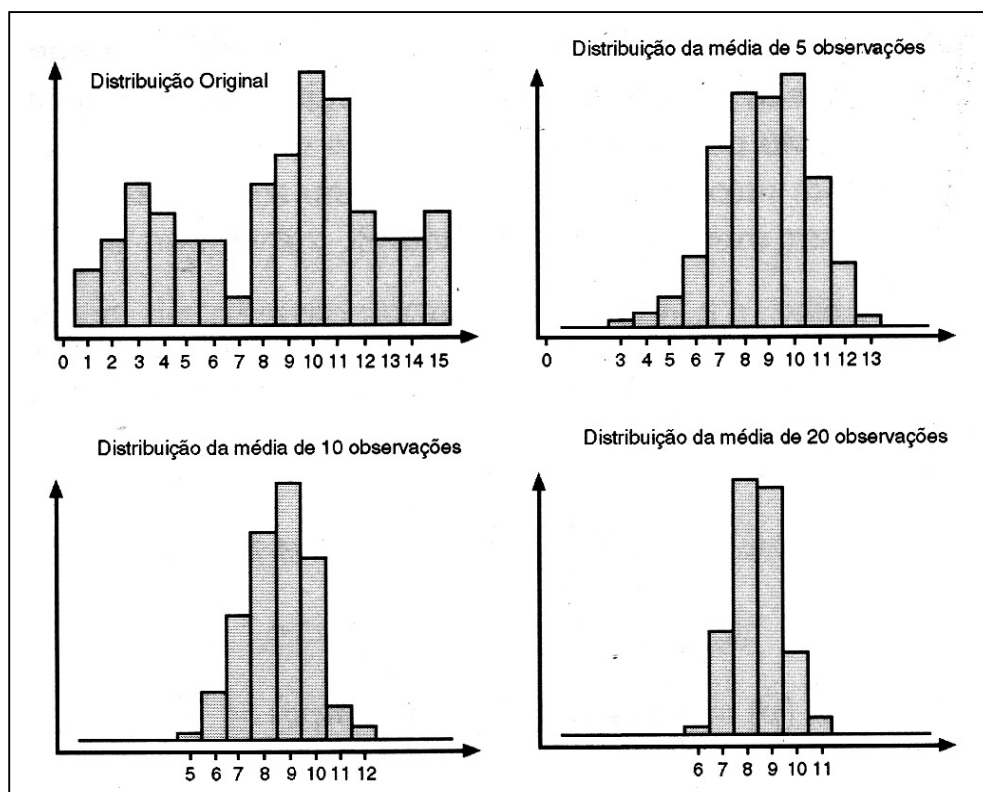


Figura 6 – Distribuição bimodal com 300 repetições
 Fonte: FARIAS; SOARES; CÉSAR, 2003, p. 141.

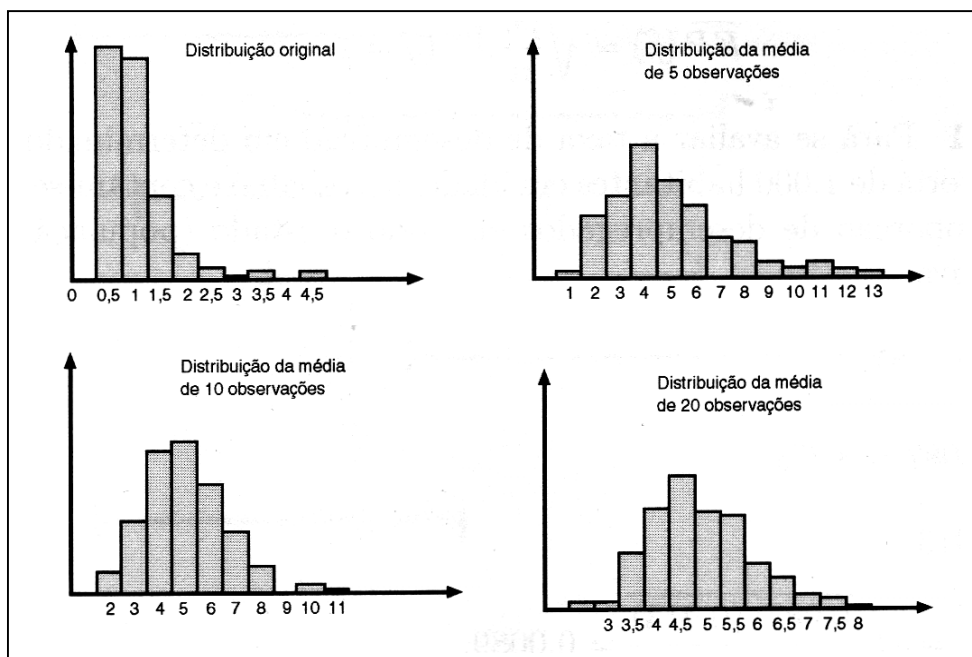


Figura 7 – Distribuição exponencial com 300 repetições
 Fonte: FARIAS; SOARES; CÉSAR, 2003, p. 141.

Os autores apresentam quatro tipos de distribuições: uniforme, logonormal, bimodal e exponencial, de modo que o leitor possa compreender que, de acordo com o número de observações, a distribuição da $\bar{X}$, média da amostra, estará centrada na média da população, μ , não importando que tipo é a distribuição, ou seja, simétrica ou não. Além disso, não apresentam nenhuma aplicação direta sobre o Teorema Central do Limite, impossibilitando-nos uma análise ecológica do saber com mais afinco, pois não há aprofundamento sobre o tema. A aplicação será subjacente a outros temas da Inferência Estatística.

6.1.2.3 Atividades do livro-texto A3

Os autores desse livro-texto exploram, além do conceito do Teorema Central do Limite, a utilização de softwares para gerar números aleatórios, no sentido de enfatizar os diferentes tipos de distribuição, simétricas ou não, dependendo do número de amostras. Vale a aplicação do teorema na mesma proporção em que se aumenta o número dessas amostras.

- **Contexto matemático**

O enunciado é o seguinte:

Na figura 8.1 apresenta-se a distribuição da variável Y , tal como se encontra definida na tabela, bme como as distribuições da média amostral $\bar{Y}$ (para $N = 5, 10, 20, 50, 100$ e admitindo que as amostras são aleatórias simples)

Interessante notar que os autores designam o N , em maiúsculo, para o número de amostras, fato este não encontrado em outros livros-texto da nossa investigação. Vejamos a figura representada pela tabela:

Tabela 8.7. Distribuição da variável Y (Exemplo 8.7)	
y	$p(y)$
0	0.80
1	0.20

Figura 8 – Tabela da distribuição da variável Y
Fonte: GUIMARÃES; CABRAL, 1997, p. 240.

Quanto aos gráficos, são apresentados conforme a figura abaixo:

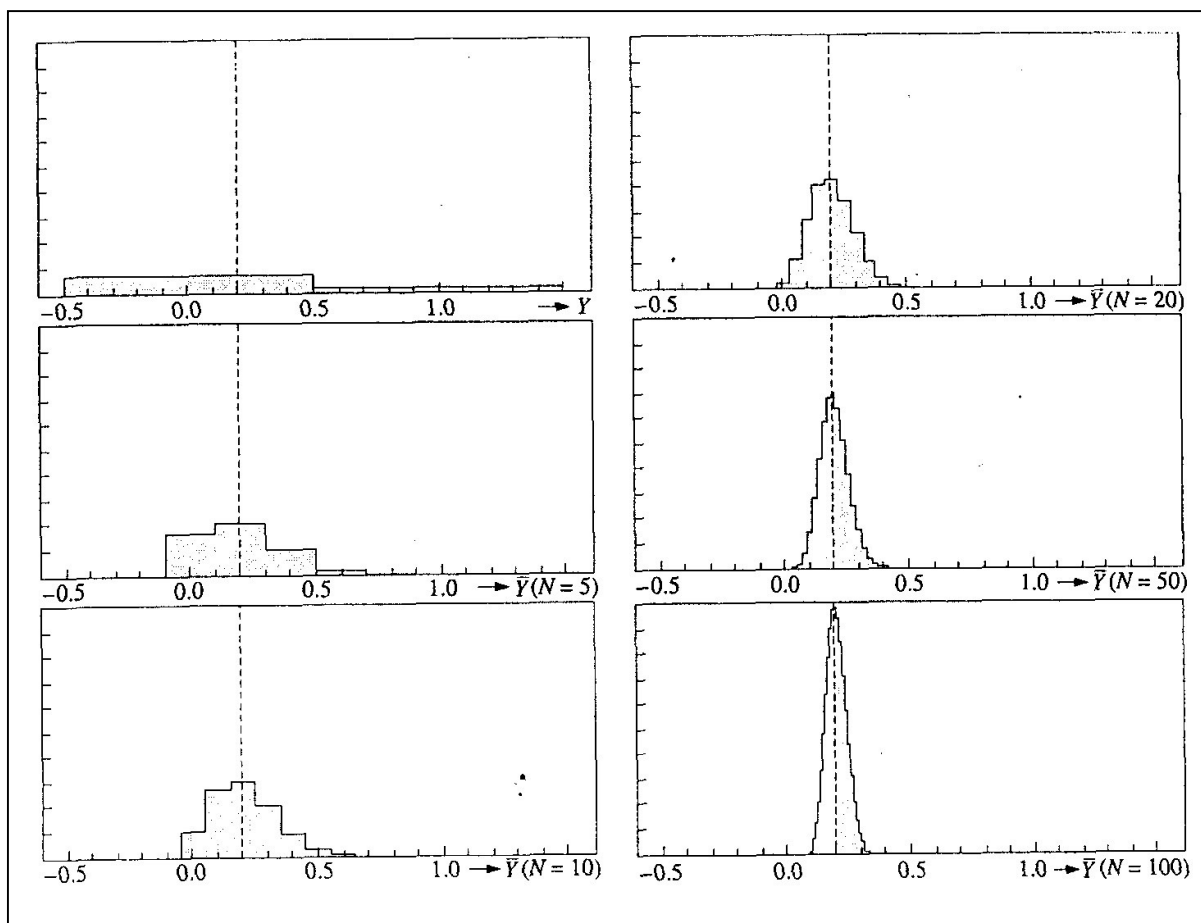


Figura 9 – Distribuição de Y e $\bar{Y}$
Fonte: GUIMARÃES; CABRAL, 1997, p. 241.

Os autores fazem os comentários sobre os objetos estatísticos sem mencionar no contexto cotidiano. Entretanto, posteriormente, citam que é um exemplo para o caso de uma situação real. Alguns comentários que eles fazem sobre o exemplo merecem atenção: apresentam os histogramas, em que “a barra

correspondente a $Y = 0$ cobre os valores de $Y = -0,5$ até $Y = 0,5$ ". Por outro lado, "enquanto que a barra correspondente a $Y = 1$ cobre os valores de $Y = 0,5$ até $Y = 1,5$."; alerta para às ordenadas dos diagramas, pois a intenção era de que as áreas dos diferentes diagramas fossem todas idênticas, à altura de cada barra foi atribuído o valor que resultou de dividir a função probabilidade pela largura da barra; e, por último, alegam que a distribuição Y é discreta e também é uma distribuição de média amostral e afirmam que a distribuição normal é uma boa aproximação da distribuição de média amostral para $n \geq 50$.

Tal qual o exemplo do livro-texto anterior, A2, os autores alegam que, se n é suficientemente "grande", portanto, pode-se dizer uma "boa aproximação normal". Da forma como foi exposta neste exemplo, essa alegação nos restringe em relação à análise ecológica do saber. Portanto, vejamos a próxima atividade de A3, que é num contexto cotidiano.

- **Contexto cotidiano**

Vejamos o enunciado exposto pelos autores:

Um elevador de acesso a um grupo de galerias de uma mina tem capacidade nominal de 3800 kg. Os mineiros que usam regularmente o elevador são 650 e têm pesos que seguem uma distribuição com valor esperado de 75 kg e desvio padrão de 8 kg. Calcule a probabilidade de ser excedida a capacidade nominal do elevador quando nele se encontrarem 50 mineiros. (GUIMARÃES; CABRAL, 1997, p. 261)

A atividade selecionada de A3 é similar à A1, o que dispensa sua análise. As diferenças estão na quantidade que suportam, em peso (massa), o elevador e o local em que ocorre o contexto. A escolha desse problema foi também intencional, pois chamamos atenção para os tipos de contextos que as obras apresentam, isto é, algumas vezes, as ideias e os problemas são basicamente os mesmos.

6.1.2.4 Atividades do livro-texto A4

Levin e Fox (2004) não apresentam a definição do Teorema Central do Limite nem exemplos ou exercícios de aplicação diretas, tanto para o contexto matemático como para o contexto cotidiano. Entretanto, a partir da definição das distribuições amostrais para as médias na forma de curva normal, encontramos indícios para que as ideias do teorema possam surgir sem necessariamente mencioná-lo.

Mais adiante, ao apresentarmos a definição que cada livro-texto atribui ao Teorema Central do Limite, analisaremos com mais detalhes a ausência explícita do teorema nessa obra. Portanto, a escolha dela também foi intencional, tendo em vista que são autores renomados no campo da Estatística e que possuem várias obras nesse ramo da Ciência.

6.1.2.5 Atividades do livro-texto A5

O livro-texto A5 não apresenta atividade com contexto matemático, no entanto, apesar de apresentar apenas cinco atividades com aplicação do teorema, percebe-se que explora o tema ao longo do texto em sua obra, enfatizando sempre que é pelo “efeito” do Teorema Central do Limite que se pode aplicar determinadas situações-problema. Por isso, optamos a escolha de uma atividade para análise, pois irá confrontar com alguns conceitos da Estatística. Vejamos:

O tempo X que um técnico necessita para fazer manutenção preventiva em um aparelho de ar-condicionado é regido pela distribuição exponencial, cuja curva de densidade aparece na figura 5. O tempo médio é $\mu = 1$ hora e o desvio padrão é $\sigma = 1$ hora. Sua empresa utiliza 70 dessas unidades. Qual é a probabilidade de que seu tempo médio de manutenção exceda 50 minutos? (MOORE, 2005, p. 224-225)

A figura 10 é a seguinte:

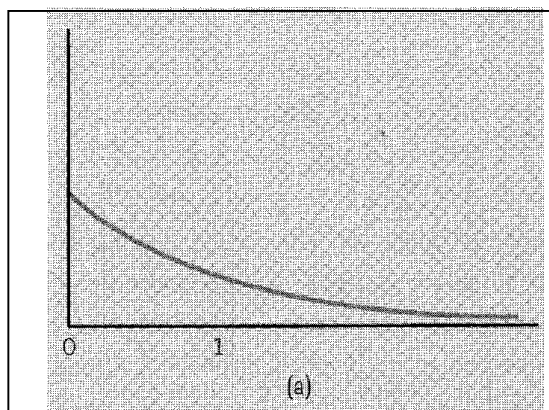


Figura 10: A distribuição de observação.
Fonte: MOORE, 2005, p. 224.

Moore (2005, p. 225) expõe inicialmente o seguinte: “O teorema central do limite afirma que a média amostral $\bar{x}$ (em horas) gastas trabalhando em 70 unidades tem aproximadamente a distribuição normal com média igual à média populacional $\mu = 1$ hora de desvio padrão”:

$$\frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{70}} \cong 0,12 \text{ hora}$$

(9)

Quando o autor apresenta esse resultado, não fica claro o que significa “0,12 hora”, mas é o desvio padrão da amostra de 70 unidades. Como exemplo resolvido, entendemos que, apesar do destaque ao teorema, talvez exista ênfase para a técnica, sem tecnologia. Logo em seguida diz: “a distribuição $\bar{x}$ é, portanto, aproximadamente $N(1; 0,12)$ ”. E, neste caso, difere da representação já apresentada anteriormente pelo A1, em que $N(\mu, \sigma^2)$ significa a representação da média e da variância, enquanto para A5 é a média e o desvio padrão. Em seguida apresenta a curva Normal de linha contínua e a tracejada a distribuição exata, conforme a figura 11:

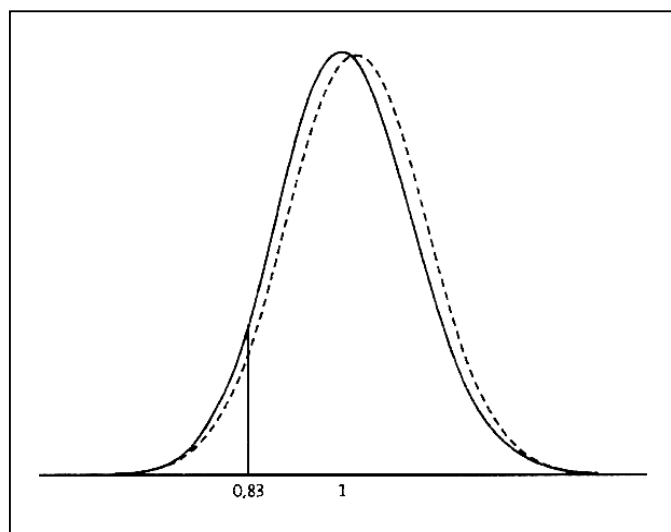


Figura 11 – A distribuição exata (tracejada) e a aproximação Normal a partir do Teorema Central do Limite (contínua) para o tempo médio necessário para manutenção de um ar-condicionado.
Fonte: MOORE, 2005, p. 225.

E o autor justifica: “como 50 minutos são 50/60 de uma hora, a probabilidade que queremos é $P(\bar{x} > 0,83)$. Um cálculo de distribuição Normal fornece que essa probabilidade é 0,9222. Essa é a área à direita de 0,83 sob a curva Normal contínua” na Figura 11.

O autor explica a técnica (τ) para a tarefa (T), mas não justifica como encontrou a área que corresponde à região desejada, quando menciona “um cálculo de distribuição normal”, resultando na probabilidade 0,9222. Assim, comprometeu a tecnologia (θ). Em seguida, ele esclarece que, utilizando “mais matemática, poderíamos, a partir distribuição exponencial, achar a curva de densidade real de $\bar{x}$ para 70 observações.” Depois explica que a curva é a que está tracejada conforme a figura 11, alegando que se pode observar que a curva Normal contínua é uma boa aproximação. Portanto, “a aproximação Normal do teorema central do limite está apenas afastada por cerca de 0,07”.

Observamos que, apesar de o autor não priorizar as técnicas, existe uma estrutura ecológica para a resolução proposta por Moore nessa atividade. Quando afirma que “a curva Normal contínua é uma boa aproximação”, efetivamente, ele opera a tecnologia da tecnologia, isto é, apresenta a teoria (Θ) da tarefa (T). Outro item interessante a ser ressaltado, neste caso, é de que não há uma organização praxeológica local, mas, sim, um conjunto de organizações praxeológicas regionais,

consistindo-se, assim, numa organização praxeológica global. Isso é justificado pela série de técnicas (identificar as subunidades da hora e encontrar o valor do desvio padrão da amostra para 70 unidades, por exemplo), tecnologias (as explicações e justificativas em relação às técnicas utilizadas) e teoria (justificativa da justificativa) ao realizar a tarefa.

6.1.2.6 Atividades do livro-texto A6

Este livro-texto foi selecionado a partir dos critérios expostos nos procedimentos metodológicos, porém um detalhe deve ser esclarecido. Apesar de o ano da edição não ser recente, verificamos alguns itens interessantes a comentar sobre o teorema, além do caráter eclético da obra, pois seu título é específico para curso de Administração.

Um dos pontos interessantes notados por nós foi uma das afirmações do autor: “em sentido estrito, o *Teorema Central do Limite* só se aplica a médias amostrais.” (STEVENSON, 1981, p.182) (Grifo nosso). Esse tipo de aplicação não foi explorado explicitamente em outras obras. Vejamos as atividades desta obra.

- **Contexto matemático**

O seu enunciado é:

Uma população muito grande tem média 20,0 e desvio padrão 1,4. Extrai-se uma amostra de 49 observações. Responda: (a) Qual a média da distribuição amostral? (b) Qual o desvio padrão da distribuição amostral? (c) Qual a percentagem das possíveis médias amostrais que diferirão por mais de 0,2 da média da população? (STEVENSON, 1981, p. 182-182)

Essa atividade foi extraída do primeiro exemplo após a definição do teorema, o que nos foi bastante oportuno averiguar que, de certa forma, a metodologia do livro privilegia as técnicas em detrimento das situações-problema do cotidiano. Uma

característica também comum em outras obras que apresentam atividades em ambos contextos: matemático e cotidiano.

Para a resolução, o autor parte do pressuposto de que, se o tamanho da amostra é superior a 30, $n > 30$, pode-se supor “normal a distribuição amostral”. Nesse caso, a abordagem do tamanho da amostra a ser considerado já havia sido “estipulado” anteriormente, como por conveniência. E, assim, para responder ao item (a), chega-se à conclusão de que “a média da distribuição amostral é sempre igual à média da população. Logo, $\mu_{\bar{x}} = 20,0$.” (STEVENSON, 1981, p. 183). A tarefa (T) foi a de encontrar a média da distribuição amostral; a técnica (τ) refere-se ao resultado que o conceito do Teorema Central do Limite garante, o que justifica a técnica, que é a tecnologia (θ). Já a teoria (Θ) é o próprio conceito do teorema, portanto a praxeologia é local, não estendendo para regional.

Para responder ao item (b), o autor aplica diretamente a fórmula do desvio padrão da distribuição amostral:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{1,4}{\sqrt{49}} = \frac{1,4}{7} = 0,2 \quad (10)$$

Nota-se que privilegia a técnica do *fazer*, pois não justifica por que é necessário o desvio padrão da distribuição amostral ser menor do que o da população. Outro item que nos chamou atenção foi em relação à notação que atribui para o desvio padrão da amostra e o desvio padrão da população, $\sigma_{\bar{x}}$ e σ_x , respectivamente, pois, em geral, o da população não há o “índice” com x .

O item (c) exige um pouco mais de detalhes, principalmente com os objetos matemáticos, como, por exemplo, o conceito de área e a ideia do que vem a ser “diferir por mais de 0,2 da média populacional”, a partir da ilustração do gráfico que está representada na figura 12. No entanto, apresenta os cálculos para encontrar os valores reduzidos, z :

$$\begin{aligned} \frac{20,2 - 20}{0,2} &= +1\sigma_{\bar{x}} && \text{proporção: } 0,1587 \\ \frac{19,8 - 20}{0,2} &= -1\sigma_{\bar{x}} && \text{proporção: } \underline{0,1587} \\ \text{Total: } &&& 0,3174 \end{aligned}$$

Em seguida, apresenta a ilustração:

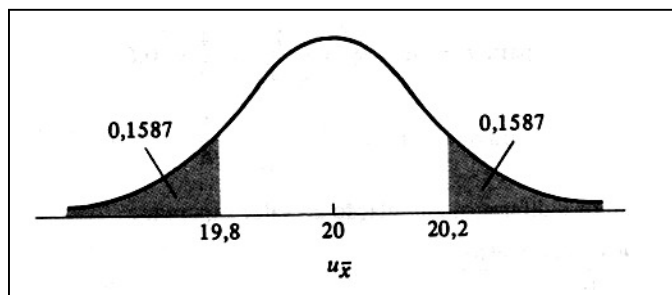
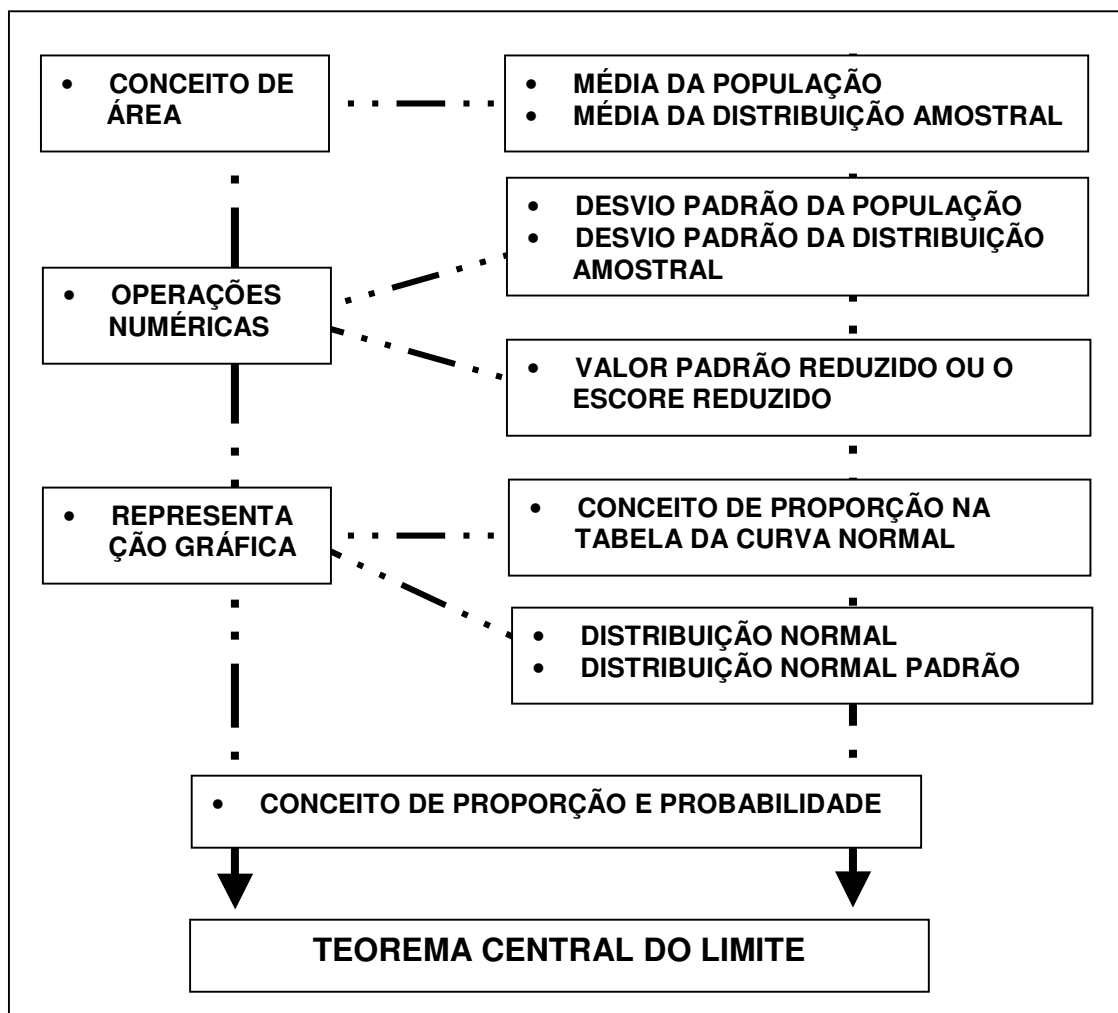


Figura 12 – As porções sombreadas são iguais à probabilidade de uma média amostral inferior a 19,8 ou superior a 20,2
Fonte: STEVENSON, 1987, p. 183

A realização dessa atividade requer algumas habilidades que nem sempre são tão praticadas assim, como, por exemplo, o que significa diferir por mais de 0,2 da média da população, conforme já havíamos dito. Para que o estudante entenda essa linguagem, ele ou ela precisa estar familiarizado(a) que, nesse contexto, significa diferir abaixo e acima da média populacional em 0,2, isto é, representam as áreas que residem nas caudas da curva Normal, cujas referências são os valores menores que 19,8 ou maiores que 20,2. Se não fosse a ilustração, talvez o estudante demorasse mais tempo para chegar à estratégia a ser utilizada. Isso faz parte da técnica (τ) da praxeologia dessa atividade.

Além disso, a apresentação do valor do escore reduzido não está claro, pois a notação, $+1\sigma_{\bar{x}}$, que significa mais um do desvio padrão da distribuição amostral, não é frequentemente utilizada em outros livros-texto. Em seguida, indica a proporção de 0,1587 e também para $-1\sigma_{\bar{x}}$, com o mesmo valor, para finalmente apresentar o total. Didaticamente, a resolução parece mais uma forma de “adivinhação” do que uma sequência didática, contrapondo, assim, a proposta da construção dos conceitos por parte do estudante, no entorno do teorema. Evidentemente, a representação gráfica foi fundamental para compreensão da resolução proposta pelo autor, mas, ainda assim, não permite a evolução das ideias matemáticas e estatísticas presentes na atividade.

Uma estrutura ecológica dessa atividade, somente para o item (c), seguindo a evolução dos saberes, poderia ser descrita conforme o esquema 9:



Esquema 9 – Ecossistema para o Teorema Central do Limite na atividade (c).

Fonte: Dados da pesquisa.

À esquerda, estão os objetos matemáticos necessários para realização da atividade, e os da direita, os da Estatística, o que mostra a necessidade dos saberes que são imprescindíveis para que o teorema “sobreviva”, mesmo no contexto matemático da questão, isto é, com aplicação de técnicas (τ), nem sempre devidamente justificado, isto é, desprovido da tecnologia (θ), segundo a resolução do livro-texto A6.

- **Contexto cotidiano**

Ainda nessa obra, A6, optamos também por expor um exemplo, no sentido de analisarmos a estrutura ecológica da atividade proposta.

Um fabricante de baterias alega que seu artigo de primeira categoria tem uma vida esperada (média) de 50 meses. Sabe-se que o desvio padrão correspondente é de 4 meses. Que percentagem de amostras de 36 observações acusará vida média no intervalo de 1 mês em torno de 50 meses, admitindo ser a verdadeira vida média das baterias? Qual será a resposta para uma amostra de 64 observações? (STEVENSON, 1981, p. 183-184)

Vejamos como o autor expõe a sua resolução.

Ele estabelece um diálogo com o leitor para explicar a tarefa:

Sabemos que, como $n > 30$, a distribuição de médias amostrais será aproximadamente normal, com média igual à média populacional e desvio padrão igual ao desvio padrão populacional dividido pela raiz quadrada do tamanho da amostra. (STEVENSON, 2001, p. 183)

Nesse item, o autor apresenta a técnica (τ) para a tarefa (T). E continua: A figura 7 “ilustra a probabilidade desconhecida. A resolução envolve a determinação do número de desvios padrões que 49 e 51 meses distam da média” e, assim, o autor sugere consultar a tabela de áreas sob a curva normal padronizada, dispensando, assim, aplicação do modelo matemático para o cálculo da área, e obtendo as probabilidades desejadas. Nesse caso, podemos dizer que a técnica (τ) para tarefa (T) foi substituída pela justificativa de que já existe uma tabela em que foram calculados os valores das áreas que correspondem às probabilidades desejadas, por meio do escore reduzido, variando de 0 a 3 e, a área, de 0 a 0,5, que também é chamada de Tabela da Curva Normal Padronizada. Por outro lado, exige outras técnicas (τ), como, por exemplo, o cálculo do desvio padrão da distribuição amostral:

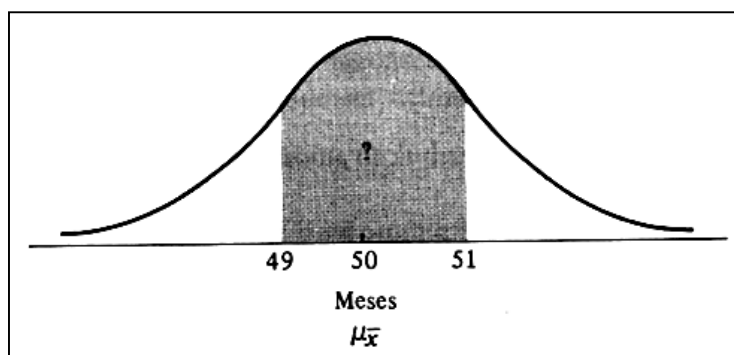


FIGURA 13 – Determinação da área sombreada da distribuição amostral.

Fonte: STEVENSON, 2001, p. 184.

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \begin{cases} \text{Para } n=36; \sigma_{\bar{x}} = \frac{4}{\sqrt{36}} = \frac{4}{6} \cong 0,67 \\ \text{Para } n=64; \sigma_{\bar{x}} = \frac{4}{\sqrt{64}} = \frac{4}{8} = 0,50 \end{cases}$$

(11)

Em seguida, determina a diferença em relação ao valor esperado, isto é, o escore reduzido para cada uma das variáveis que distam da média: $x_1 = 49$ e $x_2 = 51$, sendo que estes valores são expressos por:

$$z = \frac{\text{estatística} - \text{parâmetro}}{\sigma_{\bar{x}}}$$

(12)

Assim, fica explicada e justificada a técnica (τ):

Para $n = 36$: $z_1 = -1,5\sigma_{\bar{x}}$ e $z_2 = 1,5\sigma_{\bar{x}}$;

Para $n = 64$: $z_1 = -2,0\sigma_{\bar{x}}$ e $z_2 = 2,0\sigma_{\bar{x}}$;

(13)

Enfim, para a determinação das áreas, por meio da tabela da curva Normal padronizada, tem-se que:

Para $n = 36$, $P(49 < \bar{x} < 51) = P(-1,5 < \bar{x} < 1,5) = 0,8644$;

Para $n = 64$, $P(49 < \bar{x} < 51) = P(-2,0 < \bar{x} < 2,0) = 0,9546$.

(14)

Percebe-se que há também preocupação do autor em justificar esses resultados (tecnologia, θ), quando afirma que:

Note-se que, mesmo permanecendo constante o intervalo de 49 a 51, as respostas para as amostras de tamanhos 36 e 64 são diferentes. A probabilidade de obter uma média amostral no intervalo dado é maior para amostras de 64 observações do que para amostras de 36 observações, devido ao fato de o desvio padrão da distribuição amostral *decrecer* quando n aumenta. (STEVENSON, 1981, p. 184)

A conclusão deixa implícita a aplicação do Teorema Central do Limite. No entanto, podemos dizer que essa atividade contemplou o complexo praxeológico ($T / \tau / \theta / \Theta$), pois, para atingir as organizações regionais ao redor de uma teoria, exigiu-se também a combinação entre as organizações locais e pontuais sob a forma ($T_i / \tau_i / \theta / \Theta$) que “centradas sobre uma tecnologia, θ , determinada e depois em organizações regionais ($T_{ij} / \tau_{ij} / \theta / \Theta$) formadas ao redor de uma teoria Θ ”. (CHEVALLARD, 1999, p. 228), constitui-se, portanto, uma organização global.

A estrutura ecológica para esse caso é similar aos anteriores.

6.1.2.7 Atividades do livro-texto A7

No último livro-texto analisado, A7, não há nenhuma atividade contendo somente contexto matemático. Portanto nos restringimos a analisar apenas uma atividade com contexto cotidiano conforme feito anteriormente com outras obras. Vejamos seu enunciado:

Na engenharia humana e no projeto de produtos, frequentemente é importante considerar os pesos das pessoas, de modo que não haja sobrecarga em aviões ou elevadores, as cadeiras não se quebrem, e não ocorram outros acontecimentos perigosos ou embaraçosos. Dado que a população de homens tem pesos distribuídos normalmente com média de 173 lb e desvio padrão de 30 lb (com base em dados do National Health Survey dos EUA), determine a probabilidade de que: (a) um homem escolhido aleatoriamente pese mais de 180 lb; (b) em 36 homens escolhidos aleatoriamente, o peso médio seja superior a 180 lb.²⁴ (TRIOLA, 1999, p. 129)

²⁴ Uma libra (massa) corresponde aproximadamente a 0,4536 kg.

Esse exemplo, exposto por Triola (1999, p. 129), apresenta dois itens para serem solucionados e comparados. O item (a) refere-se à massa de um homem, escolhido aleatoriamente e, portanto, não constitui numa amostra. Calculando o valor padronizado ou o escore reduzido, já que foi afirmado que os pesos (massas) são “distribuídos normalmente”, podemos entender que se trata de uma distribuição de curva Normal. Assim:

$$z = \frac{x - \mu}{\sigma} = 0,23 \quad (15)$$

Ele também recorre à tabela da curva Normal padrão para encontrar a área que corresponde a 0,4090, isto é, $P(x > 180) \cong P(z > 0,23) = 0,5 - 0,0910 = 0,4090$. Portanto, a probabilidade de o homem pesar mais de 180 lb é 0,4090 e a representação gráfica dada por:

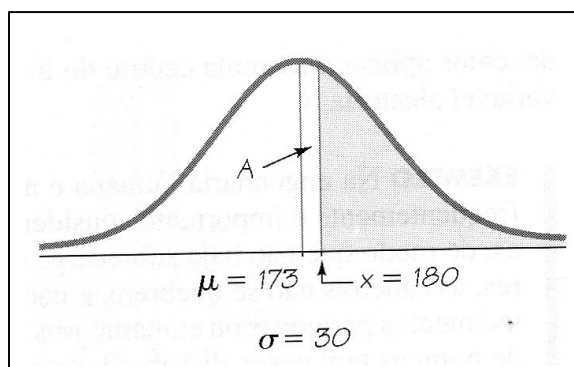


Figura 14: Distribuição de pesos de homens.
Fonte: TRIOLA, 1999, p. 130(a).

Para o item (b), o autor indica que “*utilize o teorema central do limite* (porque estamos lidando agora com a *média para um grupo* de 36 valores, e não um valor individual).” Em seguida, ele justifica que agora se trata de uma distribuição de médias amostrais, daí a necessidade de utilizar os parâmetros $\mu_{\bar{x}}$ (média das médias amostrais) e $\sigma_{\bar{x}}$ (desvio padrão da amostra), que são:

$$\mu_{\bar{x}} = \mu = 173 \quad (16)$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{30}{\sqrt{36}} = 5 \quad (17)$$

E, para determinar a área sombreada conforme a figura 14, deve-se calcular o valor padronizado z , ou o escore reduzido, incidindo a utilização da tabela da curva Normal padrão. Então,

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{180 - 173}{\frac{30}{\sqrt{36}}} = \frac{7}{5} = 1,40$$

(18)

E, a partir da figura 15, a representação gráfica da área indicada é $0,5 - 0,4192 = 0,0808$, que é a probabilidade de os 36 homens terem peso (massa) médio superior a 180 lbs.

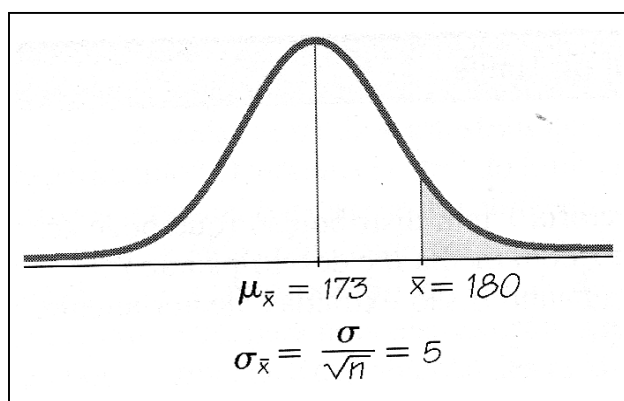


Figura 15 – Distribuição de médias amostrais de 36 homens.
Fonte: TRIOLA, 1999, p. 130(b).

O autor ainda justifica que “há uma probabilidade de 0,4090 de um homem pesar mais de 180 lb, mas a probabilidade de 36 homens terem peso médio superior a 180 lb é de apenas 0,0808.” Essa conclusão se dá, ao menos intuitivamente, pois é muito mais fácil um único indivíduo se afastar da média, do que um grupo de 36 indivíduos. “Um peso extremo entre os 36 perderá seu impacto quando considerado em conjunto com os outros 35 pesos.” (TRIOLA, 1999, p. 129).

Analisando essa atividade sob o ponto de vista praxeológico, não há muita ênfase na técnica, o que, no entanto, não significa que não há utilização. O autor opta por utilizar mais a retórica do que símbolos, em comparação às outras atividades analisadas. Também não há ênfase, por parte do autor, em representar o modelo matemático para indicar as áreas que correspondem às probabilidades encontradas, tendo em vista que todos os problemas são contextualizados para o

cotidiano, isto é, ele simula situações hipotéticas reais para que haja aplicação do Teorema Central do Limite.

Entretanto, para análise praxeológica, podemos dizer que a tarefa T exigiu técnicas τ (o que se deseja calcular), justificadas por uma tecnologia θ , quando é ou não caso de aplicação do Teorema Central do Limite. Para justificar a tecnologia, tem-se a teoria Θ , principalmente quando isso ocorre para comparar os casos: ao aplicar o teorema (b) e não aplicar, (a).

Diante desses resultados, percebemos que cada obra tem suas próprias características ao tratar-se do Teorema Central do Limite. Com exceção de A4, todos outros apresentam uma definição para o teorema. Vejamos como eles abordam no próximo subitem.

6.1.3 Definição do Teorema Central do Limite nos Livros-Texto.

De acordo com as análises das atividades dos sete livros-texto, nos foi possível apontar algumas características próprias de cada obra. As mesmas ocasionalmente são exitosas, mas apresentam também restrições que podem servir como entraves no processo de aprendizagem. Em contrapartida, determinadas restrições ou limitações enriquecem o *design didático*, pois servem como ferramentas excelentes para reflexão e discussão entre especialistas, professores e alunos.

Entretanto, alguns pontos comuns nos foi possível identificar. Todos os livros-texto apresentam primeiramente a definição do teorema, não incorrendo em sua demonstração e, em seguida, algumas aplicações sob forma de problemas de contexto matemático e/ou do cotidiano, reforçando as ideias em torno do Teorema Central do Limite. Segundo Bussab e Morettin (2003, p. 273), “a demonstração completa desse teorema exigiria recursos dos quais não dispomos, portanto não será dada, mas o importante é sabermos como esse resultado pode ser usado.”

Diante do que constatamos na análise dos livros-texto e a afirmativas desses autores, podemos dizer que um dos **saberes dispensáveis do Teorema Central do Limite**, no processo de ensino e de aprendizagem, é a demonstração de como se chega a esse resultado, porém defendemos que o contexto histórico deveria ser

ênfatisado, principalmente porque, na nossa investigação, os atores são os licenciandos em Matemática.

Essa abordagem nos proporcionou o ponto de vista do professor ao elaborar suas atividades didáticas, ao tratar-se do teorema. Tauber (2001, p. 69) interpreta essa dinâmica como um “significado institucional de referência”, pois serve como base de seleção de objetos significativos que estão de acordo com o enfoque para a construção de uma sequência didática. Por isso achamos interessante apresentar como o teorema é definido pelos autores, pois todos alegam sua extrema importância, com exceção de A4, porque não o menciona. Assim, vejamos como cada um deles define o Teorema Central do Limite:

(continua)

OBRA	DEFINIÇÃO DO TEOREMA CENTRAL DO LIMITE
A1	<i>Teorema:</i> Para amostras aleatórias simples $(X_1, \dots, X_n)$, retiradas de uma população com média μ e variância σ^2 finita, a distribuição amostral da média $\bar{X}$ aproxima-se, para n grande, de uma distribuição normal, com média μ e variância σ^2/n . (p. 273-274)
A2	Para uma distribuição <i>não</i> -normal com média μ e desvio padrão σ , a distribuição da média amostral $\bar{X}$ para amostras de tamanho n suficientemente grande é aproximadamente normal com média μ e desvio padrão $\frac{\sigma}{\sqrt{n}}$, isto é: $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0; 1) .$ (p. 139-141)
A3	Sejam $X_1 \dots X_n$ variáveis aleatórias independentes com a mesma distribuição, que se admite ter variância finita (quase todas as distribuições com interesse prático têm variância finita, pelo que esta condição não é particularmente restritiva). Qualquer que seja a forma da distribuição destas variáveis, se o valor N for suficientemente grande, a variável soma $S = \sum_{n=1}^N X_n$ segue aproximadamente uma distribuição Normal. Esta distribuição é inteiramente especificada através do valor esperado e da variância de S , que são dados por $\mu_S = N \cdot \mu_X$ e $\sigma_S^2 = N \cdot \sigma_X^2$, onde μ_X e σ_X^2 , representam o valor esperado e a variância das variáveis X_n . (p. 241)
A4	NÃO DEFINE O TEOREMA CENTRAL DO LIMITE
A5	Extraia uma AAS, <i>amostra aleatória simples</i> , de tamanho n de qualquer população com média μ e desvio padrão finito σ . Quando n é grande, a distribuição amostral da média amostral $\bar{x}$ é aproximadamente Normal: $\bar{x} \text{ é aproximadamente } N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) .$ (p. 223)

(conclusão)

A6	1. Se a população sob amostragem tem distribuição normal, a distribuição das médias amostrais também será normal para todos os tamanhos de amostra. 2. Se a população básica é não-normal, a distribuição de médias amostrais será aproximadamente normal para grandes amostras. (p. 181-182)
A7	Dado: 1. A variável aleatória x tem distribuição (que pode ser normal, ou não), com média μ e desvio-padrão σ . 2. Amostras de tamanho n são extraídas aleatoriamente dessa população. Conclusões: 1. Na medida em que o tamanho da amostra aumenta, a distribuição das médias amostrais $\bar{x}$ tende para uma distribuição normal. 2. A média das médias amostrais será a média populacional μ . 3. O desvio-padrão das médias amostrais será $\sigma / \sqrt{n}$. (p.129)

QUADRO 2- Definição do Teorema Central do Limite nos livros-texto.**Fonte: Dados da pesquisa.**

De acordo com o Quadro 2, pode-se notar que as definições acerca do Teorema Central do Limite são bastante similares. Em síntese, dada uma distribuição de variáveis aleatórias de distribuição Normal ou não, pode-se afirmar que, na medida em que aumenta o tamanho das amostras, a média das médias amostrais será a média da população; e, o desvio padrão, das médias amostrais, será $\sigma / \sqrt{n}$.

A comparação entre esses livros-texto sobre a definição do teorema pode ser sintetizada da seguinte forma: A1 e A2 distinguem-se no tipo de distribuição de amostras, isto é, se é, inicialmente, Normal ou não, tendo em vista que a primeira obra não especifica o tipo, e a segunda chama atenção para distribuições não-Normais. A4 não apresenta definição, mas mostra o comportamento pela representação gráfica, conforme a figura 16.

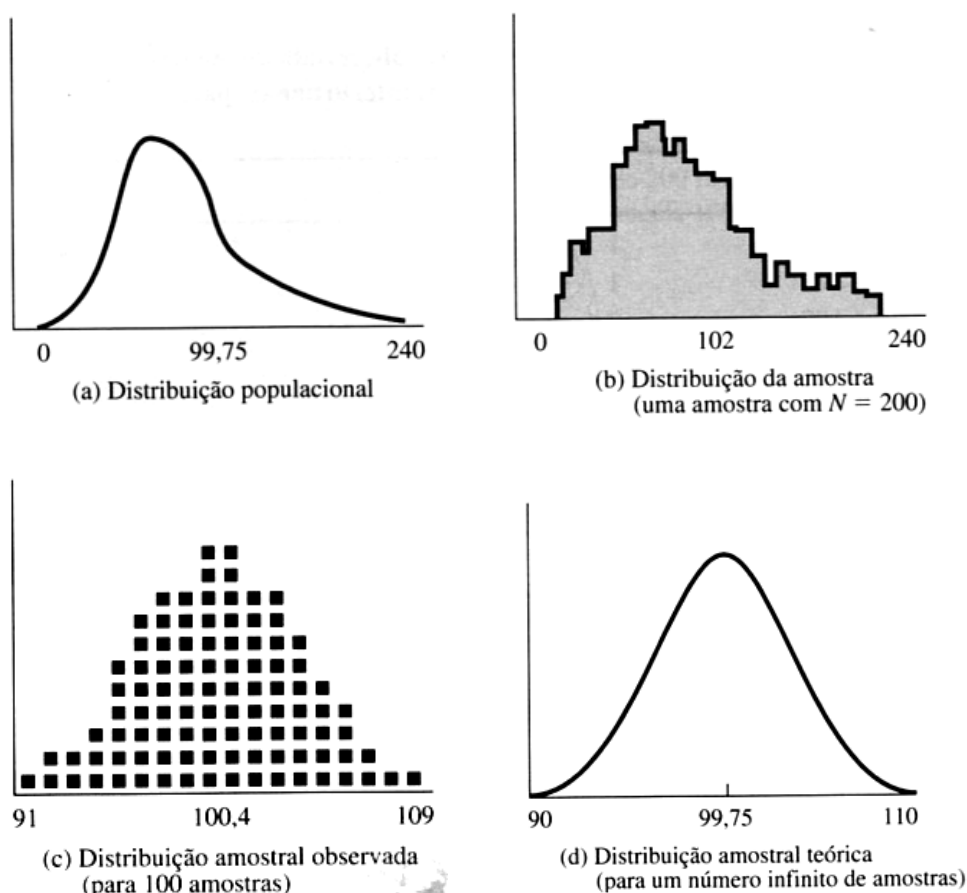


FIGURA 16 - Distribuições: populacional, da amostra, amostral observada e amostral teórica
Fonte: LEVIN; FOX, 2004, p. 188.

Vejamos como os autores interpretam esses gráficos que dizem respeito ao Teorema Central do Limite, mas sem mencioná-lo:

[...] a variabilidade de uma distribuição amostral é sempre menor do que a variabilidade em toda a população ou em qualquer uma das amostras. A *figura (10a)* mostra a distribuição populacional de chamadas interurbanas, com média (μ) de 99,75 (em geral não dispomos dessa informação). A distribuição é assimétrica à direita: mais famílias despendem menos tempo do que a média de 99,75 minutos em uma chamada interurbana, mas algumas, na cauda direita, parecem não querer largar o telefone. A *figura (10b)* mostra a distribuição das durações de chamadas em uma determinada amostra de 200 residências. Note que ela tem forma semelhante à distribuição da população, com média aproximadamente igual ($\bar{X} = 102$). A *figura (10c)* mostra a distribuição amostral de médias (as médias de 100 amostras do excêntrico). Ela se afigura razoavelmente normal, e não assimétrica, tem média (100,4) quase igual à média da população e menos variabilidade, seja em relação à distribuição populacional em (10a) ou em (10b), o que se pode ver comparando os valores da reta-base. Se o pesquisador tivesse continuado a extrair amostras de 200 residências, o gráfico das médias dessas amostras se assemelharia a uma curva normal, conforme mostra a *figura (10d)*. Essa é a verdadeira distribuição amostral. (LEVIN; FOX, 2004, p. 187) (Grifo nosso)

Apesar de A4 não ter definido e nem mencionado o Teorema Central do Limite, apresenta uma situação em que parte do princípio de uma situação hipoteticamente real, de tempo médio mensal de duração de chamadas interurbanas, considerando-a como uma população de residências, de média 99,75 minutos de duração e, neste caso, a curva é assimétrica à direita, conforme mostra a figura 16(a). Na figura 16(b), é uma distribuição de amostra de 200 residências, com média aproximada de 102 minutos. A diferença entre as situações 16(a) e 16(b) está no fato de que o primeiro trata de população e o segundo, de uma amostra da população de 200 residências. Agora, na figura 16(c), não é nem a população e nem uma amostra de um determinado número de residências, mas sim de uma distribuição de médias de duração de interurbanos com número de 100 dessas amostras. O gráfico, neste caso, já possui uma tendência à normalização quanto a sua curva. E, por fim, o gráfico 16(d) simula o comportamento do gráfico para 200 amostras de distribuições de médias amostrais de aspecto similar ao de um número infinito de residências.

Ora, o que está implícito é o próprio Teorema Central do Limite, tendo em vista que sua definição garante que a média das distribuições de médias amostrais tende a ser a média da população. Para o desvio padrão, os autores afirmam que “variabilidade” tende a ser sempre menor quando é extraída de distribuições de médias amostrais.

Os livros-texto A5 e A6 seguem a mesma linha que A1. Já A3 e A7 são mais detalhistas; a primeira não especifica o tipo de distribuição, apesar de ser mais específico nos modelos matemáticos; a obra A7 alerta para o fato de que a distribuição pode ser normal ou não, contanto que o tamanho da amostra seja suficientemente grande.

Além das definições apresentadas, vale ressaltar o aspecto didático na obra (A5) de Moore. A justificativa para tal se deve ao fato de que o objeto matemático, Teorema Central do Limite, não se esgota em um tópico, mas se projeta ao longo do livro-texto, sendo referido em diversos itens que dizem respeito à Inferência Estatística. Como, por exemplo, nas distribuições amostrais, quando se aplica, explorando dados para inferência; na inferência sobre uma média populacional; e em situações-problema para duas médias. O extraordinário de sua obra reside no fato de o autor sempre reportar-se ao teorema, chamando a atenção para sua importância e aplicabilidade, fato este não constatado em outras obras.

De modo geral, “o Teorema Central do Limite nos permite usar cálculos de probabilidades Normais para responder perguntas acerca de médias amostrais de muitas observações, mesmo quando a distribuição populacional não seja Normal” (MOORE, 2005, p. 224). Esse fato confere porque quanto mais uma distribuição estiver afastada da normal, maior deve ser o tamanho da amostra.

6.2 A Análise da Investigação Sob o Ponto de Vista Ecológico do Didático: o Teorema Central do Limite e os Licenciandos

Segundo Almouloud (2007, p. 123), “as organizações matemáticas referem-se à realidade matemática que se pode construir em uma sala de aula”. Nesse contexto, a definição do teorema diz respeito ao planejamento do(a) professor(a), incidindo na construção de ideias e conceitos sobre o teorema, que será reconstruído pelos estudantes. Consequentemente, a maneira de como a definição chegará a eles sofrerá interferências no reconstruir, culminando, portanto, na praxeologia didática do Teorema Central do Limite.

Nos capítulos anteriores, constatamos a importância do Teorema Central do Limite e suas implicações de ordem histórica e de aplicabilidade, que podem contribuir para o processo de ensino e de aprendizagem no curso de Licenciatura em Matemática. Contudo, algumas limitações verificadas nos livros-texto, no estudo praxológico das atividades relativas ao teorema, nos levam a “ressignificar” a prática em sala de aula, como professores e pesquisadores.

Tais limitações, na verdade, fomentaram o nosso interesse em compreender como as ideias e os conceitos, acerca do teorema, podem ser construídos a partir de situações-problema, que dizem respeito ao cotidiano da prática do ensino e da aprendizagem em Matemática na sala de aula, no nível de Educação Básica. Entretanto, inicialmente, foi preciso buscar que tipo de saberes já estavam apropriados por eles, os licenciandos, em relação à Estatística Básica²⁵. E, para realização desse trabalho, partimos do desejo de avaliar as atividades propostas na pesquisa descrita, buscando aprofundar o conhecimento dos alunos sobre a

²⁵ Entendemos por “Estatística Básica” os temas que dizem respeito à Estatística Descritiva e à Teoria das Probabilidades.

importância do teorema na aplicação de situações-problema do cotidiano, que muitas vezes estão evidenciadas nos meios de comunicação. Reforçando a proposta de nossa investigação, seguimos uma metodologia de ensino em que privilegiamos a compreensão das ideias de quando e para quê a utilização do Teorema Central do Limite.

6.2.1 As Atividades com os Licenciandos

Para a realização de nossa investigação, caracterizamos o nosso trabalho, propondo um procedimento metodológico qualitativo, reiterando a estrutura do *design didático*. Contamos, inicialmente, com apenas quatro licenciandos do Curso de Licenciatura em Matemática, de um Centro de Ensino da cidade de Juiz de Fora, Minas Gerais.

Para a mobilização das ideias e dos conceitos estatísticos, precisamos nos situar na teoria dos campos conceituais de Vergnaud (1996, p. 155), que prima em “fornecer um quadro coerente e alguns princípios de base para o estudo do desenvolvimento e da aprendizagem das competências complexas, nomeadamente daquelas que relevam das ciências e das técnicas”. É uma abordagem que favorece os objetivos da pesquisa pelo fato de que o ensino do teorema requer, principalmente, sob o ponto de vista didático do mesmo, favorecer o processo de aprendizagem dos licenciandos.

Um outro aspecto a ser considerado por essa teoria é o que diz respeito ao quadro que nos permite “compreender as filiações e as rupturas entre conhecimentos” (VERGNAUD, 1996, p. 155). Valemo-nos do sentido proposto por Vergnaud (1996), que expressa o termo *conhecimento* tanto para o saber fazer quanto para o saber expresso. Essa abordagem destaca-se pela atenção que atribui à aprendizagem dos adultos, o que particularmente nos interessa, já que os sujeitos da pesquisa são os licenciandos.

Desse modo, quando dizemos que o objetivo de nosso estudo é desenvolver, analisar, avaliar e reconstruir, numa proposta de ensino, as principais ideias e conceitos ligados ao Teorema Central do Limite, significa desenvolver (ou criar) um olhar voltado para ações dos alunos à medida que eles expressam os objetos

matemáticos implícitos nos conceitos estatísticos. Buscamos, portanto, a análise e a compreensão dos procedimentos e argumentos de que cada um se fez valer em cada situação.

A seguir, apresentaremos as sequências didáticas que nos proporcionaram a realização dessa investigação.

6.2.2 A Primeira Atividade com os Licenciandos

De acordo com os princípios expostos, a elaboração das atividades permitiu surgir as ideias e os conceitos da Estatística Descritiva, tais como as medidas de tendência central e outras advindas da variabilidade de diferentes conjuntos de dados. Isso favoreceu o surgimento de conceitos espontâneos dos alunos sobre o tema, fato que, conseqüentemente, nos propiciou melhor compreensão de como ocorre aquisição e internalização significativa no tocante às ideias e aos conceitos estatísticos.

Conforme já citado anteriormente, contamos com quatro licenciandos²⁶ como sujeitos da pesquisa. Para conhecer a experiência que cada um tinha com a Estatística, foi realizada uma etapa exclusiva a fim de que eles dialogassem e manifestassem os conhecimentos relacionados à análise exploratória de dados e às ideias acerca do teorema. Para tal, utilizamos um questionário aberto, cujas perguntas enfatizavam as ideias estatísticas necessárias para a estrutura ecológica do Teorema Central do Limite sob o ponto de vista de objeto do saber.

Para Artigue (1988), uma investigação *prima* por registrar as manifestações que revelam as ações e seus possíveis problemas que podem ocorrer no sistema de ensino. Nessa perspectiva, surgem questões, tais como: de que forma os alunos de um curso de Licenciatura em Matemática se apropriam das ideias do conceito do Teorema Central de Limite? Pode ser por meio de um contexto relacionado com a

²⁶ O período corresponde ao 2º semestre de 2007. Devido à mudança de programa, havia alunos que cursavam vários períodos simultaneamente, por isso não houve demanda na disciplina, já que outras disciplinas estavam em curso. Apenas três alunos estavam cursando regularmente a grade curricular proposta desde ingresso dos mesmos, isto é, o sétimo período do curso de Licenciatura desse Centro de Ensino, correspondendo ao penúltimo semestre do curso.

prática do professor de Matemática em sala de aula? Quais invariantes operatórios são construídos por eles para caracterização do Teorema Central do Limite?

Para Brousseau (1996, p. 67), “o que o aluno tem em sua memória parece ser o objetivo final da atividade de ensino”. Seguindo esse pressuposto, nos propomos a conhecer os saberes prévios que os sujeitos possuem em relação às ideias sobre o Teorema Central do Limite. Para isso, estabelecemos algumas perguntas de tal forma que nos proporcionou algumas considerações para elaboração de outras atividades. Vejamos como ocorreu a primeira atividade de nossa investigação, no intuito de reconhecer os conhecimentos prévios de cada um dos licenciandos.

- ATIVIDADE 1: Responda às perguntas a seguir:

OBJETIVOS: Conhecer os elementos estatísticos que os alunos da Licenciatura possuem.

- 1) O que você entende por variável ou variáveis?
- 2) O que você entende por aleatoriedade?
- 3) E independência de dois ou mais eventos?
- 4) Para você, a variável aleatória tem alguma relação com função matemática?
- 5) Elabore uma definição para variável aleatória e variáveis aleatórias independentes.

Essa atividade²⁷ foi realizada a partir de um questionário de cinco perguntas abertas, com a finalidade de averiguar que ideias e conceitos acerca do teorema nos permitiriam auxiliar para a compreensão do teorema. No sentido de preservar a identidade dos estudantes, os chamaremos pelas letras: AO, BM, CA e FA.

(continua)

1) *O que você entende por variável ou variáveis?* Essa pergunta relaciona-se diretamente com as idéias algébricas que estão subjacentes ao teorema. As respostas foram as seguintes:

AO: Aquilo que não é constante, como, por exemplo, avaliar o grau de autoestima de alunos da 5ª Série. Os resultados obtidos são as variáveis.

BM: Possibilidade de que algo ocorre varie, ou seja, não ser um resultado constante.

CA: Valores que quero obter.

²⁷ As respostas dessa atividade encontram-se, na íntegra, no anexo 2.

(conclusão)

FA: Variável é a incógnita, ou seja, o que eu quero achar. As variáveis podem assumir infinitos valores.

A partir dessas respostas, podemos concluir que AO tem a concepção formada do que vem a ser *variável* no contexto matemático, mesmo não sabendo previamente o objetivo dessa pergunta. O estudante BM tem as mesmas concepções que AO, apesar de a resposta ser um pouco diferente. Já CA deixa a resposta um tanto vaga, não nos permitindo auferir sobre aquilo que se entende por variáveis. O estudante FA chama de variável o mesmo que incógnita, apesar de reconhecer como aquela que pode assumir infinitos valores.

2) O que você entende por aleatoriedade? A proposta dessa pergunta foi a de buscar ideias no entorno dos eventos não-tendenciosos, sendo, portanto, um item bastante importante na análise estatística. As respostas foram as seguintes:

AO: Algo que não se pode prever como, por exemplo, retirar uma bola aleatoriamente de uma urna, numeradas de um a cem.

BM: Algo que ocorra ao acaso.

AC: Não ter uma ordem para escolher alguma coisa, tomar algo ao acaso.

FA: Aleatoriedade é o acaso. Num evento aleatório, não há escolha e sim a ocorrência é o acaso.

Com exceção do estudante AO, todos associaram diretamente aleatoriedade com o acaso. No entanto, com outras palavras, AO também atribuiu para o acaso. Vale ainda ressaltar que AC registrou que se toma “algo ao acaso”, alegando que “aleatoriedade significa não ter uma ordem de escolha”. Ele recorreu diretamente à ideia de ordenação, fato esse que confirmou restrição quanto à sua afirmativa.

(continua)

3) E independência de dois ou mais eventos? A seguir, as respostas:

AO: Quando um evento não depende da outra para ser estudado.

BM: É quando a ocorrência do evento não depende das dos demais.

CA: É quando dois ou mais eventos podem ser ‘definidas’ sem levar em consideração outro, ou melhor, podemos analisá-lo sem nos preocupar com outro.

(conclusão)

FA: Independência de dois eventos ocorre quando um não depende do outro para ser analisado.

Nesse estágio, pressupôs-se que os estudantes já conheciam o conceito de probabilidade e, portanto, poderiam concluir sobre o significado de eventos. Por isso, foi conveniente apresentar uma sequência de questionamentos que os levassem à definição do Teorema Central do Limite. Talvez, por esse motivo, todos chegaram a uma mesma conclusão.

4) Para você, a variável aleatória tem alguma relação com função matemática?

AO: Sim.

BM: Sim.

CA: Não, pois a variável aleatória poderá assumir infinitos valores em um intervalo finito, além de não seguir uma 'lei' (ordem).

FA: Sim, pois, assim como na função matemática, a variável aleatória pode assumir infinitos valores e está associada a outra variável.

Os dois primeiros estudantes apenas concordaram que existe relação entre a variável aleatória e a função matemática. Os outros dois divergiram em suas respostas. CA confirma o significado de variável, mas confronta com a definição de função matemática. Já FA responde que sim, justificando sua resposta de acordo com as ideias de função matemática, tendo em vista que a pergunta foi um tanto ampla.

(continua)

5) Em caso afirmativo ao item anterior, elabore um 'conceito' para variável aleatória e variáveis aleatórias independentes.

AO: Variável aleatória é aquela que os valores estão em função da aleatoriedade, ou seja, à medida que varia a amostra, variam-se os dados correspondentes. Variáveis aleatórias independentes têm o mesmo conceito anterior, porém, de forma independente uma da outra."

BM: Variável aleatória → Um experimento ocorre ao acaso.

Variáveis aleatórias independentes → Um experimento que não depende dos demais."

(conclusão)

CA: NÃO RESPONDEU.

FA: *Variável aleatória → as variáveis aleatórias podem assumir infinitos valores em um intervalo.**Variáveis aleatórias independentes → podem assumir infinitos valores, porém sem depender do valor da outra variável.*

O estudante CA não respondeu, sendo coerente com sua resposta anterior. AO se depara com a ideia de que a variável estabelece uma relação com aleatoriedade, surgindo, assim, uma função matemática. BM confirma dando a justificativa baseada no que respondeu anteriormente, e FA responde de forma similar.

Isso nos permitiu confirmar que a estrutura do questionário foi uma tentativa de dar um contorno ecológico aos saberes envolvidos com o Teorema Central do Limite. Os conceitos surgiram de forma natural, mesmo reconhecendo que as respostas sofreram influências do aprendizado matemático, tendo em vista que foram adquiridos ao longo da formação básica de cada um dos estudantes que participaram dessa atividade.

As atividades posteriores foram baseadas na situação-problema relacionada à autoestima dos alunos em Matemática, da 5ª Série, ou 6º ano, do Ensino Fundamental. E, para conhecer um pouco sobre os dados coletados, pois estes serão trabalhados para a construção do conceito do Teorema Central do Limite, vamos esclarecer alguns itens pertinentes à Educação Matemática Emocional.

6.2.3 Noções Preliminares das Atividades

Para nos situarmos das atividades propostas aos licenciandos, partimos de uma situação-problema no contexto real com o qual a maioria dos professores se depara, é o confronto com a autoestima dos alunos, da Educação Básica, com a Matemática, porque frequentemente os “pequenos” se sentem “incapazes” de compreender essa disciplina, seja por crença extremamente arraigada na nossa cultura, seja por “fracassos” que tenham enfrentado, entre outros fatores.

Diante das respostas obtidas pelos licenciandos, nos foi possível explorar a *literacia* estatística, nem sempre explicitamente, mas subjacente às suas respostas. Os dados da situação-problema, conforme já citados anteriormente, foram extraídos de uma pesquisa de Iniciação Científica que envolveu mensuração de atitude, que “é uma organização duradoura de crenças e cognições em geral, dotada de carga afetiva pró ou contra um objeto definido, que predispõe a uma ação coerente com as cognições e objetos relativos a este objeto” (BUNCHRAFT; CAVAS, 2002, p. 105), neste caso, a disciplina Matemática como elemento definido.

Daí a relevância da seguinte indagação como contexto para a nossa atividade com os licenciandos: por que a autoestima dos alunos é tão baixa em relação à disciplina de Matemática? É uma questão bastante discutida, mas que ainda persiste na reflexão entre educadores matemáticos. De acordo com Chacón (2003), a autoestima de cada sujeito é um dos componentes responsáveis pelo desempenho no processo de aprendizagem.

A pertinência desse assunto para os licenciandos é corroborada por Charnay (1996, p. 37), quando afirma que a relação entre a situação-problema e os alunos deve atender aos seguintes requisitos: uma atividade deve propor um verdadeiro *problema por resolver* para o aluno, isto é, deve ser compreendido por eles; deve permitir ao aluno que utilize os conhecimentos anteriores; mas, por outro lado, deve oferecer oportunidades de modo que o aluno vá além daquilo que já se conhece, isto é, um movimento de superação a partir do momento em que seus conhecimentos anteriores são gradativamente transformados, surgindo novos objetos do saber; por fim, o ideal é que a *validação* não seja produzida pelo professor, mas da própria situação em que o protagonista é o aluno e, no nosso caso, os licenciandos.

Por outro lado, como a nossa investigação está sob o jugo da Didática da Matemática, de acordo com Almouloud (2007, p. 34), “uma *situação didática* se caracteriza pelo jogo de interações do aluno com os problemas colocados pelo professor”, o que contempla as diretrizes dessa pesquisa. Assim, o desenvolvimento das atividades propostas será do próprio estudante em testemunhar ou vivenciar a construção de ideias e conceitos na dinâmica da teoria das situações didáticas.

Assim, portanto, estaremos efetivamente tentando reponder à problemática ecológica do objeto existente, o Terorema Central do Limite, que, segundo Artaud (1998), é uma forma de questionar o real: o que existe e por que existe? É nessa direção que conduzimos nossa pesquisa, dando oportunidade para o aluno se

expressar, descentralizando do professor a fala e o desenvolvimento de modos de pensar sobre o conteúdo estatístico em estudo. Por isso, optamos por recorrer a uma diversidade de situações dentro de um mesmo contexto, uma vez que é por meio de situações e de problemas a resolver que um conceito adquire sentido para o sujeito. Contudo, haverá também interferências controladas, por parte dos pesquisadores, pois, como já exposto, o professor também terá sua atuação como pesquisador, fundamentado nos pressupostos da pesquisa-ação de Thiollent (1986).

De modo geral e de acordo com a Atividade 1, podemos afirmar que nenhum deles questiona sobre o significado de *variável*, especificamente, tal como foi proposto, com exceção de AO, que apresentou alguns exemplos. Esse fato pode ser atribuído à ausência de comunicação que, supostamente, deveria ocorrer durante as aulas de Matemática na formação do indivíduo.

De fato, Medeiros (2005) confirma:

O recurso quase exclusivo às técnicas algébricas, cujo objetivo em Matemática é o de reduzir a linguagem, economizá-la, impede a construção da generalização e das abstrações matemáticas pelo aluno. A abstração é algo a ser atingido no ensino da Matemática. O uso precoce e exclusivo de tais técnicas, porém, induz comumente o aluno ao automatismo segundo as regras de um jogo, com a não compreensão das operações efetuadas sobre os números e a não apreensão dos significados matemáticos presentes que se pretende resolver. (MEDEIROS, 2005, p. 20)

A sequência de respostas obtidas nos proporcionou a análise dos constrangimentos sob três dimensões a serem consideradas: epistemológica, cognitiva e didática. Para Artigue (1996), a distinção das três dimensões é:

- a dimensão *epistemológica* associada às características do saber em jogo;
- a dimensão *cognitiva* associada às características cognitivas do público ao qual se dirige o ensino;
- a dimensão *didática* associada às características do funcionamento do sistema de ensino.

[...] O ensino habitual está centrado no funcionamento do quadro algébrico. Parece, por isso, bastante natural procurá-lo, tendo em conta o objeto preciso da investigação: estudar a viabilidade de uma abordagem epistemologicamente mais satisfatória, os constrangimentos que se opõem à extensão do ensino a outros quadros. (ARTIGUE, 1996, p. 200)

Nessa perspectiva, podemos citar os constrangimentos que se opõem à extensão do quadro conceitual do Teorema Central do Limite:

- Na esfera epistemológica: a complexidade do teorema frente aos obstáculos que surgiram ao longo da história da Matemática, desde seu nascimento e desenvolvimento até tal como é efetivamente sua ação no processo da transposição didática nos dias de hoje;
- Na esfera cognitiva: a mobilização entre os quadros algébricos e geométricos, considerando a complexidade da demonstração algébrica. Além disso, o tratamento do teorema é comumente feito pela definição e, em seguida, a passagem para representação gráfica, sendo, portanto, para o nosso caso, um tratamento qualitativo para o teorema;
- Na esfera didática: a concentração dos esforços por parte do professor, de orientar o aluno pelos meios que favoreçam aprendizagem, porém nem sempre passíveis de sucesso. Dessa forma, fica instaurado o caráter de estudo qualitativo.

Nesse sentido, podemos afirmar, de acordo com as respostas registradas, que as três esferas citadas acima foram contempladas.

Historicamente, a disciplina Matemática é temida por muitos estudantes da Educação Básica, pois é reconhecida como aquela que mais retém os alunos nas séries em que estudam. Nessa direção e voltando ao nosso contexto da situação-problema proposta, vejamos os dados relativos ao índice de autoestima que os alunos, da 5ª Série, ou 6º ano, têm com a Matemática, para que os licenciandos, os futuros professores da Educação Básica, estejam familiarizados com crenças e atitudes no processo de ensino e de aprendizagem, que fazem parte do cotidiano dos professores de Matemática. Dessa forma, podemos favorecer os licenciandos a exercer a criatividade para elaboração de estratégias na construção das ideias e dos conceitos estatísticos pertinentes ao teorema em estudo. E, nesse sentido, podemos dizer que:

A estratégia é um elemento essencial para o planejamento de um trabalho quantitativo simples, tanto para a elaboração de um projeto, a definição de hipóteses e de variáveis, como para a escolha dos sujeitos e para o processo de coleta de dados. Vemos o pensamento analítico como uma atitude estatística, ou melhor, uma atitude crítica do estudante, não apenas em relação às técnicas, com ou sem informática, mas principalmente em relação aos resultados obtidos no contexto em que os dados se encontram inseridos (social, comunitário, político, ambiental etc.). WODEWOTZKI; JACOBINI, 2004, p. 234-235)

Assim, a partir da escolha da estratégia utilizada para os dados quantitativos pela situação-problema, viabilizamos um sentido duplo, em termos de estratégia, para a nossa pesquisa, isto é, por um lado, há o ensino e toda a estrutura que o envolve; por outro, ressaltamos a importância de favorecer a aprendizagem dos licenciandos, ao propormos os dados referentes à autoestima dos alunos da 5ª Série do Ensino Fundamental, ou 6º ano. Além disso, estaremos, implicitamente, estimulando o espírito científico, por meio de questionamentos que porventura possam surgir durante as atividades.

A mensuração do índice de autoestima proposta na situação-problema foi realizada a partir das respostas de vinte problemas básicos de algoritmo²⁸. Para estabelecer as pontuações, o critério adaptado seguiu os procedimentos que Chacón (2003, p. 214) empregou em sua pesquisa, usando os métodos estatísticos para mensuração, o modelo teórico de Likert:

A Escala de Likert é composta, em sua forma final, por vinte e cinco afirmativas favoráveis e desfavoráveis ao objeto atitudinal, seguidas de alternativas, como por exemplo: concordo plenamente, concordo, não tenho opinião, discordo, e discordo plenamente. (BUNCHAFT; CAVAS, 2002, p. 107)

Para as vinte questões adaptadas de Chacón (2003), conforme o anexo 3, as alternativas são divididas em: ‘estou seguro que está correto’, ‘acredito que está correto’, ‘aposto 50% que está correto’, ‘acredito que está incorreto’, e ‘estou seguro que está incorreto’. A pontuação segue a condição de que “cada alternativa terá um peso de acordo com o índice de intensidade de favorabilidade ou desfavorabilidade” (CHACÓN, 2003, p. 109). Normalmente os pesos mais altos são diretamente proporcionais à favorabilidade, conforme segue tabela 2:

TABELA 2
Índice de autoestima da situação-problema proposta aos licenciandos

RESPOSTA	PONTUAÇÃO
<i>Estou seguro que está correto</i>	4
<i>Acredito que está correto</i>	3
<i>Aposto 50% que está correto</i>	2
<i>Acredito que está incorreto</i>	1
<i>Estou seguro que está incorreto</i>	0

Fonte: Dados da pesquisa de Iniciação Científica realizada no CES/JF.

²⁸ Veja anexo 3.

A semissoma das pontuações será a nota global do aluno (da 5ª Série, ou 6ª ano). A faixa de índice de autoestima possível é de 0 a 40 pontos, cujos dados coletados serão apresentados, na atividade 2, aos sujeitos da pesquisa, os licenciandos.

Ao expor as atividades a eles, um outro aspecto foi discutido e estreitamente relacionado às crenças e mitos em torno da Matemática. Isso diz respeito ao que influencia diretamente a dimensão emocional de cada criança em relação a essa disciplina.

[...] acreditamos que o estilo matemático está relacionado com as emoções. Hoje, há um crescimento da consciência coletiva sobre a necessidade de desentranhar os aspectos emocionais do conhecimento, nos quais possivelmente há que se buscar a raiz de muitos fracassos de nossa vida intelectual e, em particular, de nossa educação.

Se fizéssemos um estudo das palavras utilizadas nas discussões dos professores e dos pesquisadores sobre os fatores de aprendizagem, “afetividade” e “motivação” encabeçariam a lista. Esse fato deixa claro que, no âmbito do ensino, reconhece-se a grande influência que as variáveis afetivas exercem na construção do conhecimento dos estudantes. No entanto, na pesquisa escolar, a aprendizagem foi medida pelas conquistas acadêmicas nos aspectos cognitivos. (CHACÓN, 2003, p. 13)

Para a leitura e análise dos dados, foi necessário reconhecer que objetos estatísticos essenciais para essa etapa — a tabulação; o cálculo da média e sua interpretação; e a questão da variabilidade — permeiam distintamente, por meio de pontuações, em cada turma de 5ª Série, ou 6º ano (instituições pública e particular). A partir das medidas calculadas e interpretadas, verificamos a importância das medidas descritivas, incidindo nas idéias do Teorema Central do Limite.

Nesse sentido, a pesquisa de Méndez (1991) identifica quatro propriedades básicas para a compreensão do teorema:

- 1) A média da distribuição amostral é igual à média da população e igual à média de uma amostra quando o tamanho da amostra tende ao infinito;
- 2) A variância da distribuição amostral é menor do que a variância da população ($N > 1$);
- 3) A forma da distribuição amostral tende a ser aproximadamente normal à medida que se aumenta o tamanho da amostra, independentemente à forma da distribuição da população, isto é, normal ou não;

- 4) A forma da distribuição amostral cresce na altura e decresce na dispersão à medida que o tamanho amostral cresce.

Essas propriedades servirão como base para constituir o pensamento estatístico por meio das ideias construídas acerca do Teorema Central do Limite, por parte dos licenciandos. É uma área da nossa investigação que diz respeito à amostragem, aleatoriedade e métodos estocásticos. E, para finalizar, elaboramos uma parte específica para que os licenciandos percebam o comportamento do teorema, produzindo uma definição “particular”, construída por cada um dos licenciandos, a partir de suas próprias conclusões vivenciadas pelas atividades propostas.

6.2.4 Análise das Atividades 2 e 3

Para realização das atividades 1 e 2, não foi possível contar com um dos participantes, CM. Foi-nos permitido analisar apenas três tipos de respostas.

ATIVIDADE 2: Observe os dados²⁹ referentes ao índice de autoestima, variando de 0 a 40, em relação à Matemática, de alunos da 5ª série, ou 6º ano, do Ensino Fundamental de escolas particular e pública:

Objetivos:

- Identificar e interpretar medidas de tendência central;
- Reconhecer a variabilidade num conjunto de dados e interpretá-la.

A partir dos dados, que está variando de 0 a 40, em que 40 corresponde ao maior índice, e zero, ao mais baixo, de alunos da 5ª Série, ou 6º ano, do Ensino Fundamental, de Instituições Pública e Privada, escolha uma turma de cada tipo de instituição e responda às questões abaixo:

²⁹ Anexo 4.

- 1) Determine a média dos valores observados em cada uma das turmas escolhidas. O valor encontrado é um bom representante do conjunto de dados? Justifique.
- 2) Determine o intervalo limitado por $[\bar{x} - s; \bar{x} + s]$. Qual a proporção de elementos observados nesse intervalo, em cada um dos casos?

Vejamos como cada um respondeu a esses itens:

RESPOSTA DE AO PARA ATIVIDADE 2:³⁰

1) Turma escolhida: 5B – Rede particular: $\bar{x} = 30,68$ e $\tilde{x} = 32,75$

5B – Rede pública $\bar{x} = 22,35$ e $\tilde{x} = 22,25$

A média encontrada na 5B (particular) não é bem representativa; já que, no total de 31 alunos, 20 alunos estão acima da média, por isto, há uma discrepância de dados visto que a menor nota da turma é 15,5.

Na rede pública, a média é representativa na pesquisa visto que a medida de tendência central aproxima do valor central do rol.

2)

Particular:

5A [28,01965; 39,76295] Proporção: $\frac{17}{23} = 73,92404\%$

5B [24,69193; 36,66287] Proporção: $\frac{21}{31} = 67,74193\%$

5C [27,51467; 34,77113] Proporção: $\frac{16}{28} = 57,14285\%$

5D [20,40617; 37,17083] Proporção: $\frac{17}{26} = 65,38461\%$

5E [18,34834; 40,29446] Proporção: $\frac{11}{14} = 78,57142\%$

Pública:

5A [9,35293; 26,01067] Proporção: $\frac{16}{22} = 72,72727\%$

5B [13,16677; 31,54283] Proporção: $\frac{18}{31} = 58,06451\%$

5C [14,99624; 29,94116] Proporção: $\frac{20}{32} = 62,5\%$

5D [9,80303; 23,95897] Proporção: $\frac{13}{21} = 61,90476\%$

5E [10,24619; 26,18721] Proporção: $\frac{22}{30} = 73,33333\%$

³⁰ Anexo 5.

RESPOSTA DE **BM** PARA ATIVIDADE 2³¹:

1) Turma escolhida: 5D (particular): $\bar{x} = 28,79$

5D (pública): $\bar{x} = 16,88$

Analisando os dados, vemos uma amplitude grande (na escola particular o maior valor é 40 e o menor, 9; na pública, a maior é 30,5 e a menor 5).

Isto pode deslocar a média e fazer com que ela não seja representativa. É preciso analisar juntamente com outras medidas de tendência central, como por exemplo a mediana.

Na escola particular, a turma tem $\bar{x} = 28,79$ e $\tilde{x} = 30,5$. Na escola pública: $\bar{x} = 16,88$ e $\tilde{x} = 17,0$.

Notamos que os valores estão próximos, logo podemos considerar que a média é representativa nos dois casos.

2) 5D: particular: [20,40617; 37,17083] Proporção: $\frac{17}{26} = 65,3846\%$

5D: pública: : [9,80303; 23,95897] Proporção: $\frac{13}{21} = 61,90476\%$

RESPOSTA DE **FA** PARA ATIVIDADE 2³²:

1) As turmas escolhidas foram 5C particular e 5C pública. A turma 5C particular apresentou $\bar{x}$ de 31,14, enquanto a 5C da pública apresentou $\bar{x}$ de 22,97. As médias encontradas estão representando bem o conjunto de dados, uma vez que não fugiu muito do valor da mediana de cada turma, que foi no 5C, particular, 31 e na pública, 23.

2) Particular: 5C [27,51467; 34,77113] Proporção: $\frac{16}{28} = 57,14285\%$

Pública: 5C [14,99624; 29,94116] Proporção: $\frac{20}{32} = 62,5\%$

Analisando as respostas de AO, BM e FA, nos foi possível constatar que AO não distinguiu a média da mediana verbalmente, indicando, apenas, simbolicamente,

³¹ Anexo 6.

³² Anexo 7.

o que não nos permitiu identificar se realmente conhece e distingue efetivamente as medidas centrais. Por outro lado, após a escolha das turmas, pública e privada, na questão 2, apresentou os intervalos da diferença e soma entre a média e desvio padrão de cada turma, pela apresentação da proporção existente em cada um deles, isto é, não selecionou apenas uma turma para cada instituição, mas fez para todos. Sua conclusão foi a de que a média do índice de autoestima dos alunos da escola particular, para a turma escolhida, não foi representativa, a partir da leitura sobre a amplitude total, isto é, reconhecida como discrepância entre as medidas; já a média da escola pública foi representativa. Essa análise feita por AO parece ter sido influenciada pela variabilidade que ocorre na distribuição de dados. Por outro lado, percebemos que AO esteve mais atento às técnicas (τ) que deveriam ser aplicadas na atividade proposta, o que, no entanto, não ocorreu a justificativa para as mesmas, isto é, houve a ausência da tecnologia (θ).

No caso da atividade de BM, averiguamos que, além de utilizar as técnicas (τ) no sentido de encontrar os valores de medida central, como a média e a mediana, ocorreu a tecnologia (θ) por meio da justificativa de que existem outras medidas de tendência central, chamando atenção para leitura conjunta entre a média e a mediana. O item 2 da atividade 2 foi resolvido apenas para as turmas que BM escolheu. Não incorreu numa análise mais aprofundada da questão.

O estudante FA respondeu similarmente à BM no tocante à primeira questão. Aplicou as técnicas (τ) necessárias para o cálculo da média e mediana e, além disso, explicou e justificou que as medidas estão bem próximas nos dois casos. Na segunda questão, fez também apenas para o intervalo das turmas escolhidas.

Diante desses resultados, percebemos como o enunciado interfere na execução da tarefa. No item dois da Atividade 2, foram solicitados apenas o intervalo e a proporção dos índices de autoestima existente neles, entre a diferença e soma da média e do desvio padrão. Nenhum dos três manifestou interesse em justificar os intervalos, até mesmo porque não estava explícito. Portanto, essa análise não apenas nos proporcionou uma reflexão sobre a importância de elaborarmos uma atividade com mais clareza, sem criar expectativas do que os alunos poderiam ter respondido, mas também nos indicou uma limitação para a nossa investigação.

ATIVIDADE 3:

Objetivos:

- Construir o significado de aleatoriedade e amostragem;
- Conhecer o método estocástico para construir uma distribuição de amostras aleatórias;
- Desenvolver habilidades para interpretação de gráficos estatísticos;
- Analisar os resultados obtidos de média e desvio padrão;
- Obter os valores de média e desvio padrão para um conjunto de dados e comparar esses resultados com os obtidos anteriormente;
- Construir outros bancos de dados;
- Construir novos gráficos e comparar com os já construídos;
- Reconstruir uma definição para o Teorema Central do Limite.

Siga as instruções e responda ao que se pede:

- 1) Determine a média, $\bar{x}$, e o desvio padrão, s , para os dados da Instituição Pública e Instituição Particular.
- 2) Sorteie 20 elementos aleatórios de cada instituição e calcular a média, $\bar{x}$, e o desvio padrão, s . Repita o processo 30 vezes.
- 3) Construa um banco de dados com todos os resultados obtidos, isto é, reúna todos os dados do grupo.
- 4) Construa um gráfico que represente as médias observadas. Faça o mesmo para os valores de s , desvio padrão. O que você observa?
- 5) Determine a média dos valores obtidos em cada item (coluna) do banco de dados do item anterior, (4). O que você observa? Compare com o obtido no item 1.
- 6) Considere um novo banco de dados construído com os resultados dos demais colegas. O que você observa a partir dos resultados obtidos para média e desvio padrão?
- 7) Construa o gráfico desse novo banco de dados, isto é, do item 6, e determine a média e o desvio padrão.
- 8) Compare com o obtido no item 4. O que você observa?

A realização dessa atividade contou com uso de *Microsoft EXCEL* e do software *SPSS* de posse dos participantes, para análise estatística.

RESPOSTAS DE **AO** PARA ATIVIDADE 3³³:

(continua)

1) Determine a média, $\bar{x}$, e o desvio padrão, s , para os dados da Instituição Pública e Instituição Particular.

Resposta: *Para o número total de 136 alunos da instituição pública, a média do nível de autoestima é 19,8676 e desvio padrão, 8,32. E, para 122 alunos da instituição particular, tem-se que a média do nível de autoestima é 30,832 e o desvio padrão, 6,96.*

2) Sorteie 20 elementos aleatórios de cada instituição e calcular a média, $\bar{x}$, e o desvio padrão, s . Repita o processo 30 vezes.

Resposta: *Pelos dados aleatórios extraídos do EXCEL, a partir da amplitude total de cada uma das instituições, nos foi possível anotar as seguintes médias e seus respectivos desvios padrão, conforme anexo.*

3) Construa um banco de dados com todos os resultados obtidos, isto é, reúna todos os dados do grupo.

Resposta: *Construído no EXCEL.*

4) Construa um gráfico que represente as médias observadas. Faça o mesmo para os valores de s , desvio padrão. O que você observa?

Resposta: *Analizando as médias das 30 amostras aleatórias entre as escolas pública e particular, foi possível observar que há maior variação nas médias da escola pública, enquanto na escola particular os valores estão mais próximos.*

Quanto aos gráficos dos desvios-padrão entre as escolas pública e particular, são bem similares.

5) Determine a média dos valores obtidos em cada item (coluna) do banco de dados do item anterior, (4). O que você observa? Compare com o obtido no item 1.

³³ A atividade 3 do licenciando AO se encontra em sua íntegra no anexo 8.

(conclusão)

Resposta: *A média das médias da escola pública foi de 20,24, enquanto a da escola particular foi de 22,09. Já o desvio padrão das médias da escola pública foi 2,74 e da escola particular, 2,35. Neste sentido, podemos concluir descritivamente que a média das médias da escola particular é mais representativa, já que o coeficiente de variação é de 0,1064 e o da escola particular, 0,1354, apesar de que a média em ambos casos seja significativas.*

6) Considere um novo banco de dados construído com os resultados dos demais colegas. O que você observa a partir dos resultados obtidos para média e desvio padrão?

Resposta: *Foi possível observar que os valores foram bem próximos daqueles encontrados anteriormente.*

7) Construa o gráfico desse novo banco de dados, isto é, do item 6, e determine a média e o desvio padrão.

Resposta: *Os gráficos construídos para este item foram realizados no programa SPSS, cujas médias e desvios padrão das escolas pública e particular estão no anexo. Foi possível observar a média das médias e o desvio padrão das amostras, realizada no Excel: Escola Pública: Média das médias 19,8565 e desvio padrão das amostras: 2,257985; Escola Particular: média das médias: 22,47733 e desvio padrão, 2,061976*

8) Compare com o obtido no item 4. O que você observa?

Resposta: *Comparando as médias das escolas pública e particular, neste caso com os dados aleatórios, 600 variáveis, das 30 amostras, para escola pública, cuja média foi 19,59 e o desvio padrão, 10,41, e o mesmo número de variáveis para escola particular, cuja média foi 22,09 e o desvio padrão, 10,16, foi possível constatar que as médias têm seus valores bem próximos enquanto os desvios padrões diferem-se entre si, sendo que os das amostras são bem menores que os da população.*

Analisando as respostas de AO, percebemos que o estudante executou a tarefa (T) conforme uma sequência didática, indo além das técnicas (τ) ao solicitar softwares, tais como EXCEL e SPSS, para facilitar os cálculos e gerar os gráficos. A

organização de como AO executou a tarefa (T) também nos chamou atenção, pois foi construindo, passo a passo, as tarefas de modo que suas conclusões foram plausíveis de entendimento. Quanto à tecnologia (θ), podemos dizer que AO explicou e justificou no item (8), quando concluiu que os valores das médias das amostras é bem próxima da população aleatória construída no item (2) para os dois tipos, público e particular. Um fato extraordinário nos chamou atenção: a constatação, de AO, de que o desvio padrão das médias das amostras é bem menor que a média da população. Tal constatação resultou na justificativa da justificativa, isto é, culminou na teoria (Θ), mesmo não registrando como o Teorema Central do Limite, o que deixou para uma outra etapa da investigação em que ocorreu uma discussão acerca das atividades realizadas.

Vejamos como BM executou a atividade 3:

RESPOSTAS DE **BM** PARA ATIVIDADE 3³⁴:

(continua)

1) Determine a média, $\bar{x}$, e o desvio padrão, s , para os dados da Instituição Pública e Instituição Particular.

Resposta: *Pública: $\bar{x} = 19,8676$ e $\sigma = 8,3209$.*

2) Sorteie 20 elementos aleatórios de cada instituição e calcular a média, $\bar{x}$, e o desvio padrão, s . Repita o processo 30 vezes.

Resposta: *Anexo, resolvido no programa SPSS.*

3) Construa um banco de dados com todos os resultados obtidos, isto é, reúna todos os dados do grupo.

Resposta: em branco.

4) Construa um gráfico que represente as médias observadas. Faça o mesmo para os valores de s , desvio padrão. O que você observa?

Resposta: *As amostras foram de tamanhos 5, 10, 20, 25 e 30. À medida que o tamanho da amostra aumenta, o gráfico vai aproximando da curva normal.*

³⁴ A atividade 3 foi realizado no programa SPSS, cujo relatório se encontra no anexo 9.

(conclusão)

5) Determine a média dos valores obtidos em cada item (coluna) do banco de dados do item anterior, (4). O que você observa? Compare com o obtido no item 1.

Resposta: *À medida que aumentamos o tamanho da amostra, a média vai se aproximando da média da população e o desvio padrão vai diminuindo.*

6) Considere um novo banco de dados construído com os resultados dos demais colegas. O que você observa a partir dos resultados obtidos para média e desvio padrão?

Resposta: *Os outros chegaram à mesma conclusão.*

7) Construa o gráfico desse novo banco de dados, isto é, do item 6, e determine a média e o desvio padrão.

Resposta: deixou em branco.

8) Compare com o obtido no item 4. O que você observa?

Resposta: *O teorema central do limite.*

BM efetivou a atividade por meio do software *SPSS*, não dispôs do *EXCEL*. Os relatórios gerados pelo programa foram devidamente escaneados e anexados neste trabalho (anexo 10). Quanto às respostas de BM, no item 1, ele descreveu apenas os resultados da média e desvio padrão para escola pública. Para escola particular, o relatório também foi realizado, mas não indicado, porém foi apresentada também a representação das duas distribuições, da escola pública e da escola particular, ordenada em *ramo e folha* e em *Box-Plot*. O que não foi descrito na atividade foram os dados da escola particular. Então, dos 122 alunos que passaram pelo teste, o valor da média do índice de autoestima foi de 30,8320 e o desvio padrão, 6,96. No item 2, BM deixa indicado um relatório extraído do *SPSS*, sem indicar os números aleatórios e nem explicitar como encontrou esses dados. Para o item 3, não indicou como reuniu todos os dados aleatórios, mas faz relação com o item 4, já que a opção de BM foi a de selecionar diferentes números das amostras para as médias. Isto é, ele optou por escolher tamanhos de amostras diferentes e observar o comportamento dos respectivos gráficos para os tamanhos de 5, 10, 20,

25 e 30. Nesse caso, concordamos com Chevallard, Bosh e Gascón (2001), quando afirmam:

O momento do *primeiro encontro* faz referência aos objetos matemáticos que constituem um tipo de problema; o momento *exploratório*, relaciona um determinado tipo de problema com a construção de uma técnica adequada para abordá-los; o momento do *trabalho da técnica* se refere ao domínio, precisão e nova criação de técnicas matemáticas; o momento *tecnológico-teórico* faz referência, como seu nome indica, aos dois níveis de justificativa da prática matemática; e os momentos *institucionalização* e *avaliação* se referem, finalmente, à obra matemática em seu conjunto. (CHEVALLARD; BOSCH; GASCÓN, 2001, p. 276)

Assim, podemos dizer que, quanto ao trabalho da técnica (τ), BM pareceu ser criativo, já que conhecia alguns procedimentos sobre Inferência Estatística. Além disso, os mesmos autores reiteram que “essa descrição subjaz um princípio democratizador” (CHEVALLARD; BOSCH; GASCÓN, 2001, p. 276). Outro fato que nos chamou atenção nas atividades realizadas por BM foi, talvez, um “não se importar” muito com as justificativas. Há ênfase para técnica, mas a tecnologia, de certa forma, ficou comprometida. Por isso, será relatado posteriormente como ocorreu o diálogo no final de todas as atividades entre todos os sujeitos envolvidos na investigação: os pesquisadores, particularmente o professor-pesquisador, e os licenciandos.

Para o item 5, BM responde, antecipadamente, o mesmo que esperava para o item 8. Já no item 8, faz menção ao teorema. No item 6, comenta que sua resposta é igual à dos outros. Contudo, no item 7, a tarefa foi realizada no *SPSS*, somente para os dados da escola pública, cujos relatórios se encontram nos anexos. Porém, vejamos agora as médias e os desvios padrões encontrados para os diferentes tamanhos de amostras, conforme a tabela abaixo:

TABELA 3
Atividade realizada por BM para o item 7.

Tamanho das amostras das médias	Média	Desvio Padrão
n = 5	18,70	1,58
n = 10	19,9	1,28
n = 20	19,71	1,28
n = 25	19,80	1,60
n = 30	19,73	1,53

Fonte: Dados da pesquisa.

Para BM ter concluído que a média das amostras é igual ao da população inicial e o desvio padrão é menor, à medida que aumenta o tamanho da amostra, vale retomar o valor numérico da média de índice de autoestima de todos os 136 alunos da escola pública, que foi 19, 8676 e o desvio padrão, 8,32. Esses dados proporcionaram a BM chegar à conclusão de que tratava-se do Teorema Central do Limite sem recorrer aos dados aleatórios, como um todo, conforme AO tinha realizado.

Por fim, a resposta de FA:

RESPOSTAS DE **FA** PARA ATIVIDADE 3³⁵:

(continua)

1) Determine a média, $\bar{x}$, e o desvio padrão, s , para os dados da Instituição Pública e Instituição Particular.

Resposta: *Particular: $\bar{x} = 30,36$ e $\sigma = 6,06497$.*

2) Sorteie 20 elementos aleatórios de cada instituição e calcular a média, $\bar{x}$, e o desvio padrão, s . Repita o processo 30 vezes.

Resposta: Em branco.

3) Construa um banco de dados com todos os resultados obtidos, isto é, reúna todos os dados do grupo.

Resposta: Em branco.

4) Construa um gráfico que represente as médias observadas. Faça o mesmo para os valores de s , desvio padrão. O que você observa?

Resposta: *Construímos um gráfico para um tamanho de amostra de 5, 10, 20, 25 e 30 e, daí, observamos que, quanto maior o tamanho da amostra, mais a curva se aproxima da normal.*

5) Determine a média dos valores obtidos em cada item (coluna) do banco de dados do item anterior, (4). O que você observa? Compare com o obtido no item 1.

³⁵ As respostas, na íntegra, encontram-se no anexo 10.

(conclusão)

Resposta: *Quando vamos aumentando o tamanho da amostra, mais a média se aproxima da média da população e, ao contrário, o desvio padrão vai diminuindo.*

6) Considere um novo banco de dados construído com os resultados dos demais colegas. O que você observa a partir dos resultados obtidos para média e desvio padrão?

Resposta: Em branco

7) Construa o gráfico desse novo banco de dados, isto é, do item 6, e determine a média e o desvio padrão.

Resposta: Em branco.

8) Compare com o obtido no item 4. O que você observa?

Resposta: Em branco.

Analisando a resposta de FA, ela parece ter sido realizada concomitantemente com BM, já que os itens 2, 3, 6, 7 e 8 foram deixados em branco. E, além disso, FA registrou no item 1 o que BM não tinha feito, que é o registro da média dos dados iniciais na Atividade 3, para a escola particular com relação à média do índice de autoestima dos 122 alunos, conforme o material coletado³⁶ de BM. Outro fato que reitera essa afirmativa foi a técnica utilizada para o item 4, pois FA seguiu da mesma forma, diferindo de AO, ou seja, ele atribuiu tamanhos de amostras diferentes para as médias dos números aleatórios no que concerne aos índices de autoestima.

Contudo, percebemos algumas diferenças quanto às justificativas, pois a resposta do item 5 foi bem similar, apresentando características de como as técnicas (τ) utilizadas foram gradativamente levando à conclusão final, expressa somente oralmente para o item 8, pois FA deixa em branco, mas fala sobre sua conclusão, na próxima etapa de nossa investigação.

A partir dos dados registrados dos três licenciandos AO, BM e FA, podemos dizer que, de certa forma, essa análise atendeu a nossa hipótese, ao menos em parte, já que essa investigação não foi de cunho quantitativo, e, sim, qualitativo.

³⁶ Anexo 9.

Dessa maneira, poderíamos chegar à reconstrução do que vem a ser o conceito do Teorema Central do Limite.

Diante dos resultados obtidos das atividades realizadas, dizemos que a estrutura ecológica do saber e do didático foi contemplada, no sentido de que delimitamos os saberes indispensáveis para se chegar à definição do teorema. Outro item a ser destacado se refere aos saberes que efetivamente estão subjacentes aos objetos matemáticos e estatísticos, que nem sempre são explicitados. No entanto, significa que são elementos que permitem a “sobrevivência” do teorema em si, enquanto objeto do saber. Como exemplos, temos a leitura gráfica, a construção das ideias em relação à variabilidade de uma distribuição de dados, a centralidade da variável em jogo, entre outros.

Em contrapartida, os saberes “dispensáveis” pertinentes ao teorema, nos dias atuais, também foram constatados, porque não comprometeram a compreensão do mesmo. E um dos elementos do saber é precisamente a demonstração matemática do teorema durante o desenvolvimento dos saberes que estão no entorno dele. Isso não significa que podemos desvincular sua importância, principalmente sua abordagem histórica, já que devemos justificar por meio de argumentos plausíveis a necessidade de seu surgimento e do porquê de sua importância fundamental na Inferência Estatística. Para compreendermos a aplicabilidade do teorema nesse ramo da Estatística, precisamos nos recorrer a interferências, controladas ou não, no sentido de perceber a *literacia* estatística, por parte dos licenciandos.

6.3 O Diálogo: Intervenções Controladas no Design Didático

Os mesmos procedimentos de análise nos permitiram conduzir uma sequência de diálogos para perceber o *reconhecer*, *analisar*, *orientar* e *reconstruir*, conforme o esquema construído a partir dos fundamentos da pesquisa-ação de Thiollent (1986). Esclarecemos que os diálogos foram divididos em duas etapas. A primeira diz respeito à atividade um. Posteriormente, ocorreram intervenções para as atividades 2 e 3. Para a realização do diálogo, podemos caracterizar a primeira etapa como o do *reconhecer* os objetos matemáticos e estatísticos que estariam envolvidos. A segunda, a prática propriamente dita. Exporemos alguns diálogos que

ocorreram no segundo momento de nossa investigação, após as atividades realizadas. Chamaremos o professor-pesquisador de PE, e os licenciandos mantêm as mesmas nomeações anteriormente mencionadas. Os registros abaixo correspondem à transcrição da gravação ocorrida durante a discussão.

PE: O que vocês acharam das atividades?

AO: Achei interessante, mas bastante cansativo quando tive que achar os números aleatórios.

BM: Eu, particularmente, segui o programa de SPSS e o próprio programa gerou os resultados.

FA: Eu aproveitei os dados de BM e realizei as minhas tarefas.

AO: Tive mais trabalho, porque optei por usar EXCEL e SPSS, simultaneamente.

Esse primeiro diálogo diz respeito à atividade como um todo. Vejamos como seguimos:

PE: Independentemente de ser trabalhoso ou não, o que vocês perceberam que sempre aparece para ser calculado?

AO: A média e o desvio padrão.

BM: É mesmo...

FA: Também observei isso!

É interessante notar que os licenciandos observaram as presenças da média e do desvio padrão, o que não significa que tenham percebido as suas devidas importâncias para o teorema. Por conseguinte, entendemos o quanto é importante a Inferência Estatística, mas, no nosso caso, o Teorema Central do Limite serve como base para inúmeros trabalhos que envolvem distribuição de probabilidades e curva Normal. E uma das grandes vantagens de se trabalhar com o teorema é que uma “[...] distribuição pode ser aproximada normal de probabilidade sem importar a origem dos dados iniciais. A distribuição normal de probabilidade equivale à função erro de Laplace.” (SALSBURG, 2009, p. 80-81).

Essa vantagem ainda é corroborada pelo número de parâmetros a ser considerado, isto é, precisa-se conhecer apenas a média, não necessariamente a da

população, mas ao menos de um número significativo de amostras, e do desvio padrão, que também pode ser desconhecido, a partir do momento em que são consideradas as proporções. Ou seja, basta trabalhar com dois parâmetros da Estatística: a média e o desvio padrão das amostras, conquanto o número de amostras seja significativo. Em geral, isso simplifica o trabalho dos estatísticos, e, além disso, suas implicações estendem para importantes testes que sustentam as pesquisas quantitativas em diversas áreas do conhecimento.

Para que os licenciandos compreendessem, efetivamente, as vantagens de se trabalhar com poucos parâmetros, foi preciso a intervenção do professor-pesquisador:

PE: A grande vantagem de trabalhar com esses dois parâmetros, a média e o desvio padrão, é que podemos tirar algumas conclusões, sem estender para outros tipos de parâmetros, tais como os quartis, outras medidas de dispersão, assimetria ou o grau de curtose. Assim, estaremos simplificando o modo de como podemos tirar algumas conclusões. Então, o que vocês podem concluir a partir das médias e dos desvios padrões dos dados apresentados, isto é, dos 132 alunos da escola pública e dos 122 alunos da particular?

Após essa intervenção, BM foi o primeiro a desejar manifestar sua opinião:

BM: A média dos alunos da escola pública é bem menor que o da escola particular. Parece que a autoestima dos alunos em relação à Matemática é mais baixa quanto menor o nível econômico, é isso?
AO: Pode ser, mas acho que temos que observar alguns dados individuais, já que a média leva em conta todos os valores, inclusive os extremos.
FA: Mesmo levando em conta todos os valores, não podemos desprezar esse parâmetro, pois nas atividades seguintes, com os números aleatórios, observamos que acontece a mesma coisa.

Nessa etapa do diálogo, podemos perceber que os licenciandos “ultrapassam” os conhecimentos específicos do teorema e estendem para uma leitura mais crítica dos dados. Inclusive AO alerta para o fato de que nem sempre a

média é a melhor medida central a ser considerada, chamando atenção para os valores extremos.

Vejamos como eles continuam essa discussão:

AO: Outro dado que observei foi de que a maior média das turmas da escola pública consegue ser menor que a menor média da escola particular, vocês observaram isso?

BM: Eu não tinha observado isso!

PE: Então, parece que isso faz sentido do porquê dessa atividade para vocês, afinal, vocês estarão atuando tanto na escola pública como na escola particular. Essa baixa autoestima dos alunos da escola pública com relação à disciplina Matemática poderá ser revertida se vocês atuarem de forma que rompa com os mitos e as crenças existentes em torno da Matemática, não é? Mas agora vamos voltar para o nosso foco. O que vocês observaram quando verificaram que o desvio padrão das amostras é menor que o desvio padrão da população, ao menos AO, fez amostras aleatórias para 600 alunos da escola pública e 600 para alunos da particular, não foi?

Apesar da importância sobre a leitura dos dados, nesse instante, o objetivo estava centrado no número de vezes em que as amostras aleatórias foram geradas. Ulteriormente, no seu devido tempo, essa questão será retomada.

FA: Todos que foram construídos a partir de amostras foram menores até que os dados originais. Mesmo fazendo com tamanho de amostras diferentes.

BM: Isso realmente aconteceu.

AO: Então eu perdi tempo construindo todas as 30 amostras de 20 elementos aleatórios?

PE: Claro que não! Foi, inclusive muito interessante, já que você optou por extrair todos do EXCEL. Assim, nos foi possível inclusive discutir sobre isso! Mas o que será que está por “trás” disso? Isto é, de quais objetos matemáticos e estatísticos vocês precisam para que isso ocorra?

A construção de dados aleatórios foi bastante providencial, porque permitiu que os alunos percebessem como o teorema faz sentido na medida em que se aumenta o número de amostras. Em seguida, o pesquisador-professor indagou sobre que tipos de saberes estão implícitos nas atividades propostas. Vejamos o que eles falaram:

BM: A partir dos dados coletados, extraímos a média e o desvio padrão de cada instituição. Depois atribuímos o tamanho das amostras extraídas aleatoriamente a partir das 30 amostras de 20 elementos cada. Os objetos matemáticos eu não sei, mas os estatísticos tem a ver com a média, o desvio padrão e o tamanho da amostra.

FA: Não seria o objeto matemático a divisão que faríamos quando desejamos encontrar o desvio padrão da amostra? Por que precisávamos exatamente dos três dados que BM falou?

AO: A manipulação no EXCEL e no SPSS também podem ser considerados objetos matemáticos e estatísticos?

PE: FA e BM, vocês podem responder para AO sua pergunta?

BM: Acho que sim, neste caso vejo apenas objetos estatísticos, tais como média e o desvio padrão. Ainda bem que FA falou da idéia do quociente como objeto matemático. Interessante como tudo tem matemática.

Durante essa discussão, podemos auferir que os conhecimentos prévios emergem a partir da própria experiência em *saber-fazer* de cada um. A indagação, por parte do professor-pesquisador, de quais objetos matemáticos e/ou estatísticos podem ser apontados a partir das atividades realizadas representa uma intervenção controlada, isto é, uma orientação dentro de um contexto pelo qual o pesquisador deseja fazer com que os alunos reconstruam seus próprios conhecimentos.

Nessa abordagem, os licenciandos ainda não tinham esgotado todos os objetos matemáticos observados, já que ficaram um pouco “inseguros” quanto aos elementos que poderiam ser apontados.

FA: Mas quando escolhemos um determinado item a ser resolvido no software, não estamos escolhendo, de certo modo, uma função matemática?

PE: Isso mesmo! Os objetos matemáticos, principalmente, nem sempre estão explícitos. Muitas vezes aparecem, mas não no primeiro nível, eles são subjacentes ao que desejam encontrar. Agora, qual foi a conclusão unânime entre vocês, sobre a atividade 3?

AO: Eu entendi que mesmo não sendo simétrico, se aumentarmos o tamanho da amostra, a média vai ser cada vez mais próxima da média da população e o desvio padrão é sempre menor que o da população. Mas sempre cai numa curva normal?

FA: Tudo indica que sim!

Nesse ponto podemos dizer que chegamos à definição do Teorema Central do Limite, construída a partir de uma situação-problema, inicialmente apresentada com um número limitado de observações. O pensamento e o raciocínio estatístico estão presentes na leitura já que eles chegam a conclusões acerca do comportamento dos números calculados, para a média e para o desvio padrão.

No entanto, a representação gráfica também foi uma forte aliada para a compreensão do teorema.

BM: Pelo experimento de AO, parece que sim. Cada vez que aumentamos o tamanho da amostra, os histogramas que construímos no SPSS indicou que seria normal.

PE: Vamos dizer que aproximadamente normal, tudo bem? Mas será que existe um modelo matemático que descreva esse comportamento?

BM: É o teorema central do limite! Eu já tinha estudado isso antes.

AO: Eu não sabia que era um teorema.

FA: E eu não sabia que tinha um modelo que fizesse que tal coisa acontecesse!

A partir dos histogramas gerados pelos dados das amostras, eles chegaram à definição do teorema, cada um a seu modo, mostrando-nos o quanto são diferentes as percepções que cada um tem sobre as ideias e os conceitos no entorno de um teorema. A conclusão de que um teorema gera um comportamento “padrão”, causa, muitas vezes, surpresa aos alunos. Contudo, nos atentamos para o contexto em si,

de uma situação que ocorre no dia a dia dos professores de Matemática. A autoestima dos alunos acaba por interferir na aprendizagem nessa disciplina. São reflexões pertinentes para licenciandos em Matemática, pois farão parte de sua realidade quando estiverem em exercício do magistério.

A interação entre um tema fundamental para a Inferência Estatística e a prática de sala de aula causa também motivação para todos os sujeitos envolvidos. Assim, para encerrar o diálogo, o desfecho foi sintetizar os objetos matemáticos e estatísticos que estiveram presentes no estudo do Teorema Central do Limite.

PE: Então, tudo que realizamos foi para chegar a essa conclusão: existe um modelo matemático chamado o Teorema Central do Limite, pois, a partir desse teorema, podemos dizer que, mesmo que uma curva seja assimétrica, se aumentarmos o tamanho das amostras, a curva tenderá cada vez mais aproximar-se da curva Normal. E, segundo os autores de livros-texto, trabalhar com curva Normal tem suas vantagens, principalmente porque precisamos de usar apenas dois parâmetros: a média e o desvio padrão. Infelizmente, a demonstração desse teorema demandará muito tempo e, além disso, é bastante complexo, não desmerecendo habilidade de vocês com os objetos matemáticos, mas, por enquanto, não é o nosso foco de estudo, e sim, as implicações que esse teorema tem no dia a dia. Vocês poderiam citar outros exemplos?

AO: Deve ter bastante exemplos. Mas agora nada me vem em mente.

FA: Todas as distribuições que podem ser reduzidas a uma curva normal, acho que aplicamos o teorema, não é?

BM: Deve ter um monte de exemplos, mas podemos ficar por aqui, não?

PE: Tudo bem. Mas deixo como lição de casa uma reflexão sobre aplicação do teorema, tudo bem?

Para o encerramento da discussão, foi preciso um elo para que os licenciandos pudessem pensar sobre que tipos de situações-exemplo poderiam estar associados ao teorema. A partir da análise dessa discussão, parece que o nosso propósito foi atingido, mesmo considerando tão poucas observações. Porém, a ideia central era que eles percebessem o quanto era necessária a existência do teorema, principalmente na Inferência Estatística.

Outro aspecto relevante foi a abordagem ecológica subjacente nos diálogos, pois, de posse do *fazer-produzir*, eles se apropriaram do saber por diferentes técnicas (τ), contemplando as justificativas e explicações acerca do que entenderam, tendo em vista que, durante as atividades, nem sempre foi possível detectar nos registros escritos. Isso porque, nas discussões, os *insights*³⁷ ocorreram mais naturalmente, após as atividades realizadas.

Outro fator que podemos destacar é o que diz respeito à *literacia* estatística, reiterando Scheaffer, Watkins e Landwehr (1998), que apontam os elementos que devem ser explorados no ensino de Estatística: senso numérico; compreensão de variáveis; interpretação de tabelas e gráficos; aspectos de planejamento de pesquisa ou experimento; processos de análise de dados; relação entre probabilidade e estatística; raciocínio inferencial. E, dentro dessa abordagem, chamamos atenção para *literacia* estatística, que se fez presente durante todos os argumentos utilizados pelos licenciandos durante as atividades. Segundo Walmann (1993), a *literacia* estatística é uma:

[...] habilidade para compreender e avaliar criticamente os resultados estatísticos que permeiam as informações do nosso dia a dia – em conjunto com a habilidade de apreciar as contribuições que o pensamento estatístico nos proporciona a tomar decisões e agir nas dimensões públicas e privadas, profissionais e pessoais. (WALMANN, 1993, p. 1) (Tradução nossa)³⁸

Um outro ponto que não deve ser despercebido, já que o nosso trabalho está direcionado aos licenciandos de Matemática, é no tocante à formação do professor. O contexto no qual as atividades 2 e 3 foram elaboradas tem relação direta com essa linha de pesquisa, que é o da formação continuada dos professores, apesar de não haver ênfase para tal, até mesmo porque o nosso objetivo está bem delimitado numa abordagem ecológica do saber e do didático. Em contrapartida, a autoestima dos alunos da 5ª Série, ou 6º ano, do Ensino Fundamental, nos propiciou alguns momentos de reflexão sobre esse assunto. Assim, segundo Chevallard, Bosch e Gascón (2001):

³⁷ Entendemos como *insight* a habilidade de discernir sobre o assunto a ser abordado.

³⁸ “Statistical Literacy” is the ability to understand and critically evaluate statistical results that permeate our daily lives – coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions.” (WALMANN, 1993, p. 1).

[...]: não há momentos “nobres” e momentos “menos nobres”, como também não há momentos “mais matemáticos” e momentos “mais didáticos”. O episódio da aula de prática e os comentários didáticos subsequentes evidenciam a importância crucial de um dos momentos mais desprestigiados – o momento do trabalho da técnica – e a necessidade de que essa dimensão do processo de estudo seja aceita nos dispositivos didáticos escolares. (CHEVALLARD; BOSCH; GASCÓN, 2001, p. 276)

Portanto, todo momento foi privilegiado, sem diminuir ou aumentar outro, mas, desde já, ressaltamos a importância de focar o objeto do nosso tema, Teorema Central do Limite, e suas abordagens ecológicas tanto para o saber em si, quanto para o didático.

Para finalizar, talvez não seja de grande pretensão de nossa parte achar que esse trabalho vem preencher as lacunas existentes sobre a Inferência Estatística, muito pelo contrário: talvez seja uma forma de chamarmos atenção para que, efetivamente, os estudantes de pós-graduação, pesquisadores e outros interessados tenham um olhar mais atento para essa linha de pesquisa, que é a Educação Estatística, um campo que emerge da Educação Matemática como meio de estimular mais pesquisas.

7 CONSIDERAÇÕES FINAIS

Essa investigação surgiu a partir do desejo de conhecer um pouco mais sobre as questões que envolvem o Ensino da Estatística para os futuros professores de Matemática e, particularmente, de um objeto que atendesse tanto ao aspecto matemático quanto ao estatístico. O Teorema Central do Limite foi o tema providencial para a nossa escolha, por fazer parte da Inferência Estatística. Esse teorema é abordado, na maioria dos livros-texto para o Ensino Superior, normalmente com um comentário breve, porém são feitas afirmações do quanto o aprendizado dele é fundamental, sem que se esclareça o porquê de ser tão importante.

Deparamo-nos com diversos obstáculos, o que nos permitiu reconhecer o quanto somos limitados diante das intenções iniciais. Para conseguir realizar esse trabalho, iniciamos com quatro licenciandos, mas apenas três chegaram ao final de nossa investigação. Muitas vezes não contamos com esses percalços, entretanto, na realidade, eles existem e devemos saber como superar as nossas próprias limitações para atingir, ao menos, o caminho que nos projetamos para tal.

Entendemos o quanto é importante a Estatística no mundo atual, extremamente influenciada por meios de comunicação e tecnologia de última geração. De fato, não há como acompanhar os avanços tecnológicos num mundo ditado por uma gama de informações, que podem gerar até mesmo o sentimento de insegurança por parte de quem procura mais e mais informações. Portanto, seria pretensão de nossa parte finalizar as discussões aqui explicitadas, já que devemos entender que este trabalho deve servir como mais um meio de *sensibilizar* a classe de educadores matemáticos que lidam com o Ensino da Estatística no Ensino Superior e, em especial, os que fazem parte da formação dos futuros professores de Matemática que estarão atuando na Educação Básica. É um ciclo que não permite um início e não há um ponto final. Deve ser trabalhada, ininterruptamente, a importância da Estatística como leitura de mundo, na formação dos professores, no ato de inferir sobre as informações e saber efetivamente como os dados foram trabalhados até chegar ao conhecimento de todos.

Nesse contexto, não há como desconsiderar a importância desse objeto do saber como modelo matemático para explicar fenômenos da natureza, das Ciências

Sociais, das Ciências Biológicas, entre outras ciências. Apesar de o foco de nossa investigação ter sido no processo de ensino e de aprendizagem do teorema, vale ressaltar que é a partir da Educação que podemos acreditar numa mudança de paradigma, rompendo, assim, com crenças arraigadas de que a Matemática, e consequentemente, a Estatística, serve apenas para um pequeno grupo. Muito pelo contrário, o enfoque desse tema na Educação visa efetivamente à possibilidade de ampliar o conhecimento dos futuros professores, no sentido de que mudar significa, metaforicamente, despir de todo o *pré-conceito* e se permitir aceitar suas próprias limitações e desejar buscar meios que possam ser superados mediante o processo de estudo (CHEVALLARD; BOSCH; GASCÓN, 2001), a serviço do futuro, isto é, daqueles ou daquelas que estão diante das carteiras, sedentos pelo saber. É claro que a realidade não é bem assim. Existem outras variáveis a serem consideradas, tais como indisciplina, desmotivação, a desvalorização do professor como profissional, a estrutura particular de cada instituição de ensino, entre outras. Todavia, devemos acreditar ser um ponto de partida, mesmo com todos esses obstáculos.

Como a linha de nossa pesquisa está voltada para Educação Matemática, em particular, a Educação Estatística, não foi dada ênfase na demonstração do teorema. Isso é reforçado até mesmo na revisão da literatura e no estudo praxeológico dos livros-texto consultados. São poucos os casos em que há demonstração efetiva do teorema, até mesmo porque a sua demonstração só foi efetivamente concretizada no meio da turbulência, na primeira metade do século XX, entre as duas Grandes Guerras Mundiais. Segundo Salsburg (2009, p. 84), “por volta de 1934, o(s) teorema(s) central(ais) do limite não era(m) mais conjectura. Tudo que se tinha a fazer era provar que as condições de Lindeberg-Lévy³⁹ se mantinham. Então o teorema central do limite se sustenta, [...]”. Além disso, outros métodos foram inseridos de modo a simplificar os procedimentos utilizados em termos de ser definido tal como é apresentado nos livros-texto atualmente. Para os interessados na demonstração do teorema, sugerimos conhecer as condições de Lindeberg-Lévy, conforme anexamos neste trabalho.

³⁹ “Jarl Waldemar Lindeberg, da Finlândia e Paul Lévy, da França, descobriram de modo independente, um conjunto de condições sobrepostas necessárias para a conjectura se tornasse verdadeira”, e, neste caso, a conjectura central do limite passaria a ser o teorema central do limite. (SALSBURG, 2009, p. 84).

Outro fato importante a ser destacado é o da *literacia estatística*, pois ela tem sido alvo de muitas discussões recentemente na Educação Estatística. Como precursores dessa temática, têm-se os trabalhos de Iddo Gal, Joan Garfield, Robert delMas, Rosemary Callingham, Jane Watson, Katherine Walmann e outros mais que podemos encontrar, principalmente nos artigos publicados do SERJ e nos encontros internacionais referentes à Educação Estatística, tal como o ICOTS.

Portanto, as habilidades específicas da *literacia estatística* não podem ser desconsideradas, mas sim trabalhadas, principalmente no âmbito educacional, porque não devemos esquecer que, de certa forma, os educadores são também formadores de opinião. E, para que efetivamente tenhamos em mente pessoas que podem pensar e agir criticamente, devemos estar cientes das contribuições que a *literacia estatística* apresenta para a nossa classe de profissionais.

Nessa perspectiva, esperamos que este trabalho seja apenas um ponto de partida para outros que poderão iniciar no campo da Educação Estatística, no âmbito de incrementar as pesquisas voltadas para Inferência Estatística, principalmente no tocante aos elementos que dizem respeito ao processo de estudo, culminando no processo de ensino e de aprendizagem nos diferentes níveis de ensino. A pesquisa em si limitou-se à abordagem ecológica do saber e do didático, cujo objeto foram o Teorema Central do Limite e os sujeitos da pesquisa, os licenciandos e os pesquisadores, inclusive o professor-pesquisador, já que nos fundamentamos na metodologia da pesquisa-ação, apresentando um *design didático* para a estrutura desta tese. E, assim, esperamos que este trabalho seja uma contribuição, mesmo que ínfima, no campo da Educação Estatística, para os futuros pesquisadores que desejarem dar continuidade nesta linha de pesquisa em diferentes enfoques.

REFERÊNCIAS

ALMOULOUD, Saddo Ag. **Fundamentos da didática da matemática**. Curitiba: UFPR, 2007.

ALVARADO, H.; BATANERO, C. Designing a study process of the central limit theorem for engineers. In: **International Conference On Teaching Statistics**, 7, 2006. Salvador, Bahia, Brazil. Disponível em: <http://www.stat.auckland.ac.nz/~iase/publications/17/C404.pdf>. Acesso em: 26 mar. 2006.

ARTAUD, M. Introduction à l'approche écologique du didactique, L'écologie des organisations mathématiques et didactiques. **Actes de la neuvième École d'Été de didactique des mathématiques**. Houlgate: Bailleul, 1998, p.101-139.

ARTIGUE, Michèle. Ingénierie didactique. **Recherches en Didactique des Mathématiques**, v. 9, n. 3, p.281-308. Grenoble, 1988.

_____. Ingeniería didáctica. In: ARTIGUE, M.; MORENO, R.; GÓMEZ, L. **Ingeniería didáctica en educación matemática**. México: Grupo Iberoamérica, 1995, p. 33-59.

_____. Engenharia Didáctica. In: BRUN, Jean. **Didáctica das Matemáticas**. Trad. Maria José Figueiredo. Lisboa: Instituto Piaget, 1996, p.193-217.

_____. Proceedings of NORMA08. In: **Nordic Research in Mathematics Education**. Paris: C. Winslow Sense Publ., 2009.

BALACHEFF, Nicolas, **The researcher epistemology**: a deadlock for educational research on proof. n. 109. France: Les cahiers du laboratoire Leibniz, 2004.

BESSON, Jean-Louis. As estatísticas: verdadeiras ou falsas? In: BESSON, Jean-Louis (org.). **A ilusão das estatísticas**. São Paulo: Unesp, 1995, p. 25-67.

BLAIOTTA, Jimena; DELIEUTRAZ, Pablo. **Teorema central del limite**. Buenos Aires, 2004. Disponível em: http://www.union-matematica.org.ar/reunion_anual/anteriores/monografiatcl41.pdf. Acesso em: 02 fev. 2007.

BOYER, Carl B. **História da Matemática**. Trad. Elza F. Gomide. São Paulo: Blücher, 1974.

BOROVCHNIK, Manfred. **Probabilistic and statistical thinking**. 2005. Disponível em: <http://cerme4.crm.es/Papers%20definitius/5/wg5listofpapers.htm>. Acesso em: 19 set. 2008.

BRASIL, Ministério da Educação, Secretaria de Educação. **Parâmetros Curriculares Nacionais**: para o Ensino Fundamental. Brasília, 1998.

_____. Ministério da Educação, Secretaria de Educação. **INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira**. Disponível em: <http://www.inep.gov.br/internacional/pisa/> Acesso em: 30 ago 2009.

BROUSSEAU, Guy. Os diferentes papéis do professor. In: PARRA, Cecília; SAIZ, Irmã. **Didática da matemática**: reflexões psicopedagógicas. Porto Alegre: Artes Médicas, 1996, p. 48-72.

BRUTER, Claude-Paul. **Compreender as matemáticas**: as dez noções fundamentais. Trad. Luís Paulino Leitão. Lisboa: Instituto Piaget, 1998.

BUNCHRAFT, Guenia; CAVAS, Cláudio São Thiago. **Sob medida**: um guia sobre a elaboração de medidas do comportamento e suas aplicações. São Paulo: Vetor, 2002.

BUSSAB, Wilton de O.; MORETTIN, Pedro. A. **Estatística básica**. 5. ed. São Paulo: Saraiva, 2003.

CAJORI, Florian. **Uma história da Matemática**. Trad. Lázaro Coutinho. Rio de Janeiro: Ciência Moderna Ltda., 2007.

CAM, L. Lee. **The Central Limit Theorem around 1935**. 1986. Statistical Science 1, p. 78-96. Disponível em: http://projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.ss/1177013818. Acesso em: 06 abr. 2009.

CHACÓN, M. G. **Matemática emocional**: os afetos na aprendizagem Matemática. Porto Alegre: ArtMed, 2003.

CHARNAY, Roland. Aprendendo (com) a resolução de problemas. In: PARRA, C.; SAIZ, I. **Didática da Matemática**: reflexões psicopedagógicas. Porto Alegre: Artes Médicas, 1996, p. 36-47.

CHEVALLARD, Yves. **La transposition didactique**: Du savoir savant au savoir enseigné. Grenoble: La Pensée Sauvage, 1985.

_____. **La transposition didactique**. Grenoble: La Pensée Sauvage, 1991.

_____. Conceitos fundamentais da didáctica: perspectivas trazidas por uma abordagem antropológica. In: BRUN, Jean. **Didáctica das Matemáticas**. Trad. Maria José Figueiredo. Lisboa: Instituto Piaget, 1996, p.115-153.

_____. L'analyse des pratiques enseignantes en théorie anthropologique du didactique. **Recherches en Didactique des Mathématiques**. Grenoble: La Pensée Sauvage, v.12.1, 1999, p. 221-265.

CHEVALLARD, Y.; JOSHUA, M. A. **La transposition didactique**: du savoir savant au savoir enseigné. Grenoble: La Pensée Sauvage, 1991.

CHEVALLARD, Y; BOSCH, M.; GASCÓN, J. **Estudar matemáticas**: o elo perdido entre o ensino e a aprendizagem. Trad. Daisy Vaz de Moraes. Porto Alegre: Artmed, 2001.

CONTRERAS, José. **A autonomia de professores**. Trad. Sandra Trabucco Valenzuela. São Paulo, 2002.

COUTINHO, Cileda de Queiroz e Silva. **Introdução ao conceito de probabilidade**: uma visão frequentista. São Paulo: EDUC, 1996.

_____. **Introduction aux situations aléatoires dès le Collège**: de la modélisation à la simulation d'expériences de Bernoulli dans l'environnement informatique Cabri-géomètre II. 2001. Thèse (Docteur de Didactique des Mathématiques). Université Joseph Fourier, Grenoble, 2001.

DAJOZ, Roger. **Princípios de Ecologia**. Trad. Fátima Murad. 7 ed. Porto Alegre: Artmed, 2005.

D'AMORE, Bruno. **Elementos de didática da matemática**. Trad. Maria Cristina Bonomi. São Paulo: Livraria da Física, 2007.

DAVID, H. A. The history of statistics in the classroom. **Joint Statistical Meetings**. Iowa State University, 2003. Disponível em:
http://www.stat.auckland.ac.nz/~iase/publications/jsm/03_dadivpap.pdf Acesso em 06 abr. 2008.

DEVLIN, Keith. **Os problemas do milênio**: sete grandes enigmas matemáticos do nosso tempo. Trad. Michelle Dysmann. Rio de Janeiro: Record, 2004.

DIENES, Zoltan Paul. **An experimental study of mathematics-learning**. London: Hutchinson, 1963.

EVES, Howard. **Introdução à história da matemática**. Trad. Hygino H. Domingues. 2. ed. Campinas-SP: UNICAMP, 1997. (Coleção Repertórios)

FARIAS, Alfredo Alves; SOARES, José Francisco; CÉSAR, Cibele Comini. **Introdução à Estatística**. 2. ed. Rio de Janeiro: LTC, 2003.

FISCHER, Hans. **The central limit theorem from Laplace to Cauchy**: changes in stochastic objectives and in analytical methods. Germany: 2000. Disponível em: <http://mathsrv.ku-eichstaett.de/MGF/homes/didmath/seite/fischer.html>. Acesso em: 26 fev. 2006.

GAL, Iddo. Adults' statistical literacy: meanings, components, responsibilities. **Internacional Statistical Review**. 2002, 70(1), p. 1-25. Disponível em: www.stat.auckland.ac.nz/~iase/cblumberg/gal.pdf. Acesso em: 20 maio 2007.

GARFIELD, Joan; BEN-ZVI, Dani. How students learn statistics revisited: a current review of research on teaching and learning statistics. **International Statistical Review**. 2007, v. 75, n. 3, December/2007, p. 372-396.

GLENCROSS, Michael J. A practical approach to the central limit theorem. In: **International Conference On Teaching Statistics**, 2, 1986, University of Victoria, Canada. Disponível em: www.stat.auckland.ac.nz/~iase/publications/icots2/Glencross-1.pdf. Acesso em: 25 mar. 2008.

GNEDENKO B.V., KOLMOGOROV A.N. **Limit distributions for sums of independent random variables**. Translate: K. L. Chung. Massachusetts: Addison-Wesley Publishing Company, 1954.

GODINO, J. D. Un enfoque ontológico y semiótico de la cognición matemática. **Recherches en Didactique des Mathématiques**. v. 22, 2002, p. 237-284.

GRAHAM, Alan. **Developing thinking in Statistics**. London: Paul Chapman Publishing, 2006.

GUIMARÃES, Rui Campos; CABRAL, José A. Sarsfield. **Estatística**. Lisboa: McGraw-Hill, 1997.

HALD, Anders. **A history of mathematical and statistics**: from 1750 to 1930. New York: John Wiles & Sons, 1998. (Wiley Series in Probability and Statistics)

HOLMES, Peter. What sort of Statistics should be taught in schools – and why? In: LOUREIRO, C.; OLIVEIRA, F.; BRUNHEIRA, L. **Ensino e aprendizagem da Estatística**. Lisboa: Grafis, 2000.

JAMES, Barry R. **Probabilidade**: um curso em nível intermediário. 3. ed. Rio de Janeiro: IMPA, 2006.

JOSHUA, Samuel; DUPIN, Jean-Jacques. **Introduction à la didactique des sciences et des mathématiques**. 1. ed. Paris: Universitaires de France, 1993.

KLINE, Morris. **O fracasso da matemática moderna**. Trad. Leonidas Gontijo de Carvalho. São Paulo: Ibrasa, 1976.

LAKATOS, Imre. **A lógica do descobrimento matemático**: provas e refutações. Trad. Nathanael C. Caixeiro. Rio de Janeiro: Zahar, 1978.

LEITHOLD, Louis. **O cálculo com geometria analítica**. Trad. Cyro de Carvalho Patarra. 3 ed. São Paulo: Harbra, 1994.

LEVIN, Jack; FOX, James Alan. **Estatística para ciências humanas**. Trad. Alfredo Alves de Faria. 9 ed. São Paulo: Pearson Prentice Hall, 2004.

LUNSFORD, M.L.; ROWELL, G.H.; GOODSON-ESPY, T. Classroom research: assessment of student understanding of sampling distributions of means and the central limit theorem in Post-Calculus Probability and Statistics Classes. **Journal of Statistics Education**. v. 14, n. 3, 2006.

Disponível em: www.amstat.org/publications/jse/v14n3/lunsford.html. Acesso em: 10 dez. 2006.

MAGALHÃES, Marco Nascimento; LIMA, Antonio Carlos Pedroso de. **Noções de probabilidade e estatística**. 4. ed. São Paulo: Universidade de São Paulo, 2002.

MANN, Prem S. **Introdução à estatística**. Trad. Eduardo Benedito Curtolo; Teresa Cristina Padilha de Souza. 5. ed. Rio de Janeiro: LTC, 2006.

MEDEIROS, Cleides Faria de. Por uma educação Matemática como intersubjetividade. In: BICUDO, M.A.V. (org.). **Educação Matemática**. São Paulo: Centauro, 2005, p. 13-44.

MÉNDEZ, H. **Understanding the central limit theorem**. Tesis Doctoral. 1991. Universidad de California. UMI.

METHER, Max. **The history of the central limit theorem**. Mat-2.108 Sovelletun Matematiikan erikoistyöt, 2003. Disponível em: <http://www.sal.tkk.fi/Opinnot/Mat-2.108/pdf-files/emet03.pdf>. Acesso em: 26 fev. 2006.

MEYER, Paul L. **Probabilidade**: aplicações à Estatística. Trad. Ruy de C. B. Lourenço Filho. 2. ed. Rio de Janeiro: LTC, 1983.

MOORE, David S. **A estatística básica e sua prática**. Trad. Cristiana Filizola Carneiro Pessoa. 3. ed. Rio de Janeiro: LTC, 2005.

PAULASKAS, Vygantas. **Lindeberg's CLT in multidimensional and Banach spaces**. Vilnius University Naugarduko, Lithuania, 2006. Disponível em: <http://www.stat.fi/isi99/proceedings/arkisto/varasto/paul0312.pdf>. Acesso em: 19 nov. 2007.

RODRIGUEZ, Juan Antonio Argüelles. **Historia de la matematica**. Madrid: Akal, 1989.

SALSBURG, David. **Uma senhora toma chá...: como a estatística revolucionou a ciência do século XX**. Trad. José Maurício Gradel. Rio de Janeiro: Zahar, 2009.

SENETA, E. The central limit problem and linear least squares in pre-revolutionary Russia. The background. **Mathematical Scientific**. v. 9, 1984, p. 37–77.

SCHEAFFER, R.L.; WATKINS, A.E.; LANDWEHR, J.M. . What every high-school graduate should know about statistics. In: ERLBAUM, Lawrence. **Rejections on statistics: Learning, teaching and assessment in grades K-12**. Mahwah-NJ: S.P. Lajoie, 1998, pp. 3-31.

SHULMAN, L. S. Those who understand: knowledge growth in teaching. **Educational Researcher**, 15 (2), 1986. p. 4-14.

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

SKEMP, R. R. **The psychology of learning mathematics**. Baltimore: Penguin, 1971.

SOARES, Magda. **Letramento: um tema em três gêneros**. 2. ed. Belo Horizonte: Autêntica, 2005.

STRUIK, Dirk J. **História concisa das matemáticas**. Trad. João Cosme Santos Guerreiro. 2 ed. Lisboa: Ciência Aberta, 1987.

STEVENSON, William J. **Estatística: aplicada à Administração**. Trad. Alfredo Alves de Farias. São Paulo: Harper e Row do Brasil, 2001.

TAUBER, Liliana Mabel. **La construcción del significado de la distribución normal a partir de actividades de análisis de datos**. 2001. 276 f. Tesis (Doctoral en Didáctica de las Matemáticas). Universidad de Sevilla, Sevilla, Spain.

THIOLLENT, M. **Metodologia da pesquisa-ação**. São Paulo: Cortez, 1986.

TRIOLA, Mário F. **Introdução à estatística**. Trad. Alfredo Alves de Farias. 7 ed. Rio de Janeiro: LTC, 1999.

USPENSKY, J. V. **Introduction to mathematical probability**. New York: McGraw-Hill Book, 1937.

VERGNAUD, Gerard. A teoria dos campos conceptuais. In: BRUN, Jean (dir.) **Didáctica das matemáticas**. Lisboa: Instituto Piaget, 1996, p. 155-191. (Coleção Horizontes Pedagógicos).

WALMANN, Katherine K. **Enhancing statistical literacy**: enriching our society. 1993. Disponível em: <http://www.stat.auckland.ac.nz/~iase/islp/def>. Acesso em: 8 jun. 2009.

WILD, C.J.; PFANNKUCH, M. Statiscal thinking in empirical enquiry. **International Statistical Review**, 67, 1999, p.223-265. Disponível em: <http://www.stat.auckland.ac.nz/~iase/publications/isr/99.Wild.Pfannkuch.pdf>. Acesso em: 05 jun. 2007.

WITTE, Robert S.; WITTE, John S.; **Estatística**. Trad. Teresa Cristina Padilha de Souza; Eduardo Benedito Curtolo. 7. ed. Rio de Janeiro: LTC, 2005.

WODEWOTZKI, Maria Lucia L.; JACOBINI, Otavio Roberto. O ensino de estatística no contexto da educação matemática. In: BICUDO, Maria Aparecida Viggiani; BORBA, Marcelo de Carvalho (orgs.) **Educação Matemática**: pesquisa em movimento. São Paulo: Cortez, 2004, p. 232-249.

WONNACOTT, Thomas H. e WONNACOTT, Ronald J. **Estatistique**: économie, gestion, sciences, médecine (avec exercices d'application). 4. ed. Paris: Economica, 1991.

APÊNDICE

APÊNDICE A - Método de Laplace para Teorema Central do Limite

No trabalho publicado em 1785, Laplace chamou atenção para a soma dos erros de uma distribuição. Isto é, anunciou que um erro poderia assumir qualquer valor entre -1 e 1 com igual probabilidade. Já para a soma dos erros de duas distribuições, seria para todos os valores entre $-2n$ e $2n$, isto é, $-2n, -2n + 2, \dots, -2, 0, 2, \dots, 2n - 2, 2n$, de probabilidades que correspondem aos coeficientes binomiais $(1+1)^{2n}$.

Ele apresentou um binômio simples, pois estava ‘definindo’ a função característica, cujo termo médio do binômio é y_n (FISCHER, 2000), e também é o termo independente e^{it} ⁴⁰ do desenvolvimento binomial $(e^{it} + e^{-it})^{2n}$. Se multiplicar por dt e integrar de 0 a π , a expressão será igual a πy_n . Então, tem-se que:

$$y_n = \frac{1}{\pi} \int_0^\pi (e^{it} + e^{-it}) dt. \quad (1)$$

Partindo de que $(e^{it} + e^{-it}) = 2\cos t$,⁴¹ tem-se a fórmula:

$$y_n = \frac{2^{2n}}{\pi} \int_0^\pi \cos^{2n} t dt. \quad (2)$$

O ponto chave dessa questão foi Laplace ter notado que poderia estender essa idéia para encontrar o termo médio para o trinômio $(1 + 1 + 1)^{2n}$, para o quadrinômio $(1 + 1 + 1 + 1)^{2n}$ e assim por diante. Cada um desses polinômios corresponde a ter erros entre seus termos (por exemplo, $-1, 0, 1$ para o caso do trinômio). Sua generalização permite que os erros sejam os valores entre

$$(-m + 1, \dots, -1, 0, 1, \dots, m - 1, m), \quad (3)$$

⁴⁰ Lembrando que a função geradora exponencial para $(1, 1, 1, \dots)$ é

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^r}{r!} + \dots,$$

e nesta expansão o coeficiente de $x^r/r!$ é igual a 1 , para todo r , esta é a função geradora exponencial da sequência $a_r = 1$, para todo $r = 0, 1, 2, \dots$

⁴¹ “ $(e^{it} + e^{-it}) = 2\cos t$, porque é uma função cosseno hiperbólica, cujo domínio é o conjunto de todos os números reais e a imagem é o conjunto de todos os números no intervalo $[1, +\infty)$ ”. (LEITHOLD, 1994, p. 515)

no caso de polinômio de ordem m . Substituindo o emésimo-nomial 1 por s com e^{it} por s (e s é a soma dos erros), obtém-se:

$$(e^{imt} + e^{i(m-1)t} + \dots + 1 + \dots + e^{i(-m+1)t} + e^{-imt})^{2n} \quad (5)$$

Os parênteses podem ser novamente simplificados por cossenos e assim,

$$\left(1 + 2 \sum_{k=1}^n \cos tk \right). \quad (6)$$

A probabilidade para cada termo é igual a $(2m + 1)^{-1}$, então, multiplicando o resultado por $(2m + 1)^{-1}$ e, cujas funções características são:

$$\psi(t) = E(e^{itx}). \quad (7)$$

Para o primeiro caso, erros -1 e 1 é $\psi(t) = \frac{1}{2}(e^{it} + e^{-it})$. E, no caso geral:

$$\psi(t) = \frac{1}{2m+1} \left(1 + 2 \sum_{k=1}^n \cos tk \right) \quad (8)$$

Agora, novamente usando a função característica para calcular as probabilidades para $s_n = 0$ (onde s_n é a soma dos n possíveis erros) pela fórmula:

$$P(s_n = 0) = \frac{1}{\pi} \int_0^\pi \Psi(t) dt = \frac{1}{\pi(2m+1)^n} \left| \left(1 + 2 \sum_{k=1}^m \cos tk \right)^n \right|_0^\pi dt. \quad (9)$$

Dessa forma, Laplace desenvolveu fórmula acima e, em seguida, concluiu que a aproximação seria então:

$$P(s_n = 0) \approx \frac{\sqrt{3}}{\sqrt{2\pi nm(m+1)}}. \quad (10)$$

Sob o ponto de vista do TCL, Laplace chegou a este resultado em 1785. Apesar de estar bastante perto do resultado do TCL, ele restringiu a prova apenas para $P(s_n = 0)$, não expandindo o resultado para qualquer outro valor de s_n . Contudo, foi ele mesmo quem deu continuidade a este resultado em 1810, quando afirmou ter obtido uma generalização para TCL. No trabalho publicado em 1820, Laplace

começa por provar o teorema para algumas distribuições de probabilidades, no caso, para as distribuições discretas e contínuas. Mether (2003) apresenta apenas os resultados obtidos por Laplace para o caso de uma distribuição discreta arbitrária.

Admitindo que se tenha uma variável aleatória discreta x que toma os valores $-m, -m + 1, \dots, m - 1, m$, com probabilidades correspondentes a $p_{-m}, p_{-m+1}, \dots, p_{m+1}, p_m$. Considerando que $\frac{1}{\pi} \int_{-\pi}^{\pi} e^{-itx} \cdot e^{isx} dx = \delta_{t,s}$, obtém-se:

$$P(s_n = j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ijt} \psi^n(t) dt \quad (11)$$

Neste caso a função característica é: $\psi(t) = \sum_{k=-m}^m p_k e^{ikt}$. Inserindo a função característica em (11) e, expandindo o resultado para e^{ikt} , tem-se que:

$$P(s_n = j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ijt} \left(1 + it \sum_{k=-m}^m p_k k - \frac{t^2}{2} \sum_{k=-m}^m p_k k^2 + \dots \right)^n dt \quad (12)$$

Aproximando a expressão dentro dos parênteses, particularmente o logaritmo natural, ocorre

$$\ln \psi^n = n \ln \left(1 + it \sum_{k=-m}^m p_k k - \frac{t^2}{2} \sum_{k=-m}^m p_k k^2 + \dots \right) = n \left(it \mu_x - \frac{t^2}{2} \sum_{k=-m}^m p_k k^2 + \frac{1}{2} t^2 \mu_x^2 + \dots \right) \quad (13)$$

Neste caso, μ_x é o valor esperado de x . Foi utilizada a série de Taylor para o logaritmo natural: $\ln(1+x) = x - \frac{1}{2}x^2 + \dots$. A variância para x pode ser descrita por

$\sigma_x^2 = \sum_{k=-m}^m p_k k^2 - \mu_x^2$. A partir de (13) e simplificando toda a expressão:

$P(s_n = j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-ijt + itn\mu_x - \frac{1}{2}n\sigma_x^2 t^2 + \dots) dt$. Em vez disso, pelo cálculo de

$P(s_n - n\mu_x = s)$, tem-se que:

$$\begin{aligned} P(s_n - n\mu_x = s) &= \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-ist - \frac{1}{2}n\sigma_x^2 t^2 + \dots) dt = \frac{1}{2\pi} \exp\left(-\frac{s^2}{2n\sigma_x^2}\right) \int_{-\pi}^{\pi} \exp\left[-\frac{1}{2}n\sigma_x^2 \left(t + \frac{is}{n\sigma_x^2}\right)^2\right] dt = \end{aligned}$$

$$= \frac{1}{2\pi\sqrt{n}\pi} \exp\left(-\frac{s^2}{2n\sigma_x^2}\right) \int_{-\pi\sqrt{n}\sigma_x}^{\pi\sqrt{n}\sigma_x} \exp\left[-\frac{1}{2}\left(z + \frac{is}{\sqrt{n}\sigma_x}\right)^2\right] dt$$

(14)

admitindo que s é o de maior ordem de $\sqrt{n}$, tem-se que:

$$P(s_n - n\mu_x = s) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{s^2}{2n\sigma_x^2}\right).$$

(15)

Isso mostra que $s_n - n\mu_x$ é assintótica e aproxima da distribuição normal $N(0, n\sigma_x^2)$. Laplace também observou que se m é parte da fórmula somente pela μ_x e σ_x , o resultado pode ser válido para uma distribuição discreta com os valores tendendo ao infinito, desde que a média μ_x e o desvio padrão σ_x , na distribuição, sejam conhecidos. Segundo o mesmo autor, para o caso contínuo, Laplace desenvolveu a demonstração de modo similar ao da discreta (METHER, 2003).

APÊNDICE B - FUNÇÃO CARACTERÍSTICA

De acordo com James (2006), Se X e Y são variáveis aleatórias em (Ω, A, P) , então $Z = X + iY$ é chamada uma *variável aleatória complexa*. Notemos que Z é uma função definida em Ω e que assume valores complexos, com $Z(\omega) = X(\omega) + iY(\omega)$ para $\omega \in \Omega$. A *esperança* EZ é definida pela linearidade, $EZ = EX + iEY$, se EX e EY são finitas.

A fórmula de Euler consiste de $e^{ix} = \cos x + i \sin x$, $x \in \mathbb{R}$, em que a variável aleatória complexa $e^{iX} = \cos X + i \sin X$ sempre possui esperança finita, para toda variável aleatória X , pois as variáveis aleatórias $\cos X$ e $\sin X$ são limitadas. Assim, a esperança na definição seguinte é finita e pode-se garantir que a função característica está bem definida.

Definição: Seja X uma variável aleatória. A *função característica* de X é a função

$$\varphi(t) = \varphi_X(t) = Ee^{itX} \quad (1)$$

e define-se

$$Ee^{itX} = E\cos(tX) + iE\sin(tX), \quad t \in \mathbb{R}. \quad (2)$$

Propriedades:

- i. A função característica é limitada por 1 : $|\varphi_X(t)| \leq 1, \forall t \in \mathbb{R}$.
- ii. A função característica assume o valor 1 no ponto 0: $\varphi_X(0) = 1$.
- iii. $\overline{\varphi_X(t)} = \varphi_X(-t)$, onde $\bar{c}$ é o complexo conjugado de c .
- iv. φ_X é uniformemente contínua na reta.
- v. Se X e Y são independentes, então $\varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t) \forall t \in \mathbb{R}$.
- vi. A função característica de uma variável aleatória X determina a função de distribuição de X .
- vii. A variável aleatória X tem distribuição simétrica em torno de zero se, e somente se, $\varphi_X(t)$ é real para todo t . (Por definição, X tem distribuição simétrica em torno de zero se $P(X \leq x) = P(X \geq -x), \forall x \in \mathbb{R}$. Às vezes dizemos neste caso que X é simétrica em torno de zero.)

APÊNDICE C - Função geratriz de momentos

Meyer (1983) chama de função geratriz de momentos da seguinte forma: Seja X , uma variável aleatória discreta, com distribuição de probabilidade $p(x_i) = P(X = x_i)$, $i = 1, 2, \dots$. A função M_X , denominada função geratriz de momentos de X , é definida por:

$$M_X(t) = \sum_{j=1}^{\infty} e^{tx_j} p(x_j). \quad (1)$$

Se X for uma variável aleatória contínua com função densidade de probabilidade, chama-se de função geratriz de momentos por:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) d(x). \quad (2)$$

Em qualquer dos casos, o discreto ou o contínuo, $M_X(t)$ é apenas o valor esperado de e^{tX} . Por isso, pode-se combinar as expressões e escrever:

$$M_X(t) = E(e^{tX}). \quad (3)$$

APÊNDICE D - ARTIGO DE POISSON

Seja uma variável aleatória Y_i no intervalo $[a, b]$ com densidade contínua $f_i(y)$ = $F'_i(y)$, o qual $F_i(y) = P(Y_i \leq y)$. Seja $\alpha \leq x_i \leq \beta$, ($\alpha\delta = a$ e $\beta\delta = b$) uma variável aleatória discreta, e δ um intervalo pequeno e estabelece que:

$$p_{x_i} = f(x_i\delta)\delta, \quad x_i = \alpha, \alpha + 1, \dots, \beta. \quad (1)$$

A função característica para x_i é:

$$\psi_{y_i}(\theta) = \sum_{x_i=\delta}^{\beta\delta} p_{x_i} e^{itx_i} = \sum_{x_i=\alpha\delta}^{\beta\delta} f_i(x_i\delta) e^{i(t/\delta)x_i\delta} \quad (2)$$

Estabelecendo $t = \theta\delta$ no membro direito da igualdade (3) tende a:

$$\psi_{y_i}(\theta) = \int_1^b f(y) e^{i\theta y} dy. \quad (3)$$

Assim, deseja-se encontrar a distribuição de probabilidade para soma:

$$s_n = Y_1 + \dots + Y_n, \quad (4)$$

conforme segue:

$$P(\delta \sum x_i = \delta s) = \frac{1}{2\pi} \int_1^b e^{its} \prod_{i=1}^n (\psi_{x_i}(y)) dt. \quad (5)$$

Daqui em diante, Poisson continuou e, finalmente, alcançou o seguinte resultado:

$$P(c - \varepsilon \leq s_n \leq c + \varepsilon) = \frac{1}{\pi} \int_{-\infty}^{\infty} \left(\prod_{i=1}^n \int_1^b f_i(y) e^{i\theta y} dy \right) e^{-i\theta c} \frac{\text{sen}(\varepsilon\theta)}{\theta} \quad (6)$$

Poisson não conseguiu condições de modo a restringir demonstração do teorema com rigor para a fórmula geral, mas examinou a validade para um caso específico: $n = 1$. Assim, substituindo em (6), escreve-se:

$$P(c - \varepsilon \leq Y_1 \leq c + \varepsilon) = \frac{1}{\pi} \int_1^b \int_{-\infty}^{\infty} e^{i\theta(y-c)} \frac{\text{sen}(\varepsilon\theta)}{\theta} d\theta f_1(y) dy \quad (7)$$

Com o auxílio da fórmula:

$$\frac{2}{\pi} \int_0^{\infty} \frac{\text{sen} kx}{x} dx = \begin{cases} 1, & \text{se } y \in]c - \varepsilon, c + \varepsilon[\\ -1, & \text{se } y \notin]c - \varepsilon, c + \varepsilon[\end{cases} \quad (8)$$

E, a partir de (8), chega-se no resultado final:

$$P(c - \varepsilon, c + \varepsilon) = \int_{-\varepsilon}^{\varepsilon} f_1(y) dy, \quad (9)$$

o que conclui a demonstração de Poisson.

Em síntese, Methner (2003) descreve a versão de Poisson para o TCL da seguinte forma: Seja $Y_1, \dots, Y_n$ variáveis aleatórias independentes com função densidade desaparecendo além do intervalo fixado $[a, b]$. Se para os valores absolutos P_j das funções características de $Y_i(\psi_j(\theta) = p_j \cdot e^{k p_j})$, então existe uma função $r(\alpha)$ independente de j e $0 \leq r(\alpha) \leq 1$ para todo $\alpha \neq 0$ tal que $p_j \leq r(\alpha)$. Logo, para valores arbitrários γ e γ' tem-se que:

$$P\left(\gamma \leq \frac{\sum (Y_i - EY_i)}{\sqrt{2 \sum Var(Y_i)}} \leq \gamma'\right) \approx \frac{1}{\sqrt{\pi}} \int_{\gamma}^{\gamma'} e^{-u^2} du. \quad (10)$$

Assim, a aproximação é tanto melhor quanto n maior e, a diferença entre o membro esquerdo e o direito vem a ser ‘infinitamente pequeno’ com n tendendo ao infinito.

APÊNDICE E - PRIMEIRA VERSÃO DO TEOREMA CENTRAL DO LIMITE

Bessel (1784-1846), Dirichlet (1805-1859), Cauchy e Ellis (1814-1890) são alguns exemplos daqueles que trabalharam, por meio de várias tentativas, a primeira versão do TCL de Laplace com ‘rigor matemático’.

Dirichlet e Bessel seguiram fielmente cada etapa da demonstração de Laplace e de Poisson, acrescentando o ‘fator de descontinuidade’ em seus trabalhos. Com isso, viabilizaram a equação de Poisson:

$$P(c - \varepsilon \leq s_n \leq c + \varepsilon) = \frac{1}{\pi} \int_{-\infty}^{\infty} \left(\prod_{i=1}^n \int_1^b f_i(y) e^{i\theta y} dy \right) e^{-i\sigma\theta} \frac{\text{sen}(\varepsilon\theta)}{\theta}, \quad (1)$$

da demonstração de Poisson para um valor de n , o qual Poisson não o tinha feito.

Segundo Methner (2003), o fator de descontinuidade de Dirichlet é descrito como:

$$\frac{2}{\pi} \int_0^{\infty} \frac{\text{sen}(kx)}{x} \cos x dx = \begin{cases} 0, & |k| > 1 \\ 1, & |k| < 1 \end{cases}. \quad (2)$$

Foi com essa fórmula que viabilizaram a fórmula de Poisson, conforme descrita anteriormente, para um valor de n arbitrário. Bessel também chamou atenção de que Poisson, que de certa forma, atribuiu um fator de descontinuidade, ou seja, para $n = 1$. Mas apenas para esse caso.

Além disso, Dirichlet tentou estimar os erros de aproximação, porém sem sucesso. Mas fez algo diferente: estimou erros para aproximações não-probabilísticas e, em seguida, apresentou algumas técnicas que poderiam ser aplicadas tanto na teoria das probabilidades quanto para o que chamavam de matemática ‘pura’.

Cajori (2007) afirma que Cauchy é considerado como o primeiro matemático a dar um tratamento à teoria das probabilidades com rigor. Cauchy observou a demonstração do TCL, comparando com as anteriores. Primeiro encontrou um limite

máximo da diferença entre o valor exato e a aproximação. Em seguida, especificou condições para que esse limite tendesse a zero.

Cauchy apresentou, no seu trabalho, as variáveis aleatórias independentes distribuídas identicamente: $y_1, \dots, y_n$, com densidade simétrica $f(y)$ no intervalo finito $[-a, a]$, variância $\sigma^2 > 0$ e uma função característica $\Psi(\theta)$. Ele considerou a média da distribuição como:

$$z_n = \sum_{i=1}^n w_i y_i \cdot \sum_{i=1}^n w_i = 1 \quad (3)$$

Seguindo procedimento de Poisson, obteve:

$$P = P(-h < z_n < h) = \frac{2}{\pi} \int_0^{\infty} \Psi(\omega, \theta) \dots \Psi(\omega_n \theta) \frac{\sin(h\theta)}{\theta} d\theta. \quad (4)$$

Assim, Cauchy tentou encontrar um limite máximo para $|P - \Phi|$ e, em condições específicas, para este limite quando tende a zero.

No desenvolvimento de sua demonstração, ele dividiu o intervalo em três partes:

$$P - \Phi = (P - P_k) + (P_k - \Phi_k) + (\Phi_k - \Phi), \quad (5)$$

em que foi feito um estudo de cada parte separadamente, obtendo:

$$|\Phi_k - \Phi| = \left| \frac{2}{\pi} \int_k^{\infty} e^{-\beta^2 \theta^2 / 2} \frac{\sin(h\theta)}{\theta} d\theta \right| < \frac{2}{\pi} \frac{1}{\beta^2 k^2} e^{-\beta^2 k^2 / 2} \frac{2}{\pi} \quad (6)$$

$$|P_k - P| = \left| \frac{2}{\pi} \int_k^{\infty} e^{-\beta^2 \theta^2 \gamma / 2} \frac{\sin(h\theta)}{\theta} d\theta \right| < \frac{2}{\pi} \frac{1}{\beta^2 k^2 \gamma^2} e^{-\beta^2 k^2 \gamma / 2}, \quad (7)$$

onde

$$\gamma = \frac{1}{1 + \sigma^2 k^2 (\max(w_j))^2}. \quad (8)$$

Para última parte, Cauchy anunciou que:

$$|P_k - \Phi_k| = \left| \frac{2}{\pi} \int_0^k (e^{-\beta^2 \theta^2 \gamma} - e^{-\beta^2 \theta^2 / 2}) \frac{\sin(h\theta)}{\theta} d\theta \right| < \frac{2\sqrt{3}\alpha}{\pi} \ln \left(\frac{kh}{\sqrt{3}} + \sqrt{1 + \frac{k^2 h^2}{3}} \right) \quad (9)$$

em que,

$$\alpha = \max \left(e^{\frac{1}{8}\beta^2 k^4 a^2 (\max(w_j))^2}, 1 - e^{\frac{\sigma^2 \beta^2 k^4 (\max(w_j))^2}{4 + 2\sigma^2 k^2 (\max(w_j))^3}} \right) \quad (10)$$

Daí, Cauchy finalmente concluiu que se k é escolhido de $n^{1/2} < k < n^{3/4}$, então as três partes tendem para zero quando n tende a infinito. E, assim, $P \rightarrow \Phi$.

APÊNDICE F - CHEBYSHEV (1821-1894), MARKOV (1856-1922) E LIAPOUNOV (1857-1918)

Chebyshev (1821-1894), Markov (1856-1922) e Liapounov (1857-1918), segundo Senata (1984), são considerados aqueles que contribuíram significativamente para o desenvolvimento da demonstração do TCL.

A publicação do artigo de Chebyshev é, geralmente, considerada o início das provas com rigor para o TCL. Senata (1984) afirma que ele prosseguiu da seguinte forma:

Seja $z_1, z_2, \dots$, variáveis aleatórias independentes que é descrita por probabilidades densas. Se (i) $E(z_i) = 0$ para todo i e, (ii) $|E(z_i^k)| \leq C, \forall i, k \geq 2$, onde C é uma constante independente de i e k , então, como $n \rightarrow \infty$,

$$P\left(t < \frac{s_n}{B_n} < t'\right) \rightarrow \frac{1}{2\pi} \int_t^{t'} e^{-\frac{1}{2}x^2} dx, \quad (1)$$

onde

$$s_n = \sum_{i=1}^n z_i \text{ e } B_n^2 = \sum_{i=1}^n \text{var}(z_i). \quad (2)$$

Chebyshev usou o ‘método dos momentos’, que foi por ele mesmo desenvolvido anteriormente. Mais tarde, o seu trabalho recebeu a contribuição de Markov que declarou: *Uma condição necessária e suficiente a ser acrescida, em ordem, para que o teorema se faça de modo correto é o seguinte: (iii) $\frac{B_n^2}{n}$ é uniformemente tendendo a zero; (iii.a) $E(z_n^2)$ também tende a zero quando n tende ao infinito.*

Markov trabalhou arduamente para obter a generalização do método de momentos após a prova de Liapounov. Ele finalmente foi bem sucedido em 1913,

quando apresentou um artigo que continha uma prova rigorosa sobre TCL, superando a condição de Liapounov.

Liapounov não recorreu ao método de momentos, mas seguiu a idéia de Laplace, fazendo o uso das funções características. A demonstração de Liapounov foi publicada em 1901 e é considerada a ‘primeira’ demonstração do TCL. A seguir, um breve comentário declarado por Liapounov segundo Uspensky (1937):

“Sejam $x_1, \dots, x_n$ variáveis aleatórias independentes providas das seguintes propriedades: $E(x_i) = 0$ para todo i ; $E|x_i|^k \leq \infty, \forall i, k \geq 2$. Se existe um $\delta > 0$ tal que

$$\frac{\sum_{i=1}^n E|x_i|^{2+\delta}}{B_n^{2+\delta}} \xrightarrow{n \rightarrow \infty} 0.$$
(4)

Então,

$$P\left(\frac{s_n}{B_n} < t\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}u^2} du."$$
(5)

Após essa parte, Liapounov segue a demonstração de Laplace, usando funções características. Nessa demonstração, entretanto, ele usa um *lema fundamental* que é a chave da simplicidade e rigor, simultaneamente, de seu trabalho. Isto é, seja s_n uma variável dependente na integral n , com média 0 e variância 1. Se sua função característica é: $\Psi_n(t) = E(e^{is_n t})$ e converge para $e^{-\frac{t^2}{2}}$ (função característica da distribuição normal) uniformemente num intervalo finito qualquer $(-k, k)$, então, a função distribuição tende uniformemente para $\Phi(t)$ para todo valor t . (METHER, 2003)

Liapounov não distingue explicitamente o *lema fundamental* de sua demonstração, mas está implicitamente contido no contexto do desenvolvimento de seu trabalho.

APÊNDICE G - DEMONSTRAÇÃO DO TEOREMA CENTRAL DO LIMITE

I - PRELIMINARES PARA DEMONSTRAÇÃO DO TEOREMA CENTRAL DO LIMITE

James (2006) segue o problema central do limite como a convergência de uma distribuição de somas normais normalizadas,

$$\frac{S_n - ES_n}{\sqrt{\text{Var}(S_n)}} \xrightarrow{D} N(0, 1) \quad (1)$$

para a distribuição normal padrão, $N(0, 1)$. Neste caso, "supõe-se que todas as variâncias sejam finitas e que pelo menos uma delas seja estritamente positiva".

Uma questão pode ser levantada: quais as condições para os quais

$$\frac{S_n - ES_n}{\sqrt{\text{Var}(S_n)}} \xrightarrow{D} N(0, 1)? \quad (2)$$

Para responder esta questão deve-se supor que as variáveis sejam independentes e identicamente distribuídas. Se isso ocorre e possui a média μ e variância σ_k^2 , em que $0 < \sigma^2 < \infty$, então,

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} N(0, 1). \quad (3)$$

James (2006) define o teorema (Teorema Central do Limite de Lindeberg) apresentando as condições gerais para validade da convergência normal:

Teorema Central do Limite de Lindeberg: Sejam $X_1, \dots, X_n$ variáveis aleatórias independentes tais que $EX_n = \mu_x$ e $\text{Var } X_n = \sigma_n^2$, onde $0 < \sigma_n^2 < \infty$ e pelos menos um $\sigma_n^2 > 0$. Sejam $F_n = F_{X_n}$,

$$S_n = X_1 + \dots + X_n \quad \text{e} \quad s_n = \sqrt{\text{Var}(S_n)} = \sqrt{\sigma_1^2 + \dots + \sigma_n^2}$$

Então para que

$$\frac{S_n - ES_n}{S_n} \xrightarrow{D} N(0,1), \text{ quando } n \rightarrow \infty$$

é suficiente que a seguinte condição, chamada condição de Lindeberg esteja satisfeita:

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \frac{1}{S_n^2} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon S_n} (x - \mu_k)^2 dF_k(x) = 0.$$

(JAMES, 2006, p. 266).

Isto significa que se a condição de Lindeberg está satisfeita, vale a convergência normal.

O autor pondera algumas observações que são bastante pertinentes para antes da demonstração do teorema:

1) $\int_{|x - \mu_k| > \varepsilon S_n}$ é uma integração realizada em

$$\{x : |x - \mu_k| > \varepsilon S_n\} = (-\infty, \mu_k - \varepsilon S_n) \cup (\mu_k + \varepsilon S_n, +\infty),$$

(4)

lembrando que os extremos não estão incluídos na região de integração;

2) Se X_k for discreta, com função probabilidade $p_k(x_i)$, então

$$\int_{|x - \mu_k| > \varepsilon S_n} (x - \mu_k)^2 dF_k(x) = \sum_{i: |x - \mu_k| > \varepsilon S_n} (x - \mu_k)^2 p_k(x_i)$$

(5)

e, caso x_k for contínua, a função densidade $f_k(x)$ é

$$\int_{|x - \mu_k| > \varepsilon S_n} (x - \mu_k)^2 dF_k(x) = \int_{-\infty}^{\mu_k - \varepsilon S_n} (x - \mu_k)^2 f_k(x) + \int_{\mu_k + \varepsilon S_n}^{+\infty} (x - \mu_k)^2 f_k(x)$$

(6)

3) A representação da variância é:

$$\sigma_k^2 = \int (x - \mu_k)^2 dF_k(x) = \int_{|x - \mu_k| > \varepsilon S_n} (x - \mu_k)^2 dF_k(x) + \int_{|x - \mu_k| \leq \varepsilon S_n} (x - \mu_k)^2 dF_k(x)$$

(7)

de modo que a condição de Lindeberg pode ser escrita a partir de

$$s_n^2 = \sigma_1^2 + \dots + \sigma_n^2 : \forall \varepsilon > 0, \lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^n \int_{|x - \mu_k| \leq \varepsilon S_n} (x - \mu_k)^2 dF_k(x) \rightarrow 1$$

(8)

quando $n \rightarrow \infty$.

- 4) Na condição de Lindeberg, as parcelas $\frac{x_k - \mu_x}{s_n}$ da soma $\frac{S_n - ES_n}{s_n}$ são uniformemente pequenas para n grande. James (2006) apresenta um exemplo como consequência dessa condição:

$$\max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} \rightarrow 0 \text{ quando } n \rightarrow \infty, \quad (9)$$

isto é, para n grande, as variâncias das parcelas são uniformemente pequenas em relação à variância da soma. Portanto,

$$\begin{aligned} \frac{\sigma_k^2}{s_n^2} &= \frac{1}{s_n^2} \int_{|x - \mu_k| \leq \varepsilon s_n} (x - \mu_k)^2 dF_k(x) + \frac{1}{s_n^2} \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 dF_k(x) \leq \\ &\leq \frac{1}{s_n^2} \int_{|x - \mu_k| \leq \varepsilon s_n} \varepsilon^2 s_n^2 (x - \mu_k)^2 dF_k(x) + \frac{1}{s_n^2} \sum_{j=1}^n \int_{|x - \mu_j| > \varepsilon s_n} (x - \mu_j)^2 dF_j(x) \end{aligned} \quad (10)$$

Como o último termo não depende de k , pois ε^2 está na primeira parcela, tem-se que:

$$\max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} \leq \varepsilon^2 + \frac{1}{s_n^2} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 dF_k(x), \quad (11)$$

e converge para ε^2 , pela condição de Lindeberg. Como isso vale para todo $\varepsilon > 0$, o

$$\max \left(\sigma_n^2 / s_n^2 \right) \rightarrow 0. \quad (12)$$

O autor ainda ressalta que o significado dessa condição implica que as parcelas $\frac{x_k - \mu_x}{s_n}$ possuem variâncias uniformemente pequenas quando n é grande,

ou seja, nenhuma parcela tem muito peso para a soma $\frac{S_n - ES_n}{s_n}$. Esta alegação

nos reporta a aceitar, intuitivamente, a afirmação: “a soma de um grande número de pequenas quantidades independentes e de média zero, tem aproximadamente a distribuição normal” (JAMES, 2006, p.268).

- 5) A condição de Lindeberg é formalmente mais forte que a condição sobre o máximo das variâncias. Pois

$$s_n^2 = \sum_{k=1}^n \int (x - \mu_k)^2 dF(x) \text{ e,} \quad (13)$$

pela condição de Lindeberg, quando n é grande, é pequena a parte da variância da soma devida às "caudas" das X_k situadas a mais de ε desvios-padrão s_n das suas respectivas médias μ_k .

- 6) Com a existência da condição sobre o máximo, a condição de Lindeberg torna-se necessária para a validade do TCL.
- 7) A recíproca para o Teorema de Lindeberg deve-se a Feller: se $X_1, X_2, \dots$ são independentes com variâncias finitas $\sigma_1^2, \sigma_2^2, \dots$, se pelo menos um $\sigma_n^2 > 0$ e se

$\max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} \rightarrow 0$ quando $n \rightarrow \infty$, então a condição de Lindeberg é consequência da convergência normal.

De posse destas observações e anterior à demonstração do TCL, vejamos os seguintes:

Corolário 1: Se $X_1, X_2, \dots$ são variáveis aleatórias independentes e identicamente distribuídas com $EX_n = \mu$ e $\text{Var } X_n = \sigma^2$, onde $0 < \sigma^2 < \infty$, então

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1) \text{ quando } n \rightarrow \infty. \quad (14)$$

Demonstração: Pela condição de Lindeberg: $s_n^2 = n\sigma^2$ e, para $\varepsilon > 0$,

$$\frac{1}{n\sigma^2} \sum_{k=1}^n \int_{|x-\mu| \leq \varepsilon\sigma\sqrt{n}} (x-\mu)^2 dF_k(x) = \frac{1}{\sigma^2} \sum_{k=1}^n \int_{|x-\mu| \leq \varepsilon\sigma\sqrt{n}} (x-\mu)^2 dF_1(x) \quad (15)$$

já que são identicamente distribuídas. Assim,

$$\frac{1}{\sigma^2} \sum_{k=1}^n \int_{|x-\mu| \leq \varepsilon\sigma\sqrt{n}} (x-\mu)^2 dF_1(x) \xrightarrow{n \rightarrow \infty} \frac{1}{\sigma^2} \int_{-\infty}^{\infty} (x-\mu)^2 dF_1(x) = \frac{\sigma^2}{\sigma^2} = 1, \quad (16)$$

pois a convergência decorre da definição da integral imprópria de Riemann-Stieltjes. (JAMES, 2006, p. 269-270)

Este corolário baseia-se da condição de Lindeberg para $k = 1$, pois as variáveis são independentes e identicamente distribuídas. James (2006) recorre ao Teorema de Liapounov para o segundo corolário, para o caso das variáveis X_n possuírem momentos finitos de ordem maior que 2. Isto é, o teorema vale para a convergência normal se a soma dos momentos centrais absolutos de ordem $2 + \delta$ é assintoticamente pequena em relação à $s_n^{2+\delta}$.

Corolário 2: (Teorema Central do Limite de Liapounov). *Sejam $X_1, X_2, \dots$ são variáveis aleatórias independentes tais que $EX_n = \mu_n$ e $Var X_n = \sigma_n^2 < \infty$, com pelo menos um $\sigma_n^2 > 0$. Seja $s_n^2 = Var S_n = \sigma_1^2 + \dots + \sigma_n^2$. Se existir $\delta > 0$ tal que*

$$\frac{1}{s_n^{2+\delta}} \sum_{k=1}^n E|X_k - \mu_k|^{2+\delta} \xrightarrow{n \rightarrow \infty} 0 \text{ quando } n \rightarrow \infty,$$

(16)

então

$$\frac{S_n - ES_n}{s_n} \xrightarrow{D} N(0,1).$$

(17)

Demonstração: Supondo a condição de Liapounov satisfeita, verificamos a

condição de Lindeberg: para $\varepsilon > 0$, se $|x - \mu_k| > \varepsilon s_n$ então $\frac{|x - \mu_k|^\delta}{(\varepsilon^\delta s_n^\delta)} > 1$ de modo que

$$\begin{aligned} \frac{1}{s_n^2} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 dF_k(x) &\leq \frac{1}{s_n^2} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 \cdot \frac{|x - \mu_k|^\delta}{\varepsilon^\delta s_n^\delta} dF_k(x) \leq \\ &\leq \frac{1}{\varepsilon^\delta s_n^{\delta+2}} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon s_n} |x - \mu_k|^{\delta+2} dF_k(x) \leq \\ &\frac{1}{\varepsilon^\delta s_n^{\delta+2}} \sum_{k=1}^n \int_{-\infty}^{\infty} |x - \mu_k|^{\delta+2} dF_k(x) = \frac{1}{\varepsilon^\delta s_n^{\delta+2}} \sum_{k=1}^n E|X_k - \mu_k|^{\delta+2} \rightarrow 0 \text{ quando } n \rightarrow \infty. \end{aligned}$$

(18)

Após a prova do segundo corolário, James (2006) apresenta um exemplo como forma de enfatizar que o valor exato da variância não tem importância, mas

sim, a ordem de s_n^2 . Desta forma, ele utiliza o lema abaixo para tratar da ordem de séries do tipo $\sum n^\lambda$.

Lema 1: Para $\lambda > 0$,

$$\frac{1}{n^{\lambda+1}} \sum_{k=1}^n k^\lambda \rightarrow \frac{1}{\lambda+1} \text{ quando } n \rightarrow \infty, \quad (19)$$

de maneira que $\sum_{k=1}^n k^\lambda$ é da ordem de $n^{\lambda+1}$.

Demonstração: Como $x^\lambda \leq k^\lambda$ se $k-1 \leq x \leq k$ e $k^\lambda \leq x^\lambda$ se $k \leq x \leq k+1$, assim, segue-se que:

$$\int_{k-1}^k x^\lambda dx \leq \int_{k-1}^k k^\lambda dx = k^\lambda = \int_k^{k+1} k^\lambda dx \leq \int_k^{k+1} x^\lambda dx. \quad (20)$$

Somando os valores de k de 1 até n tem-se:

$$\int_0^n x^\lambda dx \leq \sum_{k=1}^n k^\lambda \leq \int_1^{n+1} x^\lambda dx. \quad (21)$$

Portanto,

$$\frac{n^{\lambda+1}}{\lambda+1} \leq \sum_{k=1}^n k^\lambda \leq \frac{(n+1)^{\lambda+1} - 1}{\lambda+1} \leq \frac{(n+1)^{\lambda+1}}{\lambda+1}, \quad (22)$$

Ou seja,

$$\frac{1}{\lambda+1} \leq \frac{1}{n^{\lambda+1}} \sum_{k=1}^n k^\lambda \leq \frac{1}{\lambda+1} \cdot \left(\frac{n+1}{n} \right)^{\lambda+1}. \quad (23)$$

Como $\left(\frac{n+1}{n} \right)^{\lambda+1} \rightarrow 1$ quando $n \rightarrow \infty$, o lema está provado. (JAMES, 2006, p.271).

II - PROVA DO TEOREMA CENTRAL DO LIMITE

A partir da condição do teorema de Lindeberg:

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 dF_k(x) = 0, \quad (24)$$

e, iniciando a prova ao mostrar que as funções características das somas parciais padronizadas convergem para a função característica da $N(0, 1)$, para todo $t \in \mathbb{R}$, tem-se que: (JAMES, 2006)

$$\begin{aligned} \frac{\phi_{(S_n - ES_n)}(t)}{s_n} &= \text{(por independência)} \\ &= \prod_{k=1}^n E e^{it \left(\frac{X_k - \mu_k}{s_n} \right)} \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}} \quad (I) \end{aligned} \quad (25)$$

Ao fixar um valor t real e recorrendo à Fórmula de Taylor aplicada à função $g(x) = e^{itx}$ em duas situações, a saber:

$$e^{itx} = 1 + itx + \theta_1(x) \frac{t^2 x^2}{2}, \text{ para } |\theta_1(x)| \leq 1 \quad (26)$$

$$e^{itx} = 1 + itx - \theta_1(x) \frac{t^2 x^2}{2} + \theta_2(x) \frac{t^3 x^3}{6}, \text{ para } |\theta_2(x)| \leq 1 \quad (27)$$

Considerando $\varepsilon > 0$ e usando a primeira Fórmula de Taylor para $|x| > \varepsilon$ e a segunda para $|x| \leq \varepsilon$, escreve-se que e^{itx} de forma generalizada como:

$$e^{itx} = 1 + itx - \frac{t^2 x^2}{2} + r_\varepsilon(x), \text{ onde } r_\varepsilon(x) = \begin{cases} (1 + \theta_1(x)) \frac{t^2 x^2}{2}, & \text{se } |x| > \varepsilon \\ \theta_2(x) \frac{t^3 x^3}{6}, & \text{se } |x| \leq \varepsilon \end{cases}, \quad (28)$$

Portanto,

$$E \exp \left\{ it \left(\frac{X_k - \mu_k}{s_n} \right) \right\} = \int \exp \left\{ it \left(\frac{x - \mu_k}{s_n} \right) \right\} dF_k(x) =$$

$$= \left\{ 1 + it \left(\frac{x - \mu_k}{s_n} \right) - \frac{t^2}{2} E \left(\frac{x - \mu_k}{s_n} \right)^2 + r_\varepsilon \left(\frac{x - \mu_k}{s_n} \right) \right\} dF_k(x) =$$

(29)

$$= \text{por linearidade} =$$

$$= 1 + it E \left(\frac{X_k - \mu_k}{s_n} \right) - \frac{t^2}{2} E \left(\frac{X_k - \mu_k}{s_n} \right)^2 + \frac{t^2}{2} \int_{|x - \mu_k| > \varepsilon s_n} \left\{ 1 + \theta_1 \left(\frac{x - \mu_k}{s_n} \right) \right\} \left(\frac{x - \mu_k}{s_n} \right)^2 dF_k(x) +$$

$$+ \frac{t^3}{6} \int_{|x - \mu_k| \leq \varepsilon s_n} \theta_2 \left(\frac{x - \mu_k}{s_n} \right) \left(\frac{x - \mu_k}{s_n} \right)^3 dF_k(x)$$

(30)

Como $EX_k = \mu_k$ e $Var X_k = \sigma_k^2$, tem-se que

$$E \exp \left\{ it \left(\frac{X_k - \mu_k}{s_n} \right) \right\} = 1 - \frac{t^2 \sigma_k^2}{2 s_n^2} + e_{n,k},$$

(31)

já que $|\theta_1(x)| \leq 1$ e $|\theta_2(x)| \leq 1$, o resto $e_{n,k}$ satisfaz a:

$$|e_{n,k}| \leq t^2 \int_{|x - \mu_k| > \varepsilon s_n} \left(\frac{x - \mu_k}{s_n} \right)^2 dF_k(x) + \frac{|t|^3}{6} \int_{|x - \mu_k| \leq \varepsilon s_n} \varepsilon \left(\frac{x - \mu_k}{s_n} \right)^2 dF_k(x) \leq$$

$$\leq \frac{t^2}{s_n^2} \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 dF_k(x) + \frac{\varepsilon |t|^3}{6 s_n^2} \int_{-\infty}^{\infty} (x - \mu_k)^2 dF_k(x).$$

(32)

Desta forma,

$$\sum_{k=1}^n |e_{n,k}| \leq \frac{t^2}{s_n^2} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 dF_k(x) + \frac{\varepsilon |t|^3}{6 s_n^2}$$

(33)

Pela condição de Lindeberg, a primeira parcela do termo à direita tende a zero quando $n \rightarrow \infty$. Desta forma, para n suficientemente grande,

$$\sum_{k=1}^n |e_{n,k}| \leq \frac{\varepsilon |t|^3}{3}.$$

(34)

Escolhendo uma seqüência de ε 's que converge para zero, se $\varepsilon = \frac{1}{m}$, existe n_m tal que para $n \geq n_m$ tal que:

$$\sum_{k=1}^n |e_{n,k}| \leq \frac{\varepsilon |t|^3}{3}, \quad (35)$$

em que os restos $e_{n,k}$ são os determinados pela fórmula baseada em $\varepsilon = \frac{1}{m}$. Assim, existe uma seqüência $n_1 < n_2 < \dots$ de inteiros positivos tal que $\sum_{k=1}^n |e_{n,k}| \leq \frac{\varepsilon |t|^3}{3}$, para $n_m \leq n < n_{m+1}$, onde os valores são também baseados pelo fato de $\varepsilon = \frac{1}{m}$. Vale lembrar que o valor de ε que determina o resto $e_{n,k}$ depende da posição de n em relação aos n_m . Então,

$$\sum_{k=1}^n |e_{n,k}| \rightarrow 0 \text{ quando } n \rightarrow \infty. \quad (II) \quad (36)$$

Lembrando que a função característica, (I), é representado por

$$\delta_{\frac{(S_n - ES_n)}{s_n}}(t) \underset{\text{(por independência)}}{=} \prod_{k=1}^n E e^{it \left(\frac{X_k - \mu_k}{s_n} \right)} \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}} \quad (37)$$

e, substituindo (II) em (I), tem-se que:

$$\delta_{\frac{(S_n - ES_n)}{s_n}}(t) = \prod_{k=1}^n \left(1 - \frac{t^2 \sigma_k^2}{2s_n^2} + e_{n,k} \right) \quad (38)$$

com os valores de $e_{n,k}$ satisfazendo (II).

Para provar que o termo à direita converge para $e^{-\frac{t^2}{2}}$, recorre-se ao lema dos números complexos que generaliza o resultado já utilizado para provar o Teorema Central do Limite no vaso de variáveis independentes e identicamente

distribuídas, de que $c_n \rightarrow c$ implica que $\left(1 + \frac{c_n}{n}\right)^n \rightarrow e^c$.

Para finalizar a prova do TCL, considerando o segundo lema, e segundo James (2006):

Lema 2: *Sejam $c_{n,k}$ números complexos tais que $\sum_{k=1}^n c_{n,k} \rightarrow c$ quando $n \rightarrow \infty$.*

Se

$$\max_{1 \leq k \leq n} |c_{n,k}| \rightarrow 0 \text{ quando } n \rightarrow \infty \quad (39)$$

e

$$\sum_{k=1}^n |c_{n,k}| \leq M < \infty \quad (40)$$

onde M é uma constante que não depende de n , então

$$\prod_{k=1}^n (1 + c_{n,k}) \rightarrow e^c \text{ quando } n \rightarrow \infty. \quad (41)$$

Demonstração:

Para dar seguimento da prova da condição de Lindeberg e provido do segundo lema, sejam:

$$c_{n,k} = -\left(\frac{t^2 \sigma_k^2}{2s_n^2}\right) + e_{n,k} \quad \text{e} \quad c = -\frac{t^2}{2}, \text{ por (II), tem-se que:} \quad (42)$$

$$\sum_{k=1}^n |c_{n,k}| \leq \frac{t^2}{2} + \sum_{k=1}^n |e_{n,k}| \xrightarrow{n \rightarrow \infty} \frac{t^2}{2}, \quad (43)$$

o que significa que $\sum_{k=1}^n |c_{n,k}|$ é uniformemente limitado, ou seja, existe $M < \infty$ tal que

$\forall n, \sum_{k=1}^n |c_{n,k}| \leq M$. Antes de aplicar o lema, segue a verificação da condição de máximo:

$$\max_{1 \leq k \leq n} |c_{n,k}| \leq \max_{1 \leq k \leq n} \frac{t^2 \sigma_k^2}{2s_n^2} + \max_{1 \leq k \leq n} |e_{n,k}| \leq \frac{t^2}{2} \max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} + \sum_{k=1}^n |e_{n,k}|, \quad (44)$$

com o segundo termo tendendo a zero por (II). Como a condição de Lindeberg

implica que $\max_{1 \leq k \leq n} \left(\frac{\sigma_k^2}{s_n^2} \right) \rightarrow 0$, a demonstração está terminada.

ANEXOS

ANEXO 2 – Primeira atividade com os licenciandos

• RESPOSTA DE AO

ATIVIDADE 1

Estadística → CES JF. (MATEMÁTICA)
07/11/2007 7º Período

① O que você entende por variável (ou variáveis)?

Aquilo que não é constante, como por exemplo, avaliar o grau de auto-estima de alunos da 5ª série. Os resultados obtidos são as variáveis.

② E por aleatoriedade?

Algo que não se pode prever, como por exemplo, retirar uma bola aleatoriamente de uma urna numerada de um a cinco.

③ E independência de dois ou mais eventos?

Quando um evento não depende do outro para ser estudado.

④ Para você, variável aleatória tem alguma relação com função (matemática)?

Sim.

⑤ Em caso afirmativo no item ④, elabore um "conceito" para variável aleatória e variáveis aleatórias independentes. Variável aleatória é aquela que os valores estão em função da aleatoriedade, ou seja, a medida que varia a amostra, variam-se os dados correspondentes.

Variáveis aleatórias independentes tem o mesmo conceito anterior, porém, de forma independente uma da outra.

- RESPOSTA DE BM

(1) O que você entende por variável (ou variáveis)?

Possibilidade de algo ocorrer variá, ou seja, não ser um re-
do constante.

(2) E por aleatoriedade?

Algo que ocorre ao acaso.

(3) É independência de dois ou mais eventos?

É quando a ocorrência do evento não depende dos demais.

(4) Para você, variável aleatória tem alguma relação com função (Matemática)?

Sim

(5) Em caso de afirmativo no item (4), elabore um "conceito" variável aleatória e variáveis independentes.

Variável aleatória → Um experimento ocorre ao acaso

Variáveis independentes → Um experimento que não depende dos demais.

- RESPOSTA DE CA

Atividade 1

① O que você entende por variável (ou variáveis)?

Valores que quero obter.

② É por aleatoriedade?

Não ter uma ordem para escolher alguma coisa, tomar algo ao acaso.

③ É independência de dois ou mais eventos?

É quando dois ou mais eventos podem ser "dependentes" sem ter um contexto sobre, ou melhor, podemos analisá-lo sem nos preocupar com o outro.

④ Para você variável aleatória tem alguma relação com função (matemática)?

Não, pois a variável aleatória pode ser analisada separadamente em um intervalo finito, além de não seguir uma "lei" (ordem).

⑤ Em caso afirmativo no item (4) elabore um "conceito" para variável aleatória e variáveis aleatórias independentes.

• RESPOSTA DE FA

Atividade 1

de matemática

① O que você entende por variável (ou variáveis)?

Variável é a incógnita, ou seja, o que eu quero achar. As variáveis podem assumir infinitos valores.

② E por aleatoriedade?

Aleatoriedade é o acaso. Num evento aleatório, não há escolha e sim a ocorrência é ao acaso.

③ E independência de dois ou mais eventos?

Independência de dois eventos ocorre quando um não depende do outro para ser analisado.

④ Para você, a variável aleatória tem alguma relação com funções (matemática)?

Sim, pois assim como na função matemática, a variável aleatória pode assumir infinitos valores e está associada a outra variável.

⑤ Em caso afirmativo no item (4), elabore um "conceito" para variável aleatória e variáveis aleatórias independentes.

Variável aleatória → As variáveis aleatórias podem assumir infinitos valores em um intervalo.

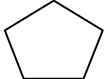
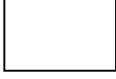
Variáveis aleatórias independentes → Podem assumir infinitos valores, porém sem depender do valor da outra variável.

ANEXO 3 – Questionário realizado na situação-problema apresentado aos licenciandos

5ª Série do Ensino Fundamental

Responda às questões e, em seguida, avalie sua resposta.

	Resposta	Estou seguro de que está correto	Acredito que está correto	Aposto 50% que está correto	Acredito que está incorreto	Estou seguro de que está incorreto	Pontuação (Espaço exclusivo dos pesquisadores)
Exemplo: 6×4	24	X					
1. Hoje Carolina completa 9 anos. Será que ela já viveu 3.000 dias de vida? 							
2. Dona Maria saiu de casa com uma certa quantia na bolsa. Ela gastou 17 reais no supermercado e, no caminho de volta, retirou 45 reais em uma caixa eletrônica. Se ela chegou em casa com 52 reais, com quanto saiu? 							
3. Qual destes decimais é maior? 1,7 0,46 1,0 1,56							
4. Agora responda: 2,4 tem quantos décimos a menos que 3?							
5. Responda: Quantos gramas são 0,5 Kg?							
6. O feirante comprou um saco com 3 Kg de feijão. Ele vai vender sacos de 0,5 Kg de feijão. Quantos sacos de 0,5 Kg ele pode fazer com a quantidade que tem?							
7. Imagine que 1 Kg de carne custe R\$ 8,00. Quanto você pagará se comprar 3,5 Kg de carne?							
8. Quantas latas de água de 8 litros são necessárias para encher uma caixa d'água na qual cabem 296 litros?							
9. Seu Luiz é caminhoneiro. Ele leva cargas de uma cidade para outra. Ele acabou de fazer uma viagem do Rio de Janeiro a Salvador em 4 dias, percorrendo a mesma distância cada dia. Quantos quilômetros seu Luiz rodou por dia?							
10. Uma escola tem 630 alunos. $\frac{1}{3}$ deles está na 5ª série. Quantos são esses alunos? 							
11. Em uma padaria, o preço da pizza inteira é R\$16,00. Mas, se você quiser, pode comprar fatias de pizza, pagando um preço promocional. Cada pizza é dividida em 8 partes iguais. Qual o preço de $\frac{1}{8}$ (um oitavo) de uma pizza? E o de 3 fatias dessa pizza? 							

5ª Série do Ensino Fundamental (cont.)	Resposta	Estou seguro de que está correto	Acredito que está correto	Aposto 50% que está correto	Acredito que está incorreto	Estou seguro de que está incorreto	Pontuação (Espaço exclusivo dos pesquisadores)
12. Uma lojinha tinha um estoque de 1850 toalhas e não conseguia vendê-las. Então, numa terça-feira, o gerente reduziu o preço delas em 50%. Nesse dia, vendeu 418 toalhas. A notícia se espalhou e, no dia seguinte, ele vendeu outras 837. Após essas vendas, quantas toalhas sobraram no estoque da loja?							
13. Faça esta multiplicação. Verá um resultado surpreendente. $123456789 \times 9 =$ 							
14. Thiago e sua família viajaram e, para evitar problemas, ele fechou o registro de entrada de água de sua casa. Nisso, a caixa da descarga do banheiro começou a vazar, perdendo 9 litros de água por hora. Antes de começar o vazamento, a caixa d'água continha 774 litros de água. Em quantas horas ela estará vazia?							
15. Mauro tem 1,27 m de altura. Crescendo mais 5 centímetros, que altura ele terá? Depois disso, quanto faltará para Mauro ter 1,4 m de altura? 							
16. Para ir ao trabalho, Carlos gasta R\$ 1,30 no ônibus. No trabalho, ele paga R\$ 4,80 pelo almoço e, à tarde, ainda gasta R\$ 2,50 em um lanche. Depois, ele volta para casa, gastando de ônibus o mesmo que na ida. Quanto ele gasta no total?							
17. No sítio de seu Paulo, o galinheiro tem forma quadrada e seus lados medem 11 m. Quantos metros de tela ele gastou para fazer a cerca do galinheiro? 							
18. Seu Orlando faz cálculos porque vai cercar o muro do terreno ao lado. Ao todo, quantos metros de muro serão construídos se os lados medem: 15,65m; 5,10m; 6,35m; 8,70m; 12,30m? 							
19. O comprimento de um retângulo é o dobro de sua largura. Seu perímetro é de 30 cm. Quanto mede o lado menor desse retângulo? Ajuda: esse lado menor mede menos de 7 cm. 							
20. Que dia da semana é hoje? Agora descubra, sem olhar no calendário, que dia da semana será daqui a 32 dias. 							

ANEXO 4 – Dados do enunciado da atividade 2 para os licenciandos

<u>Escola</u>									
<u>Particular:</u>									
Turma	5 A	Turma	5 B	Turma	5 C	Turma	5 D	Turma	5 E
Alunos:	Notas:	Alunos:	Notas:	Alunos:	Notas:	Alunos:	Notas:	Alunos:	Notas:
1	40	1	38,5	1	37	1	40	1	40
2	40	2	37,5	2	36,5	2	40	2	40
3	39,5	3	36,5	3	36	3	39,5	3	40
4	39,5	4	35,5	4	35	4	38,5	4	39,5
5	38	5	35	5	35	5	38	5	38
6	38	6	35	6	33,5	6	38	6	36,5
7	38	7	35	7	33	7	34	7	32
8	37,5	8	35	8	33	8	32	8	31,5
9	37	9	34	9	32,5	9	32	9	26,5
10	37	10	34	10	32,5	10	31,5	10	25
11	36,5	11	34	11	32,5	11	31	11	22,5
12	36	12	34	12	32	12	31	12	17,5
13	36	13	33,5	13	31,5	13	30,5	13	17
14	35,5	14	33	14	31	14	30	14	4,5
15	35	15	33	15	31	15	29		
16	34,5	16	32,5	16	30,5	16	28,5		
17	30,5	17	32	17	30,5	17	27		
18	28,5	18	31,5	18	30	18	26		
19	28,5	19	31	19	30	19	26		
20	25	20	30,5	20	29	20	23		
21	24,5	21	38,5	21	28	21	23		
22	24,5	22	27,5	22	27,5	22	22,5		
23	20	23	27	23	27	23	22,5		
		24	27	24	36,5	24	14		
		25	24,5	25	26	25	12		
		26	23,5	26	26	26	9		
		27	23,5	27	25,5				
		28	23,5	28	23,5				
		29	21,5						
		30	18						
		31	15,5						
Média da turma	33,89	Média da turma	30,68	Média da turma	31,14	Média da turma	28,79	Média da turma	29,32

**Escola
Pública:**

Turma	5 A	Turma	5 B	Turma	5 C	Turma	5 D	Turma	5 E
Alunos:	Notas:	Alunos:	Notas:	Alunos:	Notas:	Alunos:	Notas:	Alunos:	Notas:
1	34	1	38	1	37,5	1	30,5	1	30
2	33	2	36	2	34,5	2	27	2	29
3	30	3	34	3	33	3	25	3	28
4	25	4	34	4	31	4	24	4	26
5	24	5	33,5	5	31	5	22,5	5	26
6	23,5	6	32,5	6	31	6	21,5	6	25,5
7	23,5	7	32	7	29,5	7	20,5	7	25
8	21,5	8	30,5	8	28	8	19	8	23
9	20	9	30	9	28	9	18,5	9	22
10	18,5	10	29	10	27,5	10	17,5	10	21,5
11	16,5	11	28	11	26	11	16,5	11	21,5
12	16	12	26,5	12	25	12	16,5	12	21,5
13	14,5	13	25,5	13	24	13	16,5	13	21,5
14	13,5	14	23,5	14	23,5	14	15	14	21,5
15	12,5	15	23	15	23	15	14,5	15	21,5
16	11	16	21,5	16	23	16	14	16	20,5
17	11,5	17	21	17	22,5	17	11,5	17	20
18	9,5	18	19,5	18	22	18	7,5	18	17,5
19	9,5	19	19	19	21	19	6	19	17,5
20	9	20	19	20	20	20	5,5	20	17
21	8,5	21	16,5	21	20	21	5	21	16
22	4	22	16	22	19			22	15,5
		23	15,5	23	18,5			23	13,5
		24	14	24	17,5			24	13
		25	13,5	25	17			25	11,5
		26	13	26	16,5			26	9
		27	13	27	14			27	7,5
		28	12,5	28	12,5			28	4,5
		29	11,5	29	12,5			29	0
		30	10	30	12			30	0
		31	1,5	31	11,5				
				32	7				
Média da turma	17,68	Média da turma	22,35	Média da turma	22,97	Média da turma	16,88	Média da turma	18,22

ANEXO 5 – Resposta de AO para atividade 2

Atividade 2:

- 1) Forma escolhida: 5B - rede particular $\bar{x} = 30,68$ $\tilde{x} = 32,75$
 5B - rede pública $\bar{x} = 22,35$ $\tilde{x} = 22,25$

A média encontrada na 5B (particular) não é bem representativa; já que no total de 31 alunos, 20 alunos estão acima da média, por isto, há uma discrepância de dados visto que a menor nota da turma é 15,5.

Na rede pública, a média é representativa na pesquisa visto que, a medida de tendência central aproxima do valor central do rol.

Atividade 2:

1) Particular:

- 5A [28,01965 ; 39,96295] Proporção: $17/23 = 73,9130$
 5B [24,69193 ; 36,66287] Proporção: $21/31 = 67,74193$
 5C [27,51467 ; 34,77113] Proporção: $16/28 = 57,14285$
 5D [20,40617 ; 37,17083] Proporção: $17/26 = 65,38461$
 5E [18,34834 ; 40,29446] Proporção: $11/14 = 78,57142$

Pública:

- 5A [9,35293 ; 26,01067] Proporção: $16/22 = 72,7272$
 5B [13,16677 ; 31,54283] Proporção: $18/31 = 58,06451$
 5C [14,99624 ; 39,94116] Proporção: $20/32 = 62,5 \%$
 5D [9,80303 ; 23,95897] Proporção: $13/21 = 61,9047$
 5E [10,24619 ; 26,18721] Proporção: $22/30 = 73,333$

ANEXO 6 - Resposta de BM para atividade 2

Atividade 2- Anexo 1.

Turma escolhida:

5D (particular) : $\bar{X} = 28,79$

5D (pública) : $\bar{X} = 16,88$

Analisando os dados vemos uma amplitude grande (na escola particular o maior valor é 40 e o menor 9; na pública a maior é 30,5 e a menor 5).

Isto pode deslocar a média e fazer com que ela não seja representativa. É preciso analisar juntamente com outras medidas de tendência central, como por exemplo a mediana.

Na escola particular, a turma tem $\bar{X} = 28,79$ e $\tilde{X} = 30,5$.

Na escola pública : $\bar{X} = 16,88$ e $\tilde{X} = 17,0$.

Notamos que os valores estão próximos, logo podemos considerar que a média é representativa nos dois casos.

ANEXO 7 - Resposta de FA para atividade 2

Atividade 2

- 1) As turmas selecionadas foram 5C particular e 5C pública. A turma 5C particular apresentou $\bar{x}$ de 31,14 enquanto a 5C da pública apresentou $\bar{x}$ de 22,97.

As médias encontradas estão representando bem o conjunto de dados, uma vez que não fugiu muito do valor da mediana da cada turma, que foi na 5C particular 31 e na pública 23.

2) Particular: 5C

$$[27,51467; 34,77113] \text{ proporção } \frac{16}{28} =$$

$$57,142857.$$

Pública: 5C

$$[14,99624; 29,94116] \text{ proporção } : \frac{20}{32}$$

$$62,5\%$$

ANEXO 8 – Atividade 3 realizada por AO

Atividade 3

Respostas.

1) Para o número total de 136 alunos da instituição pública, a média do nível de autoestima é 19,8676 e o desvio padrão, 8,32. E, para 2 alunos da instituição particular, tem-se que a média do nível de autoestima é 30,082 e o desvio padrão 6,96.

1) Pelos dados aleatórios extraídos do EXCEL, partir da amplitude total de cada uma das instituições, nos foi possível anotar as respectivas médias e seus respectivos desvios padrão, conforme anexo.

(Construído no EXCEL (conforme anexo))

Gráficos construídos no EXCEL.

Usando as médias das 30 amostras aleatórias entre as escolas pública e particular, foi possível observar que há maior variação nas médias da escola pública, quanto a escola particular os valores são mais próximos.

a) tabela abaixo, realizada no EXCEL.
(Ver anexo)

b) Comparando as médias das escolas públicas e particulares neste caso com dados estatísticos, 600 variáveis, das 30 amostras, para escola pública, cuja média foi 14,59 e o desvio padrão, 10,41 e, o mesmo número de variáveis para escola particular, cuja a média foi 22,09 e o desvio padrão, 10,16, foi possível constatar que as médias tem seus valores bem próximos enquanto os desvios padrões aproximam-se entre si, sendo que os das amostras são bem menores que os da população.

ATIVIDADE 3 – ITEM 3: Dados realizados por AO no *Excel*:

Escola Pública:

$\bar{x}$	20,49	20,28	18,55	19,46	22,96	19,78	20,97	18,58	21,33	18,92
S	10,31	9,45	9,71	12,23	9,40	11,25	8,89	11,14	11,10	12,45

$\bar{x}$	15,12	21,58	19,45	14,17	23,74	20,85	18,94	17,90	23,27	22,00
S	11,09	11,03	10,38	10,09	9,78	10,62	7,79	10,31	12,68	11,32

$\bar{x}$	23,13	19,49	23,42	22,49	20,35	21,04	23,36	12,31	20,92	22,46
S	10,46	11,12	9,69	10,10	10,02	9,03	11,83	9,24	11,05	10,94

Escola Particular:

$\bar{x}$	21,72	24,91	21,64	22,80	21,25	20,77	22,54	24,86	24,88	20,48
S	10,89	9,27	10,66	9,83	10,62	10,93	10,24	8,31	10,95	10,19

$\bar{x}$	24,33	16,63	24,66	21,64	18,79	22,22	20,82	22,92	17,28	27,30
S	11,46	9,27	8,43	9,85	11,65	8,33	9,89	9,64	9,33	8,72

$\bar{x}$	21,31	21,06	23,05	22,73	25,59	21,92	21,30	20,53	19,64	23,03
S	10,15	10,05	12,30	10,36	9,38	10,81	11,36	10,25	10,13	9,65

ATIVIDADE 3 – ITEM 4: GRÁFICOS REALIZADOS POR AO:

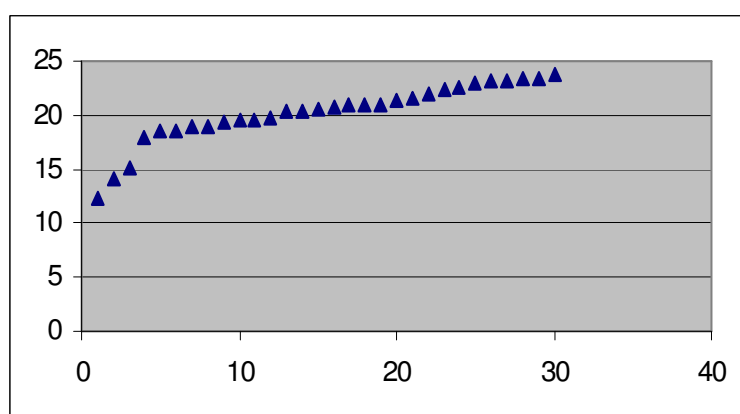


GRÁFICO 1: Médias das 30 amostras da Escola Pública
Fonte: Gráfico elaborado no EXCEL por AO.

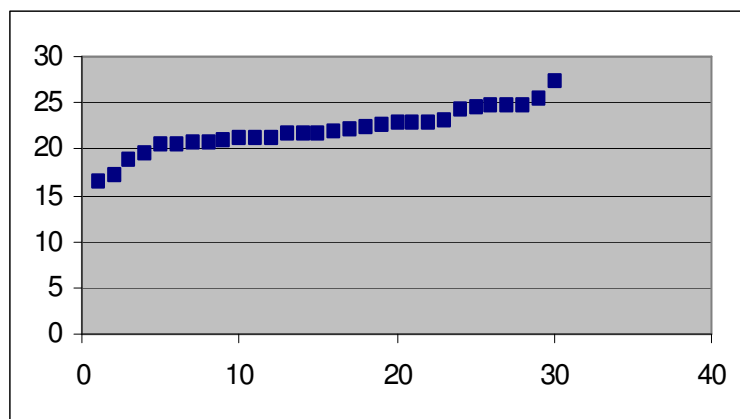


GRÁFICO 2: Médias das 30 amostras da Escola Particular
Fonte: Gráfico elaborado no EXCEL, por AO.

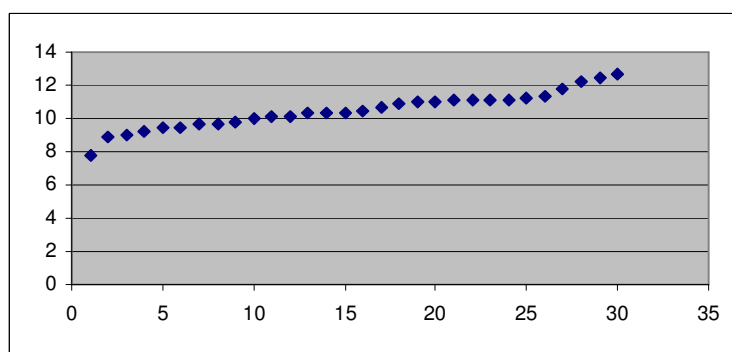


GRÁFICO 3: Desvios-padrão das 30 amostras da Escola Pública.
Fonte: Gráfico elaborado no EXCEL, por AO.

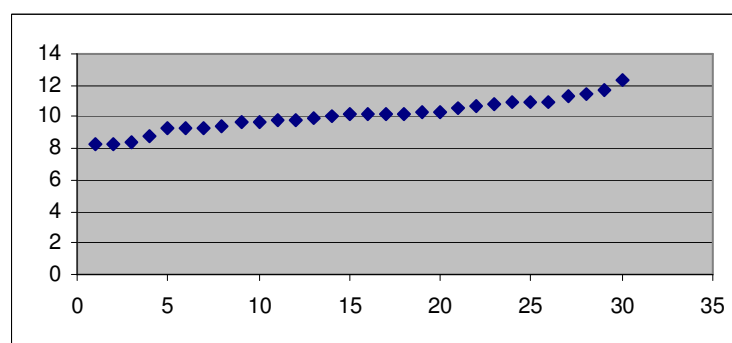


GRÁFICO 4: Desvios-padrão das 30 amostras da Escola Particular.
Fonte: Gráfico elaborado no EXCEL, por AO.

ATIVIDADE 3 – ITEM 7: GRÁFICOS REALIZADOS POR AO:

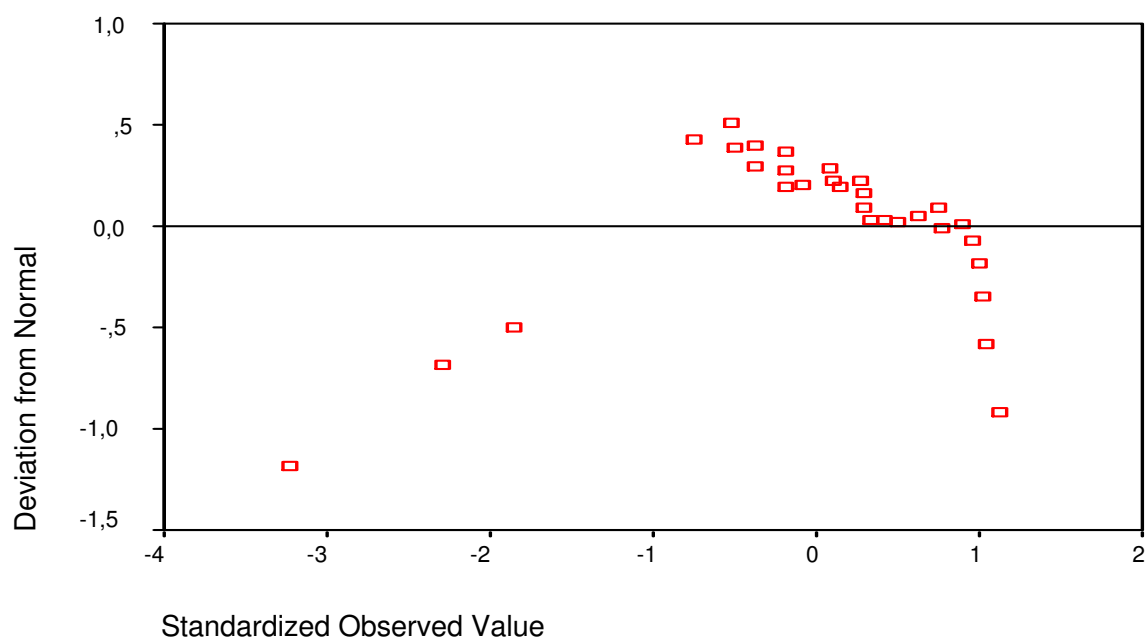


GRÁFICO 5: Distribuição das médias de dados de 60 observações da Escola Pública por AO.
 Fonte: Gráfico elaborado por AO com utilização de EXCEL e SPSS.

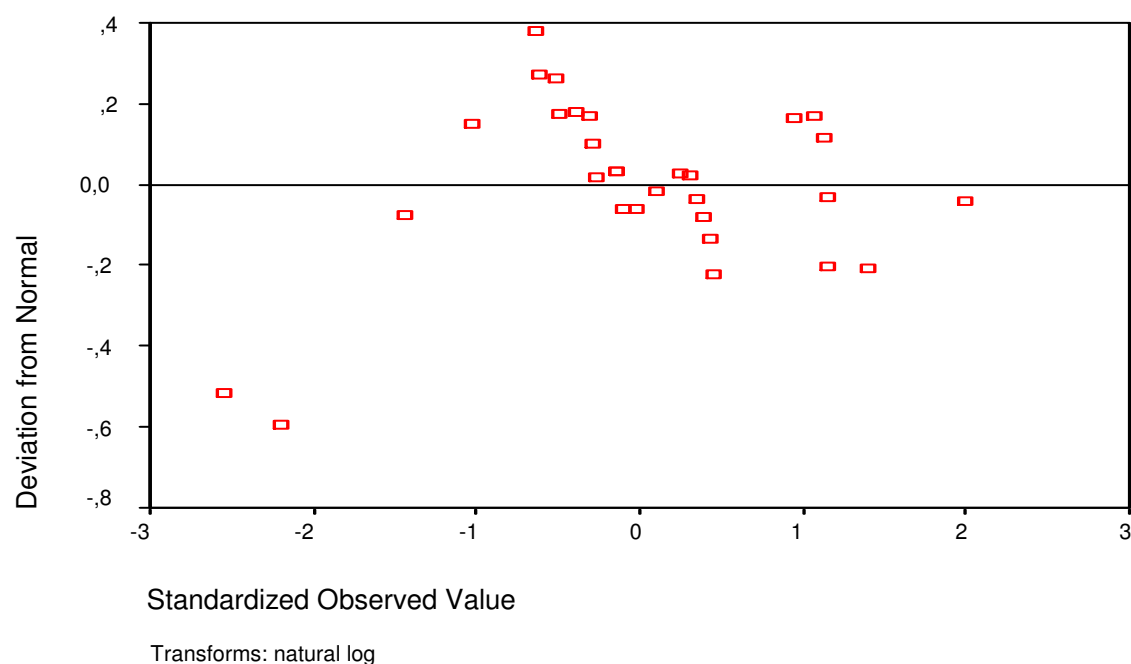


GRÁFICO 6: Distribuição das médias de dados de 60 observações da Escola Particular por AO.
 Fonte: Gráfico elaborado por AO com utilização de EXCEL e SPSS.

ANEXO 9 - Atividade 3 realizada por BM

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	,00	38,00	20,6500	11,59526
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	1,50	33,00	20,7250	8,10860
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	4,00	34,00	22,2500	8,30266
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	11,50	33,50	21,9000	6,98419
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	7,50	33,50	21,8750	7,63626
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	9,50	30,50	19,5500	6,76660
Valid N (listwise)	20				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PIIIBI TC:	20	4,00	31,00	20,5250	8,16681
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PIIIBI TC:	20	1,50	36,00	18,2750	9,49304
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PIIIBI TC:	20	9,00	38,00	20,1250	8,16423
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PIIIBI TC:	20	,00	38,00	19,3750	10,69416
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PIIIBI TC:	20	4,50	37,50	22,2000	8,34834
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PIIIBI TC:	20	5,00	38,00	20,9750	9,42348
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PIIIBI TC:	20	4,50	34,50	17,7250	7,28458
Valid N (listwise)	20				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	1,50	34,50	19,3750	9,79376
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	,00	36,00	19,4000	8,06324
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	5,00	32,00	18,6250	7,28711
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	7,00	33,00	19,9250	6,57402
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	,00	38,00	17,7500	11,07332
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	,00	38,00	18,7250	8,04506
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	1,50	33,50	19,7000	9,00351
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	9,00	32,00	19,9000	6,73092
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	10,00	37,50	19,9250	7,85104
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	,00	32,00	19,2250	8,78467
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	,00	33,50	15,9500	8,83310
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	7,00	34,50	17,4750	7,90815
Valid N (listwise)	20				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	,00	34,00	19,4750	9,92137
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	7,50	31,00	19,7750	7,08496
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	9,00	34,00	22,2500	7,22660
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	6,00	34,00	17,9250	7,27337
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	7,00	34,00	20,2750	7,30622
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PUBLIC	20	7,00	37,50	19,0250	8,77867
Valid N (listwise)	20				

ESCOLA PARTICULAR

Descriptives

Descriptive Statistics

		N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL		20	9,00	40,00	30,0250	8,62596
Valid listwise		20				

Descriptives

Descriptive Statistics

		N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL		20	9,00	40,00	30,6250	6,38847
Valid listwise		20				

Descriptives

Descriptive Statistics

		N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL		20	12,00	40,00	30,3250	7,67648
Valid listwise		20				

Descriptives

Descriptive Statistics

		N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL		20	21,50	40,00	30,6000	6,06890
Valid listwise		20				

Descriptives

Descriptive Statistics

		N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL		20	9,00	40,00	29,3000	7,42400
Valid listwise		20				

Descriptives

Descriptive Statistics

		N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL		20	4,50	39,50	30,4250	8,67585
Valid listwise		20				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	17,50	40,00	33,1250	6,47642
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	14,00	40,00	28,5500	7,84035
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	4,50	40,00	30,3500	9,38938
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	9,00	38,00	29,8750	7,42661
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	14,00	40,00	31,1750	7,00240
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	14,00	40,00	29,6000	7,42613
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	4,50	38,00	28,3500	8,88390
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	22,50	40,00	33,5250	4,79165
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	12,00	40,00	30,2000	7,63372
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	9,00	40,00	29,4750	8,45962
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	23,00	38,50	31,0500	5,42388
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	12,00	40,00	30,1250	6,91656
Valid N (listwise)	20				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	14,00	39,50	30,9250	6,29176
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	17,00	40,00	29,0250	6,39382
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	14,00	40,00	28,3500	6,96816
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	18,00	39,50	31,3500	5,34371
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	17,50	39,50	30,4250	6,07903
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	14,00	39,50	31,9750	6,64608
Valid N (listwise)	20				

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTICUL	20	14,00	40,00	30,5250	6,00104
Valid N (listwise)	20				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PARTECUL	20	15,50	40,00	31,1500	7,74104
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTECUL	20	4,50	39,50	28,7500	8,48606
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTECUL	20	22,50	38,00	30,9000	4,88984
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTECUL	20	20,00	40,00	31,7750	5,52381
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTECUL	20	4,50	40,00	28,7750	9,55245
Valid N (listwise)	20				

Descriptives**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
PARTECUL	20	23,00	40,00	31,8250	5,92558
Valid N (listwise)	20				

GRÁFICO 1

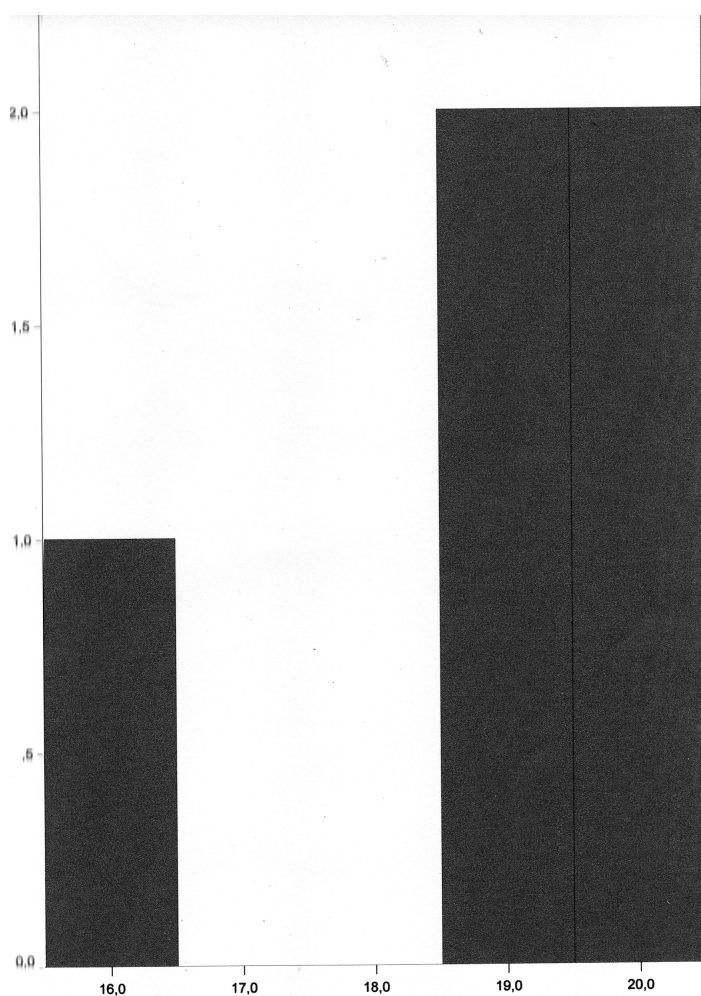


GRÁFICO 2

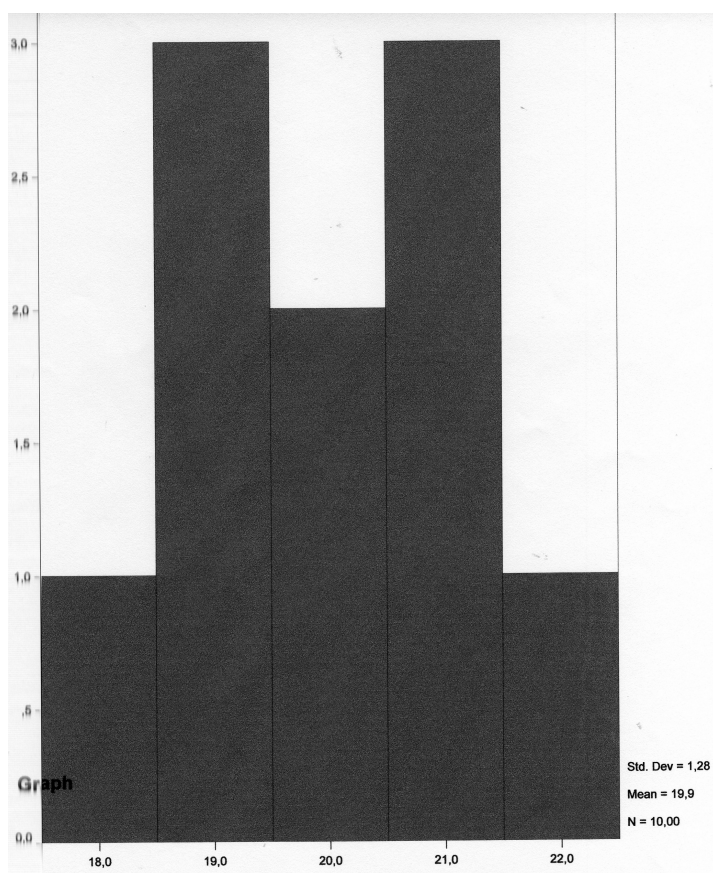


GRÁFICO 3

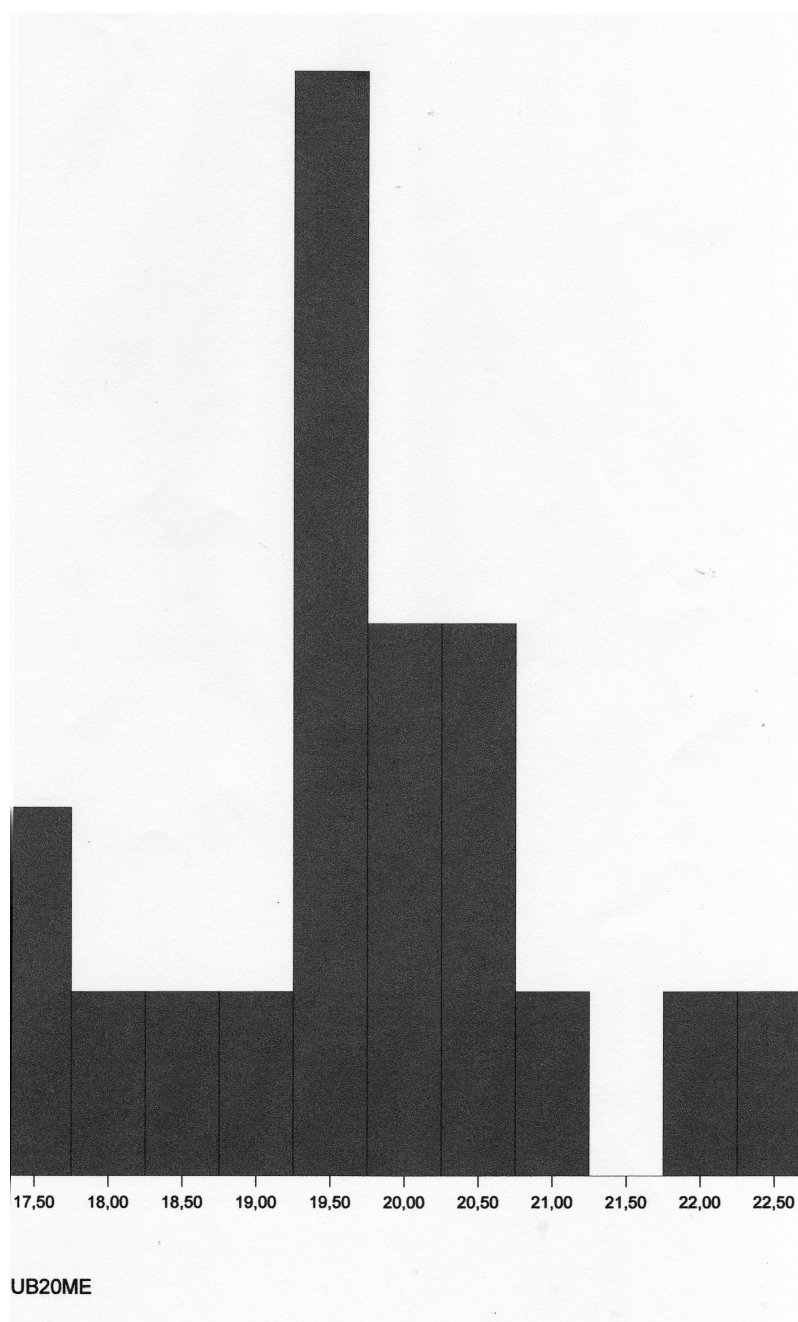


GRÁFICO 4

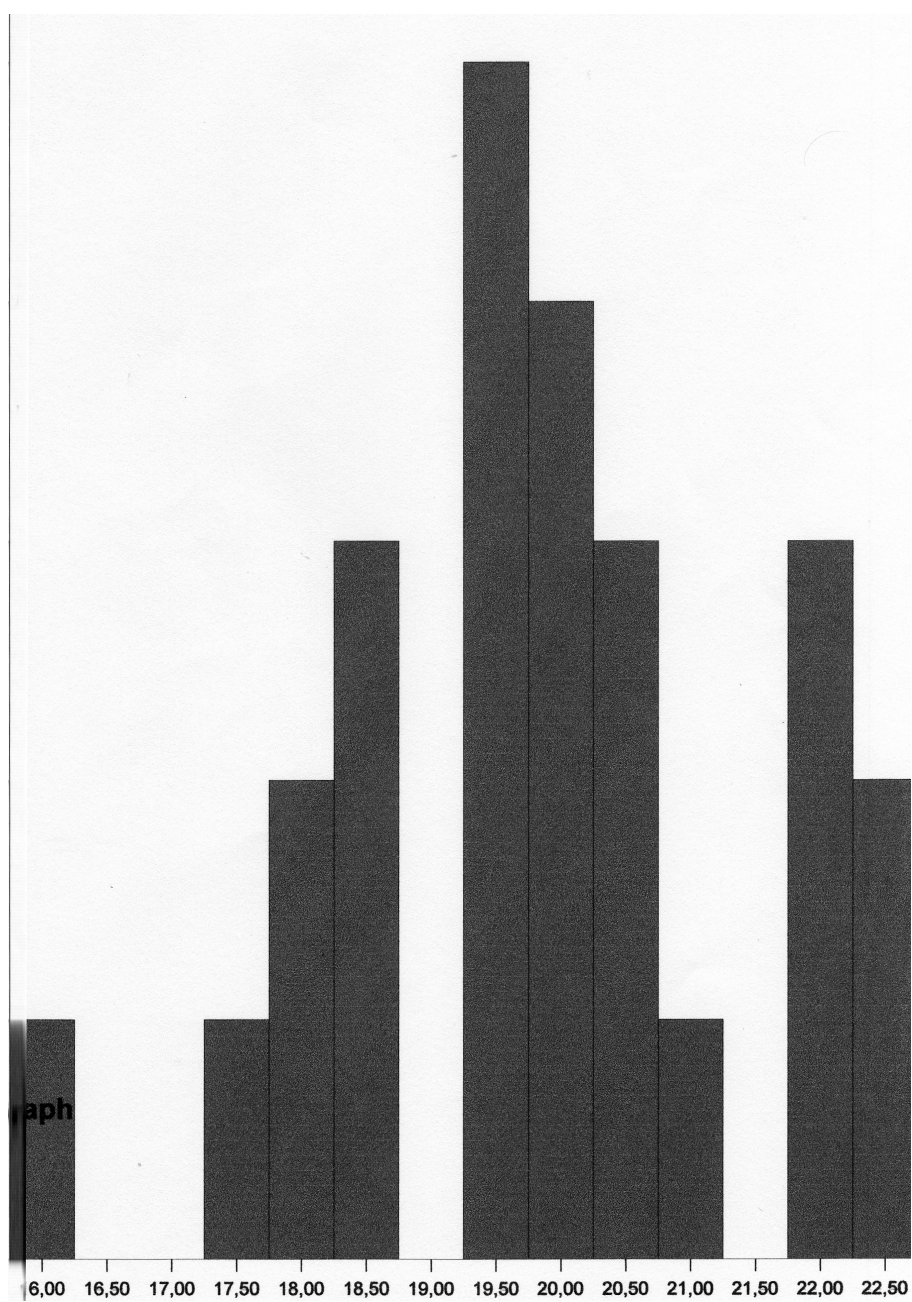


GRÁFICO 5



ANEXO 10 - Atividade 3 realizada por FA

ATIVIDADE 3

Siga as instruções e responda ao que se pede:

- 1) Determinar a média, $\bar{x}$, e o desvio padrão, s , para os dados da Instituição Pública e Instituição Particular.

Particular: $\bar{x} = 30,36$; $s = 6,96497$

- 2) Sortear 20 elementos de cada instituição e calcular a média, $\bar{x}$, e o desvio padrão, s . Repetir o processo ~~5~~₃₀ vezes.

- 3) Construir um banco de dados com todos os resultados obtidos, isto é, reunir todos os dados do grupo.

- 4) Construa um gráfico que represente as médias observadas. Faça o mesmo para os valores de s , desvio padrão. O que você observa?

Construímos um gráfico para um tamanho de amostra de 5, 10, 20, 25 e 30 e, daí, observamos

- 5) Determine a média dos valores obtidos em cada item (coluna) do banco de dados do item anterior, (4). O que você observa? Compare com o obtido no item 1.

- 6) Considere um novo banco de dados construído com os resultados das demais duplas. O que você observa a partir dos resultados obtidos para média e desvio padrão?

- 7) Construa o gráfico desse novo banco de dados, isto é, do item 6, e determine a média e o desvio padrão.

- 8) Compare com o obtido no item 4. O que você observa?