
**Tamanho Ótimo Amostral e Análise Fatorial e
Correlacional do Desempenho de Indivíduos sob
a Influência de Plataformas Computacionais de
Apoio ao Ensino**

Taffarel Brant Ribeiro



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

FACULDADE DE COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia

2018

Taffarel Brant Ribeiro

**Tamanho Ótimo Amostral e Análise Fatorial e
Correlacional do Desempenho de Indivíduos sob
a Influência de Plataformas Computacionais de
Apoio ao Ensino**

Tese de doutorado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação.

Área de concentração: Sistemas de Computação

Orientador: Renan Gonçalves Cattelan

Uberlândia

2018

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

R484t
2018 Ribeiro, Taffarel Brant, 1988-
 Tamanho ótimo amostral e análise fatorial e correlacional do
desempenho de indivíduos sob a influência de plataformas
computacionais de apoio ao ensino [recurso eletrônico] / Taffarel Brant
Ribeiro. - 2018.

 Orientador: Renan Gonçalves Cattelan.
 Tese (Doutorado) - Universidade Federal de Uberlândia, Programa
de Pós-Graduação em Ciência da Computação.
 Modo de acesso: Internet.
 Disponível em: <http://dx.doi.org/10.14393/ufu.te.2019.310>
 Inclui bibliografia.
 Inclui ilustrações.

 1. Computação. 2. Ensino auxiliado por computador. 3. Tecnologia
educacional. I. Cattelan, Renan Gonçalves, 1980- (Orient.) II.
Universidade Federal de Uberlândia. Programa de Pós-Graduação em
Ciência da Computação. III. Título.

CDU: 681.3

Maria Salete de Freitas Pinheiro - CRB6/1262



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO



Ata da defesa de TESE DE DOUTORADO junto ao Programa de Pós-graduação em Ciência da Computação da Faculdade de Computação da Universidade Federal de Uberlândia.

Defesa de Tese de Doutorado: PPGCO-08/2018

Data: 21/12/2018

Hora de início: 14 : 05

Discente: Taffarel Brant Ribeiro

Matrícula: 11423CCP009

Título do Trabalho: Tamanho Ótimo Amostral e Análise Fatorial e Correlacional do Desempenho de Indivíduos sob a Influência de Aprimoramentos Progressivos em Plataformas de Apoio ao Ensino.

Área de concentração: Ciência da Computação

Linha de pesquisa: Sistemas de Computação

Reuniu-se na sala 1B132, Bloco 1B, Campus Santa Mônica da Universidade Federal de Uberlândia, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Computação assim composta: Professores doutores: Fabiano Azevedo Dorça – FACOM/UFU, Márcia Aparecida Fernandes – FACOM/UFU, Sérgio Crespo Coelho da Silva Pinto - RCM/UFRPE, Rodrigo Lins Rodrigues – DED/UFRPE e Renan Gonçalves Cattelan – FACOM/UFU, orientador do candidato.

Ressalta-se que o Prof. Dr. Sérgio Crespo Coelho da Silva Pinto participou da defesa por meio de vídeo conferência desde a cidade de Rio das Ostras - RJ e o Prof. Dr. Rodrigo Lins Rodrigues da cidade de Recife - PE. Os outros membros da banca e o aluno participaram *in loco*.

Iniciando os trabalhos o presidente da mesa Prof. Dr. Renan Gonçalves Cattelan apresentou a Banca Examinadora e o candidato, agradeceu a presença do público e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu os conceitos finais.

Em face do resultado obtido, a Banca Examinadora considerou o candidato **aprovado**.

Esta defesa de Tese de Doutorado é parte dos requisitos necessários à obtenção do título de Doutor. O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, legislação e regulamentação interna da Universidade Federal de Uberlândia.

Nada mais havendo a tratar foram encerrados os trabalhos às 16 horas e 45 minutos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.

Participou por meio de vídeo conferência

Prof. Dr. Sérgio Crespo Coelho da Silva Pinto
RCM/UFRPE

Prof. Dr. Fabiano Azevedo Dorça
FACOM/UFU

Participou por meio de vídeo conferência

Prof. Dr. Rodrigo Lins Rodrigues
DED/UFRPE

Prof.^a Dr.^a Márcia Aparecida Fernandes
FACOM/UFU

Prof. Dr. Renan Gonçalves Cattelan
FACOM/UFU (orientador)

*“Querer não é poder.
Quem pôde, quis antes de poder só depois de poder.
Quem quer nunca há de poder, porque se perde em querer.”*

Bernardo Soares (semi-heterônimo de Fernando Pessoa)

Livro do Desassossego

Resumo

O uso de novas tecnologias no escopo educacional gera diversas questões acerca da eficiência dessas abordagens e quais benefícios elas propiciam ao âmbito acadêmico. Investigações nessa vertente abrangem uma linha de pesquisa denominada *Learning Analytics* e, na literatura, muitos trabalhos que analisam novas propostas tecnológicas têm como objetivo apenas observar as melhorias que o uso das ferramentas pode ocasionar. Tais pesquisas não analisam se o tamanho das amostras é robusto para garantir confiabilidade nos resultados ou se o aprimoramento das ferramentas tende a manter alguma influência sobre o desempenho discente. Com base nisso, esta tese determinou um tamanho ótimo amostral de 25 indivíduos para a análise do desempenho de alunos que não usam tecnologias de apoio ao ensino e de 20 discentes para turmas em contato com plataformas educacionais. Também foi desenvolvido um Gerenciador de Experimentos para se organizar a visibilidade das funcionalidades da plataforma Classroom eXperience (CX) e, empregando-se esse gerenciador, foram realizadas Análises de Variância Fatorial e Correlacional. Percebeu-se que o desempenho discente foi influenciado pela interação entre as funcionalidades do CX e as disciplinas cursadas pelos estudantes. Em todas as turmas de graduação, houve aumentos significativos no desempenho discente num comparativo entre a ausência do CX e seu uso junto às funcionalidades da plataforma. As disciplinas de graduação de natureza teórica e matemática também apresentaram correlações moderadas entre o nível de uso da plataforma e o desempenho dos estudantes. Deste modo, o emprego da plataforma influenciou positivamente as notas dos graduandos e inferiu-se que os alunos que mais interagiram com o CX também obtiveram melhores rendimentos em suas turmas. Na pós-graduação, não se observou diferença significativa no desempenho discente entre os níveis de uso do CX, nem a ocorrência de correlações que indicassem algo similar ao que ocorreu na graduação, embora também tenham ocorrido acréscimos no desempenho obtido pelos estudantes neste nível acadêmico.

Palavras-chave: Tamanho Ótimo de Amostra; Análise de Variância Fatorial; Correlação de Spearman; Learning Analytics.

Abstract

Usage of new technologies in educational scope raises several questions about the efficiency of these approaches and which benefits they provide to the academic field. Investigations in this area cover a line of research called *Learning Analytics* and, in the literature, many papers that analyze new technological proposals are only aimed at observing improvements that the use of tools can cause. Such researches do not analyze whether the sample size is robust to ensure reliability of results or whether the tool enhancement tends to maintain some influence over students' performance. Based on this, this thesis determined an optimal sample size of 25 students for the performance analysis of students who do not use teaching support technologies and of 20 students for classes in contact with educational platforms. An Experiments Manager was also developed to organize the visibility of Classroom eXperience (CX) platform functionalities and, using this Experiments Manager, a Factorial Analysis of Variance and a Correlation Analysis were performed. It was observed that students' performance was influenced by the interaction between CX functionalities and the courses taken by students. In all undergraduate classes, there were significant increases in student performance in a comparison between the absence of CX and its use with the platform functionalities. Theoretical and mathematical undergraduate courses also presented moderate correlations between the platform usage level and students' performance. Thus, the platform usage positively influenced the grades of undergraduate students and it was inferred that students who interacted more with CX also obtained the best grades in their classes. In graduate classes, there was no significant difference in students performance between CX levels of usage, nor the occurrence of correlations that indicated something similar to what happened with undergraduates, although there have also been increases in students' performance at this academic level.

Keywords: Optimal Sample Size; Factorial Analysis of Variance; Spearman's Rank Correlation; Learning Analytics.

Lista de ilustrações

Figura 1 – Representação da simetria da distribuição normal em relação à média amostral ao centro (Fonte: Adaptado de Triola (2014)).	28
Figura 2 – Salas de aula instrumentadas com dispositivos computacionais, tais como: lousas digitais, microfones, câmeras e projetores multimídia (Fonte: Brant-Ribeiro, Biase e Cattelan (2015)).	41
Figura 3 – Interface <i>Web</i> do Classroom eXperience (CX): (a) página de <i>login</i> e (b) tela de seleção de contexto (Fonte: Adaptado de Brant-Ribeiro e Cattelan (2015)).	42
Figura 4 – Interface <i>Web</i> do CX: (a) <i>home page</i> da plataforma e (b) tela de aula capturada (Fonte: Brant-Ribeiro e Cattelan (2015)).	44
Figura 5 – Estrutura do Gerenciador de Experimentos e interface de comunicação com ferramentas de apoio ao ensino (Fonte: Brant-Ribeiro e Cattelan (2018)).	46
Figura 6 – Diagrama de Componentes com destaque para o componente do Gerenciador de Experimentos.	48
Figura 7 – Acesso ao Gerenciador de Experimentos do CX: (a) <i>link</i> no menu suspenso da <i>home page</i> de usuários pesquisadores e (b) área de seleção de turmas cadastradas no sistema.	50
Figura 8 – Interface do Gerenciador de Experimentos do CX após a seleção de uma turma com estudantes matriculados.	52
Figura 9 – Telas de aulas capturadas empregando o CX: (a) aula com os módulos desabilitados para os estudantes e (b) aula com as funcionalidades de Socialização e Quizzes habilitadas.	52
Figura 10 – Relações entre os Coeficientes de Variação (CVs) e os tamanhos de subamostras para a variável desempenho dos estudantes em dez turmas de cursos da área computacional (Fonte: Brant-Ribeiro e Cattelan (2016)).	59

Figura 11 – Análise de Variância Fatorial 4x5 (<i>Two-way ANOVA</i>) dos fatores “Uso do CX” e “Disciplinas” sobre o desempenho discente.	67
Figura 12 – Influência da interação entre os fatores “Uso do CX” e “Disciplinas” sobre o desempenho dos estudantes.	69
Figura 13 – Análise Correlacional entre o desempenho obtido pelos estudantes e seus respectivos níveis de interação com o CX.	80

Lista de tabelas

Tabela 1	– Tabela <i>Visibility</i> da base de dados do Gerenciador de Experimentos com algumas tuplas contendo dados.	49
Tabela 2	– Coeficientes de Variação (CVs) unitários e de regressão obtidos a partir da análise das dez turmas utilizadas para o cálculo de tamanho ótimo amostral (Fonte: Adaptado de Brant-Ribeiro e Cattelan (2016)).	61
Tabela 3	– Análise de Variância Fatorial do desempenho discente com relação ao uso do Classroom eXperience (CX) em disciplinas de graduação e pós-graduação em cursos da área computacional (Fonte: Brant-Ribeiro e Cattelan (2018)).	68
Tabela 4	– Desempenho discente com relação ao uso do CX em disciplinas de graduação e pós-graduação em cursos da área computacional com enfoque nas disciplinas cursadas pelos estudantes (Fonte: Brant-Ribeiro e Cattelan (2018)).	70
Tabela 5	– Desempenho discente com relação ao uso do CX em disciplinas de graduação e pós-graduação em cursos da área computacional com enfoque nos níveis de utilização do CX (Fonte: Brant-Ribeiro e Cattelan (2018)).	72
Tabela 6	– Variáveis referentes aos fluxos interativos que os estudantes tiveram com a plataforma CX.	79
Tabela 7	– Análise da normalidade residual das quinze turmas que empregaram o Classroom eXperience (CX) durante os semestres letivos.	81
Tabela 8	– Correlação de Spearman entre o desempenho acadêmico obtido pelos estudantes e seus respectivos <i>logins</i> na plataforma CX.	81
Tabela 9	– Correlação de Spearman entre o desempenho acadêmico obtido por cada turma e as variáveis: Acesso às Aulas (AA), Socializações e Colaborações (SC) e Respostas aos Quizzes (RQ).	83

Lista de siglas

AA Acesso às Aulas

ABNT Associação Brasileira de Normas Técnicas

AEU Ambiente Educacional Ubíquo AEU's

ANOVA Análise de Variância

AOC Arquitetura e Organização de Computadores

AVA Ambiente Virtual de Aprendizagem AVAs

C&A Captura e Acesso

CV Coeficiente de Variação CVs

CX Classroom eXperience

DHS Diferença Honestamente Significativa

FACOM/UFU Faculdade de Computação da Universidade Federal de Uberlândia

IC Intervalo de Confiança

IDE *Integrated Development Environment* - Ambiente de Desenvolvimento Integrado

IHC Interação Humano-Computador

JSP JavaServer Pages

LA *Learning Analytics*

LM Lógica Matemática

LMS *Learning Management System* - Sistema de Gestão da Aprendizagem LMSs

MMCM Método de Máxima Curvatura Modificado

MS Módulo Social

RNA Rede Neural Artificial RNAs

RQ Respostas aos Quizzes

SC Socializações e Colaborações

SGBD Sistema de Gerenciamento de Banco de Dados

SISSI Shortcut In Sample Size Identification

SW Shapiro-Wilk

UB Unidade Básica

UbiComp Computação Ubíqua

ULSI *Ultra Large Scale Integration* - Integração em Escala Ultra Grande

VASCORLL Visualization and Analysis System for COnnecting Relationships of Learning Logs

Sumário

1	INTRODUÇÃO	19
1.1	Objetivos	21
1.2	Contribuições	22
1.3	Organização desta Tese	24
2	REVISÃO DA LITERATURA	27
2.1	Fundamentação Estatística	27
2.1.1	Normalidade e Homocedasticidade em Amostras	27
2.1.2	Reamostragem e Bootstrapping	29
2.1.3	Método de Máxima Curvatura Modificado	30
2.1.4	Análise de Variância Fatorial	31
2.1.5	Análise Correlacional	32
2.2	Trabalhos Relacionados	34
2.2.1	Estimativas de Tamanhos Amostrais	34
2.2.2	Learning Analytics para a Análise do Impacto de Plataformas de Apoio ao Ensino	36
2.3	Classroom eXperience	40
3	GERENCIADOR DE EXPERIMENTOS	45
3.1	Visão Geral da Proposta	45
3.2	Integração da Proposta ao CX	47
3.3	Interface <i>Web</i> do Gerenciador de Experimentos	50
4	ESTIMATIVA DE TAMANHO ÓTIMO AMOSTRAL PARA A ANÁLISE DO DESEMPENHO DISCENTE	55
4.1	Material e Métodos	55
4.2	Resultados e Discussões	57
4.3	Considerações Finais	61

5	ANÁLISE FATORIAL DO IMPACTO DE AEUS NO DE- SEMPENHO DISCENTE	65
5.1	Material e Métodos	65
5.2	Resultados e Discussões	68
5.3	Considerações Finais	73
6	ANÁLISE CORRELACIONAL ENTRE O DESEMPENHO DISCENTE E O NÍVEL DE INTERAÇÕES COM O CX . .	77
6.1	Material e Métodos	77
6.2	Resultados e Discussões	80
6.3	Considerações Finais	85
7	CONCLUSÃO	89
7.1	Resultados e Contribuições	89
7.2	Produção Bibliográfica	92
7.3	Limitações e Trabalhos Futuros	95
	REFERÊNCIAS	99

CAPÍTULO 1

Introdução

O uso de dispositivos tecnológicos propagou-se de modo intenso na sociedade, transformando os computadores em elementos onipresentes que se entrelaçaram por completo ao modo de vida das pessoas (WEISER, 1991; IWS, 2018). No âmbito acadêmico, o emprego de ferramentas que têm como intuito apoiar o processo de ensino/aprendizagem tornou-se uma atividade habitual, contribuindo para a validação de teorias pedagógicas ao possibilitar o desenvolvimento de novas plataformas de apoio ao ensino que empregam sistemas de hardware e software criados especificamente para esta finalidade (SMYTH; VOLKER, 2013; SCHMIDT et al., 2014; KNIGHT et al., 2015; GASEVIC et al., 2017; PUTJORN et al., 2017; CAI, 2018; ZHANG et al., 2018).

Ao se trabalhar com propostas tecnológicas implantadas em ambientes educacionais, é possível perceber que diversas questões despontam acerca da eficiência dessas abordagens e quais benefícios elas propiciam de fato ao escopo acadêmico. Trabalhos neste âmbito abrangem uma linha de pesquisa denominada *Learning Analytics* (LA), compreendida como o processo de se coletar e analisar dados de discentes¹ com o intuito de, por exemplo, estimar progressos acadêmicos, antever desempenhos futuros e compreender particularidades dos fluxos de aprendizagem (BARNEVELD; ARNOLD; CAMPBELL, 2012; CHATTI et al., 2014; YASSINE; KADRY; SICILIA, 2016). Além disso, o emprego da LA também fornece subsídios para atividades de projeto, desenvolvimento e refinamento de tecnologias educacionais, utilizando métodos de engenharia capazes de levantar requisitos de software que podem ser empregados para criar heurísticas específicas para a manutenção de sistemas no escopo acadêmico (BODILY et al., 2018).

¹ Para esta tese, os termos “discente”, “aluno” e “estudante” (e suas variações) possuem significados equivalentes e foram utilizados com o intuito de garantir uma melhor fluidez textual.

Tendo sido formalizada inicialmente em 2010, a área de LA é uma temática de pesquisa recente, originada a partir da interseção de diversas vertentes científicas, tais como a Psicologia, a Ciência da Computação, a Pedagogia e a Ciência da Aprendizagem (CHATTI et al., 2012; FERGUSON, 2012; DAWSON et al., 2014; KORTEMEYER, 2017). Em razão disso, tem como propósito elaborar e investigar técnicas para se recuperar e analisar informações educacionais, empregando procedimentos oriundos de diversos campos científicos, tais como a Estatística e a Visualização de Dados.

Pelo seu caráter contemporâneo e multidisciplinar, a área de LA ainda não dispõe completamente de um arcabouço bem fundamentado de heurísticas, métodos e regras para a realização e replicação de pesquisas científicas. Conforme pontuado por Chatti et al. (2014), o processo de aprendizagem tem se expandido em largas escalas, abordando, inclusive, variáveis de tempo, espaço e mídias utilizadas. Consequentemente, grandes volumes de dados a respeito de indivíduos em atividades de assimilação de conhecimento têm sido gerados de maneira contínua.

Tais dados podem ser interpretados como rastros que as pessoas deixam quando interação com ambientes educacionais e utilizados para se compreender como este fenômeno tende a acontecer e como otimizar o processo de aprendizagem em estudantes. Na literatura, trabalhos sobre desempenho² e adaptabilidade de usuários são frequentemente observados e abrem margem para diversas análises, como, por exemplo, a verificação de ganhos para estudantes que fazem uso cotidiano de determinadas ferramentas acadêmicas (SETTLE; DETTORI; DAVIDSON, 2011; MOURI; OGATA; UOSAKI, 2015).

No entanto, a maioria dessas pesquisas preocupa-se apenas em fazer comparações entre usuários, expondo que existem aumentos significativos de desempenho entre grupos analisados (AIHUA, 2010; WIELING; HOFMAN, 2010; DICKSON et al., 2012; MOURI; OGATA; UOSAKI, 2015), mas não dedicando-se em averiguar inicialmente se o tamanho das amostras utilizadas é robusto o suficiente para representar de modo fiel as populações de indivíduos das quais são retiradas. Investigações que não empregam tamanhos amostrais adequados estão fadadas a apresentarem resultados enviesados e análises inconclusivas, visto que as variáveis sob estudo possuem comportamentos inerentes às suas próprias naturezas experimentais.

² Nesta tese, os termos “desempenho” e “rendimento” possuem significados equivalentes e representam as notas finais obtidas pelos estudantes ao término dos períodos letivos.

Além disso, até mesmo em trabalhos que fazem o acompanhamento de ferramentas acadêmicas por períodos mais extensos e abordam quantidades maiores de usuários (SETTLE; DETTORI; DAVIDSON, 2011; DANIELSON et al., 2014; ASARTA; SCHMIDT, 2017), observa-se a realização repetitiva de análises semelhantes sobre as mesmas funcionalidades, havendo mudanças nos indivíduos observados, mas sem a elaboração de novas questões de pesquisa com características mais robustas, como, por exemplo, se a realização de aprimoramentos progressivos nas plataformas e a inclusão de novas funcionalidades também tendem a manter influências significativas sobre os estudantes sob investigação. Percebe-se, portanto, que na literatura os mesmos objetivos de pesquisa tendem a ser empregados de maneira exaustiva, o que causa certa previsibilidade de resultados e a impressão de esgotamento de assunto sobre a temática.

Existem diversos fatores que influenciam diretamente a precisão e a confiabilidade experimentais, dentre eles o nível de heterogeneidade dos ambientes investigados e do próprio material experimental sob análise (STEEL; TORRIE; DICKEY, 1980; GOMEZ; GOMEZ, 1984). Para se diminuir o erro experimental, é importante fazer uso de amostras com dimensões apropriadas e capazes de suprir as características das populações observadas. Além disso, mostra-se imprescindível empregar um método bem definido para se organizar todos os níveis de análise que compõem uma pesquisa, de modo a evitar inconsistências durante o processo de coleta de dados e, conseqüentemente, enviesamentos dos resultados obtidos nas análises. Por meio do emprego de tamanhos amostrais adequados e métodos de pesquisa que possibilitam uma correta organização dos processos experimentais, torna-se possível garantir às pesquisas um maior nível de precisão em suas análises e confiabilidade nos resultados.

1.1 **Objetivos**

Esta tese almejou inicialmente desenvolver um módulo de gerenciamento de experimentos (Capítulo 3, pág. 45), capaz de organizar o nível de visibilidade de funcionalidades experimentais e refinamentos presentes em um Ambiente Educacional Ubíquo (AEU) e fornecer subsídios para o planejamento e realização de pesquisas nesta vertente. Tal módulo foi planejado e implementado com o intuito de facilitar a execução de experimentos ao permitir a habilitação e desabilitação de funcionalidades específicas para grupos dis-

tintos de estudantes que cursaram disciplinas dos cursos de graduação e pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia (FACOM/UFU).

Em paralelo ao desenvolvimento do Gerenciador de Experimentos, esta pesquisa também objetivou investigar e determinar um tamanho ótimo de amostra (Capítulo 4, pág. 55) para a análise do desempenho de estudantes em contato com plataformas de apoio ao ensino, tais como AEU's e Ambientes Virtuais de Aprendizagem (AVAs). Além disso, enxergou-se também a necessidade de se determinar um tamanho ótimo amostral para analisar alunos de graduação e pós-graduação que não estão em contato com tais tecnologias educacionais, de modo a ser possível traçar um paralelo entre os resultados obtidos e se recomendar uma estimativa capaz de compreender ambas as circunstâncias.

Em seguida, almejou-se realizar duas investigações empregando-se o Gerenciador de Experimentos e o Tamanho Ótimo Amostral: uma Análise de Variância Fatorial 4x5 (*Two-way ANOVA*) (Capítulo 5, pág. 65) e um estudo empregando a Correlação de Spearman (Capítulo 6, pág. 77). A ANOVA Fatorial foi realizada para se atestar a existência de diferenças significativas no desempenho acadêmico de alunos em contato com um AEU denominado Classroom eXperience (CX), à medida que o mesmo foi aprimorado por meio da inclusão de novas funcionalidades gradativamente disponibilizadas para os discentes durante os semestres letivos.

Já o estudo que empregou a Correlação de Spearman foi realizado para se averiguar possíveis relações entre o desempenho acadêmico obtido pelos estudantes e seus fluxos de interação com a plataforma CX. Desta maneira, foi possível obter maiores indícios acerca do processo de evolução de uma plataforma de apoio ao ensino e o impacto que ela causou progressivamente no desempenho de estudantes em um ambiente acadêmico real.

1.2 Contribuições

Por meio da realização desta tese, foi possível desenvolver um módulo de Gerenciamento de Experimentos (Capítulo 3, pág. 45) que realizou a organização da visibilidade das funcionalidades da plataforma CX, fornecendo subsídios para o planejamento e realização deste trabalho. A partir do uso do gerenciador, foi possível empregar um método adequado para se analisar o impacto progressivo que o CX ocasionou no desempenho aca-

dêmico de estudantes que cursaram disciplinas em cursos graduação e pós-graduação da FACOM/UFU no decorrer dos semestres letivos. O uso do gerenciador se mostrou como uma prática adequada para as análises que foram realizadas nesta pesquisa e enxerga-se que o emprego deste módulo também pode ser expandido para outras investigações que demandem organizações detalhadas do nível de acesso que grupos distintos de indivíduos devem possuir ao utilizarem ferramentas de apoio ao ensino no cotidiano acadêmico.

Além disso, também foi possível determinar tamanhos ótimos de amostra para a análise do desempenho de estudantes com e sem contato com tecnologias educacionais, como AEU's e AVAs (Capítulo 4, pág. 55), o que contribui diretamente com a garantia de precisão experimental e o aumento da taxa de confiabilidade em resultados obtidos em pesquisas da área de LA. Determinou-se, deste modo, um tamanho ótimo amostral de 25 alunos para a análise do desempenho acadêmico em turmas que não utilizam tecnologias educacionais no cotidiano e de 20 estudantes para turmas que possuem contato com tecnologias educacionais durante os semestres letivos. A partir desta contribuição, foi possível perceber que a diminuição do Coeficiente de Variação (CV) das amostras não apresentou um comportamento proporcional à elevação da quantidade de observações das mesmas, havendo reduções expressivas na variabilidade até o momento em que as amostras alcançaram tamanhos aceitáveis, mas que tenderam à estabilidade quando maiores quantidades de observações foram acrescidas.

Deduziu-se, portanto, que a elevação do tamanho amostral manteve-se conveniente apenas até determinado ponto, a partir do qual o emprego de quantidades maiores de observações não se mostrou vantajoso, devido ao ganho escasso de precisão experimental. Tais estimativas mostram-se relevantes para pesquisas que realizam estudos acerca do desempenho de alunos em circunstâncias nas quais novas tecnologias são introduzidas, de modo que o emprego destes tamanhos amostrais possa garantir uma maior precisão e confiabilidade experimentais aos resultados obtidos.

Por meio da realização da Análise de Variância Fatorial 4x5 (*Two-way ANOVA*) (Capítulo 5, pág. 65), percebeu-se que o desempenho discente foi diretamente influenciado pela interação entre as funcionalidades habilitadas no CX e a disciplina cursada pelos estudantes. Num comparativo entre as amostras que não utilizaram o CX e aquelas que empregaram a plataforma com todas as funcionalidades habilitadas, houve elevações significativas no desempenho discente em todas as turmas de graduação investigadas. Na disciplina de pós-graduação analisada, não se observou diferença significativa no desem-

penho discente entre os níveis de uso do CX, embora também tenham ocorrido acréscimos gradativos no desempenho obtido pelos estudantes deste nível de ensino.

Durante a análise que empregou a Correlação de Spearman (Capítulo 6, pág. 77), inicialmente foram observadas as notas dos estudantes e a quantidade de *logins* que cada aluno realizou no sistema. A partir deste teste, percebeu-se que o nível de utilização do CX apresentou influência sobre a maneira com que os usuários interagiram com a plataforma. Tal indício foi ratificado a partir da análise de que todas as turmas que empregaram o sistema com todas as funcionalidades habilitadas apresentaram algum nível de correlação, ao passo que as amostras que empregaram o CX de modo isolado não alcançaram resultados similares.

Devido a isso, realizou-se por fim uma nova Correlação de Spearman entre o desempenho acadêmico dos estudantes e variáveis que compreenderam diferentes níveis de visibilidade e uso do CX. Nesta circunstância, para todas as turmas de graduação investigadas ocorreram correlações baixas e moderadas. As disciplinas teóricas e matemáticas apresentaram os melhores resultados dentre os demais e acredita-se que isso tenha ocorrido devido aos professores dessas disciplinas terem empregado o CX de modo cotidiano em suas aulas, substituindo o quadro negro/branco pela lousa digital do AEU.

1.3 Organização desta Tese

Os demais capítulos desta tese estão estruturados da seguinte maneira: inicialmente, no Capítulo 2 (pág. 27) é apresentado o universo de pesquisa no qual esta tese se encontra, fornecendo noções sobre a área de concentração do trabalho e fundamentações teóricas. No Capítulo 3 (pág. 45) é apresentado o Gerenciador de Experimentos que foi desenvolvido e integrado à plataforma computacional CX para se automatizar o processo de análise experimental desta pesquisa.

No Capítulo 4 (pág. 55) são discutidas as análises realizadas e os resultados obtidos referentes ao cálculo e à determinação de tamanhos ótimos amostrais para a análise do desempenho acadêmico de estudantes. No Capítulo 5 (pág. 65) é descrita a ANOVA Fatorial 4x5 que foi realizada para se investigar o impacto que uma plataforma de apoio ao ensino ocasiona no desempenho de estudantes.

No Capítulo 6 (pág. 77) são discutidos os resultados alcançados a partir da análise realizada empregando a Correlação de Spearman, que teve o intuito de verificar a existência de relações entre o desempenho acadêmico dos estudantes e suas interações com a plataforma educacional CX. Por fim, no Capítulo 7 (pág. 89) são concluídas as discussões realizadas junto aos resultados obtidos nesta tese, apresentando considerações finais, limitações e trabalhos futuros.

Revisão da Literatura

De modo a apresentar o universo de pesquisa no qual esta tese se encontra, este capítulo fornece noções essenciais sobre a área de concentração do trabalho. Inicialmente é realizado um embasamento no qual são abordados os conceitos estatísticos necessários para o entendimento desta pesquisa (Seção 2.1). Em seguida, é feita uma discussão acerca de trabalhos relacionados sobre o cálculo de tamanhos amostrais e o uso de *Learning Analytics* (LA) para a análise de plataformas de apoio ao ensino (Seção 2.2). Por fim, a plataforma educacional utilizada nesta pesquisa é apresentada na Seção 2.3.

2.1 Fundamentação Estatística

2.1.1 Normalidade e Homocedasticidade em Amostras

Dá-se o nome de população ao conjunto de unidades que são de interesse de determinada pesquisa, na qual amostras são coletadas e, com base na realização de análises, inferências são propostas (FIELD, 2013). Nesse contexto, amostras são agrupamentos de unidades que foram extraídas de populações e que têm como finalidade representar o universo de pesquisa do qual fazem parte.

Para a realização de análises empregando testes estatísticos paramétricos¹, mostra-se necessário averiguar inicialmente se as amostras observadas possuem determinadas pressuposições, tais como a normalidade de resíduos e a homogeneidade de variâncias

¹ Estatística Paramétrica é uma vertente da Estatística que emprega testes capazes de gerar estimativas mais precisas e robustas que os testes não-paramétricos para a realização de análises comparativas (GEISSER; JOHNSON, 2006).

(homocedasticidade) (TRIOLA, 2014). Uma amostra é dita “normal” quando a mesma se aproxima da distribuição normal de probabilidades (também chamada de distribuição gaussiana) e possui, deste modo, valores de média, mediana e moda coincidentes. Isso demonstra que o conjunto de observações coletadas da variável sob análise dispersa-se (ou distribui-se) de modo natural e simétrico em torno das medidas de tendência central da amostra (Figura 1). Tal característica permite a realização de testes estatísticos paramétricos, visto que os mesmos apresentam resultados confiáveis em circunstâncias que as amostras seguem a distribuição normal ou estão próximas dessa distribuição.

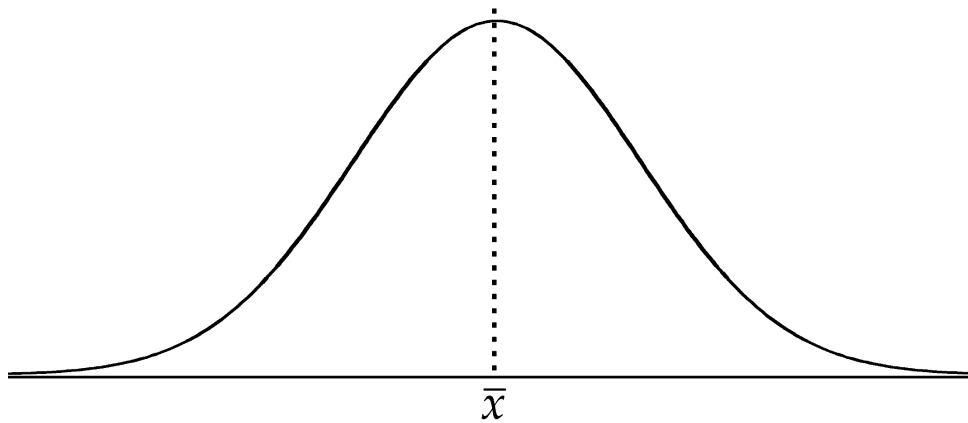


Figura 1 – Representação da simetria da distribuição normal em relação à média amostral ao centro (Fonte: Adaptado de Triola (2014)).

Existem testes específicos que podem ser empregados para se atestar a normalidade de resíduos em uma amostra, tais como Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk (SW) e Anderson-Darling. Dentre eles, o teste de SW se destaca por ser o mais indicado para amostras em que a quantidade de observações é menor que cinquenta unidades (RAZALI; WAH, 2011). A realização do teste de SW apresenta como resultado um parâmetro denominado valor-P (ou probabilidade de significância) e, caso este valor seja superior ao nível de significância α previamente estabelecido durante o delineamento da pesquisa², a amostra-alvo apresenta resíduos que seguem a distribuição normal e, portanto, suas observações permitem o uso de testes paramétricos.

O segundo pressuposto exigido para se empregar os testes paramétricos adequadamente intitula-se homocedasticidade e mostra-se necessário averiguá-lo quando se pretende realizar comparações entre amostras diferentes. Um teste de homocedasticidade

² Por convenção, normalmente emprega-se um valor de $\alpha = 0,05$ ou $\alpha = 0,01$.

indica a existência de variâncias homogêneas entre grupos distintos, ou seja, demonstra que a variabilidade entre as amostras comparadas é igual ou aproximadamente igual.

Na literatura, é possível encontrar testes específicos que podem ser utilizados para verificar o pressuposto de homocedasticidade, tais como Cochran, Levene, Bartlett e Brown–Forsythe. Dentre eles, o teste de Levene evidencia-se como o mais comumente empregado em pesquisas que realizam comparações entre amostras distintas, visto que o mesmo tende a apresentar maior robustez que os demais (RIBOLDI et al., 2014).

Assim como ocorre nos testes de normalidade, um teste de homogeneidade de variâncias também gera como resultado o parâmetro valor-P, que deve ser comparado ao nível de significância α inicialmente fixado para a realização da pesquisa. Caso seja obtida uma probabilidade de significância maior que α , as amostras comparadas apresentam a característica de homocedasticidade e, deste modo, torna-se possível realizar testes paramétricos comparativos que exigem o cumprimento dessa pressuposição.

2.1.2 Reamostragem e Bootstrapping

Reamostragem é o nome dado a um processo de simulação aleatorizada que, a partir de uma amostra já existente com n observações, permite a reorganização desses dados e a composição de inúmeras novas amostras de diferentes tamanhos por meio do reagrupamento das observações originais (GOOD, 2013). A reamostragem possibilita, dentre outras coisas, que cálculos de parâmetros médios sejam feitos empregando centenas ou milhares de amostras simuladas a partir de um mesmo conjunto de dados coletados previamente. Deste modo, torna-se possível calcular e estimar, por exemplo, o Coeficiente de Variação (CV) de cada uma das diversas amostras simuladas a partir dos mesmos dados e, em seguida, estimar um CV médio com base em todos os CVs já obtidos em cada uma das amostras simuladas.

O Método Bootstrapping é uma técnica de reamostragem que se baseia na geração aleatória de novas amostras empregando reposição de observações (EFRON, 1979). Deste modo, a cada sorteio de uma nova observação (dentre os dados já existentes na amostra original), todas as observações sempre possuirão a mesma probabilidade de serem sorteadas – até mesmo aquelas que já haviam sido selecionadas anteriormente. O Bootstrapping possibilita a simulação de circunstâncias em que as amostras possuem observações com

valores coincidentes, visto que propositalmente permite que as amostras resultantes das simulações contenham observações sorteadas mais de uma vez.

2.1.3 Método de Máxima Curvatura Modificado

A determinação de um tamanho ótimo amostral permite inferir e estabelecer uma quantidade de observações que garante confiabilidade em análises realizadas acerca de uma variável. Dentre os métodos existentes para se estimar o tamanho ótimo de amostras, o Método de Máxima Curvatura Modificado (MMCM), apresentado por Lessman e Atkins (1963) e Meier e Lessman (1971), é bastante empregado por fornecer resultados precisos ao estabelecer uma equação de regressão para explicar a relação entre os CVs e seus devidos tamanhos de amostra (SILVA et al., 2003).

Tal método também se destaca por ser capaz de ilustrar num plano cartesiano os CVs junto a seus respectivos tamanhos de amostra, realizando o cálculo do valor da abcissa referente ao ponto exato onde ocorre a máxima curvatura do modelo matemático e indicando o tamanho ótimo da amostra analisada. Para a realização do cálculo do tamanho ótimo amostral por meio do MMCM, inicialmente é necessário representar a relação entre cada CV e seu conjunto de observações obtidas por reamostragem empregando uma equação de regressão de potência (Fórmula 1):

$$CV_x = aX^{-b} \quad (1)$$

na qual: CV_x é o coeficiente de variação entre as subamostras de tamanho n , a é o coeficiente de variação de amostras de tamanho unitário, X é o tamanho das subamostras e b é o coeficiente de regressão (BARROS; TAVARES, 1995).

De modo a elucidar os parâmetros a e b utilizados na Fórmula 1: o coeficiente de variação unitário a indica o ponto no qual a linha de regressão intercepta o eixo y utilizando valores de x e y existentes – referentes aos dados já coletados na pesquisa. Já o coeficiente de regressão b retrata a taxa de mudança ao longo da linha de regressão, ou seja, representa o nível de variabilidade das amostras sob análise.

Após representar a relação entre cada CV e seu respectivo tamanho de amostra, é necessário calcular o valor referente ao ponto da abcissa no qual ocorre a máxima curvatura

do modelo utilizado, correspondente ao tamanho ótimo de amostra, denominado X_{MC} e obtido por meio da Fórmula 2:

$$X_{MC} = [a^2 b^2 (2b + 1) / b + 2]^{(1/2b+2)} \quad (2)$$

na qual: X_{MC} é o valor exato da abscissa que corresponde ao ponto de máxima curvatura do modelo, a é o coeficiente de variação de amostras de tamanho unitário e b é o coeficiente de regressão (MEIER; LESSMAN, 1971).

2.1.4 Análise de Variância Fatorial

A Análise de Variância (ANOVA) é um método paramétrico de comparação que possibilita verificar simultaneamente a existência de diferenças significativas entre as médias de três ou mais amostras independentes (TRIOLA, 2014). A ANOVA é útil quando se deseja averiguar se as diferenças observadas nas médias das amostras são, de fato, reais (e ocasionadas por algum elemento controlável, tal como o emprego de uma nova tecnologia) ou ocasionais (e provocadas por eventuais variabilidades das amostras). Deste modo, essa análise parte da pressuposição de que a casualidade gera apenas desvios leves, ao passo que diferenças maiores são acarretadas por causas significativas.

Uma ANOVA pode ser empregada para se investigar o efeito de um único elemento sobre outro(s) ou para se analisar a influência de vários elementos (ou fatores) sobre diversas variáveis dependentes – permitindo também analisar a existência de interações entre esses fatores controláveis. Em circunstâncias que envolvem a observação e a análise do efeito de mais de um fator sobre outro(s) elemento(s), esse método recebe o nome de ANOVA Fatorial (ou *Two-way* ANOVA).

Para a realização da ANOVA, inicialmente é necessário executar o teste F de Snedecor para verificar a existência de diferenças significativas entre as médias das amostras analisadas (SNEDECOR; COCHRAN, 1980). Esse teste também permite observar a presença de “interações” entre os fatores, ou seja, se os resultados obtidos na pesquisa são influenciados pelos fatores operando de modo conjunto e não isoladamente.

A realização do teste F de Snedecor apresenta como resultado a estatística F (referente à distribuição F de Fisher-Snedecor) e o parâmetro valor-P (referente à probabilidade de

significância). Caso seja obtida uma probabilidade menor que o nível de significância α fixado para a pesquisa, é possível inferir que existem pelo menos duas amostras com médias significativamente distintas dentre todos os grupos amostrais analisados.

Conforme Markowski e Markowski (1990) e Box (1953), o teste F de Snedecor possibilita rejeitar ou não a hipótese de igualdade de médias amostrais entre diversos níveis, porém não é capaz de indicar especificamente quais são os grupos que possuem as médias diferentes. Em razão disso, mostra-se necessário empregar um teste “post-hoc” para a comparação múltipla das médias. Na literatura, é possível encontrar diversos testes para essa finalidade, tais como os testes de Duncan, Bonferroni, Tukey e Scheffé.

Dentre os testes de comparações múltiplas existentes, o teste de Tukey se destaca por possuir bastante rigor estatístico ao realizar comparações entre todos os pares amostrais analisados (TUKEY, 1949). Esse teste também é conhecido como teste da Diferença Honestamente Significativa (DHS) de Tukey e é capaz de revelar toda e qualquer diferença entre as médias de grupos ao realizar a comparação de todas as amostras existentes numa pesquisa, executando análises consecutivas entre os pares amostrais até que todas as possibilidades combinatórias sejam comparadas (TRIOLA, 2014).

Assim como ocorre no teste F de Snedecor, o teste de Tukey também apresenta como resultado um valor-P. No entanto, obtém-se uma probabilidade de significância específica para cada um dos pares amostrais comparados. A partir da análise de cada um desses valores-P resultantes, é possível concluir se cada par de amostras realmente apresentou médias significativamente diferentes, bastando averiguar se a probabilidade obtida foi menor que o nível de significância α estabelecido para a pesquisa.

2.1.5 Análise Correlacional

Correlação é o nome dado a uma relação estatística que compreende a existência de uma interdependência entre duas variáveis distintas (TRIOLA, 2014). Essas variáveis são ditas “relacionadas” quando as modificações ocorridas numa delas provoca mudanças na outra e “correlacionadas” quando a recíproca também acontece. Na literatura, é possível encontrar diferentes testes de correlação que podem ser usados para medir o nível dessa relação, tais como Pearson, Kendall e Spearman. Dentre eles, é possível destacar o teste paramétrico de correlação de Pearson – capaz de analisar a relação estritamente linear entre duas variáveis contínuas – e o teste não-paramétrico de correlação de Spearman

– que avalia o quanto uma função arbitrária é capaz de descrever uma relação entre duas variáveis, mas sem realizar pressuposições iniciais sobre a distribuição de frequências dessas variáveis (HAUKE; KOSSOWSKI, 2011).

De maneira distinta de um teste de comparação de médias, uma análise correlacional não busca a significância na diferença entre duas amostras, mas no quão semelhante é a taxa de crescimento ou diminuição em ambas. Por causa disso, a realização de um teste de correlação gera como resultado um parâmetro denominado coeficiente de correlação (r para Pearson ou ρ para Spearman) e seu valor pode variar de 1 a -1 .

Um coeficiente exatamente igual a um, por exemplo, $r = 1$, representa uma correlação perfeita positiva (ou seja, sempre que uma das variáveis aumenta, a outra também tende a aumentar) e um $r = -1$ indica uma correlação perfeita negativa (o que significa que o aumento de uma variável sempre levará à diminuição da outra). Caso $r = 0$, é possível concluir que as duas variáveis não dependem linearmente uma da outra, no entanto pode ser que exista uma relação não-linear entre elas e se mostre necessário realizar um teste capaz de analisar esse comportamento.

De modo a permitir uma interpretação adequada de todo o intervalo compreendido pelo resultado dos testes de correlação, Davis (1971) propôs adjetivos capazes de especificar o quão robusto é o valor obtido pelo coeficiente de correlação. Deste modo, descrevendo em valores absolutos, caso, por exemplo, $\rho = 1$, trata-se de uma correlação perfeita; $0,70 < \rho < 0,99$ indica uma correlação muito alta; se $0,50 < \rho < 0,69$, então há uma correlação substancial; $0,30 < \rho < 0,49$ caracteriza uma correlação moderada; $0,10 < \rho < 0,29$ denota uma correlação baixa e $0,01 < \rho < 0,09$ leva ao entendimento de que a correlação obtida é, na verdade, desprezível.

O uso desses adjetivos permite compreender a faixa na qual se encontram os coeficientes de correlação obtidos. No entanto, a obtenção de correlações moderadas ou baixas não significa sempre que a relação entre as variáveis investigadas é reduzida e, por causa disso, quase irrelevante. Muitas vezes, os fenômenos sob estudo possuem níveis de heterogeneidade e complexidade muito elevados, o que faz com que a percepção de sutilezas correlacionais já indiquem direções apropriadas e válidas para a realização de estudos mais aprofundados. Deste modo, o emprego dos adjetivos de Davis deve ser utilizado para se observar o intervalo no qual as relações entre as variáveis se encontram, sendo necessário também analisar o conjunto de dados investigados e o fenômeno em si para se realizar inferências adequadas acerca dos resultados obtidos.

2.2 Trabalhos Relacionados

2.2.1 Estimativas de Tamanhos Amostrais

A determinação de um tamanho ótimo amostral compreende uma temática de pesquisa relevante, pois permite o entendimento da real dimensão da variabilidade de experimentos e a estimação de um tamanho mínimo necessário para se garantir a confiabilidade dos resultados obtidos em investigações que trabalham com amostras heterogêneas (MEIER; LESSMAN, 1971). Na literatura, é possível encontrar pesquisas que determinam essa estimativa em circunstâncias específicas, tais como para melhorar o processo de classificação em atividades de mineração de dados, ajustando-se curvas de aprendizagem para determinar o tamanho de conjuntos de treinamento (LAST, 2009), ou até mesmo estimar assintoticamente um tamanho ótimo que permita a detecção de mudanças na estatística de processos (PELKOWITZ; SCHWARTS, 1987).

Há também trabalhos que abordam a determinação de tamanhos ótimos para o treinamento de Redes Neurais Artificiais (RNAs) (CUI et al., 2004), alocação de recursos para múltiplas atividades concorrentes (EU, 1989) e minimização do custo total de fabricação de circuitos digitais *Ultra Large Scale Integration* - Integração em Escala Ultra Grande (ULSI) (PHAM, 1992). Além disso, em Confalonieri et al. (2007) também é apresentado um software denominado Shortcut In Sample Size Identification (SISSI), que realiza o cálculo do tamanho ótimo amostral por meio da implementação de uma técnica de reamostragem denominada Jackknife – técnica na qual o método Bootstrapping (Subseção 2.1.2, pág. 29) foi baseado.

Ainda na literatura, o cálculo do tamanho ótimo amostral também é realizado para a escolha do número de experimentos a serem realizados para a compreensão de padrões diferenciais de expressão gênica (MÜLLER et al., 2004). No trabalho de Yuan, Lu e Batalden (2016), uma estimativa de tamanho ótimo de amostra é determinada para o projeto de experimentos que precisam trabalhar com amostras que sigam a distribuição exponencial – caracterizada como uma distribuição amplamente utilizada para a análise de risco e confiabilidade em bases de dados.

Na vertente de pesquisa em sistemas *Web*, um tamanho clássico de amostra de cinco usuários foi inicialmente proposto por Nielsen (2000), que justificou tal quantidade como o

suficiente para se compreender cerca de 80% dos problemas de usabilidade em interfaces. No entanto, este tamanho foi readequado com a criação da Regra de 10 ± 2 usuários (HWANG; SALVENDY, 2010) e, em seguida, ambos foram contestados por Schmettow (2012), que demonstrou que a estimativa de tamanhos de amostra não era algo trivial, havendo fatores que causariam a variabilidade dos dados sob análise e sendo importante se atentar a esses detalhes para que os resultados não subestimassem as características investigadas, o que vai de encontro à necessidade de se realizar pesquisas focadas na determinação de tamanhos amostrais, tal como nesta tese.

No escopo educacional, tamanhos de amostra específicos para diferentes tipos de análises foram apresentados por Mertens (2014). Dentre as indicações, recomendou-se a utilização de pelo menos 64 participantes em testes de correlação, 15 observações em testes de regressões simples ou múltiplas e um mínimo de 51 elementos para pesquisas focadas na realização de comparações, tais como analisar o impacto que novas tecnologias causam em cursos de níveis educacionais diversificados. Tais tamanhos foram obtidos por meio de análises do poder de testes estatísticos e os resultados são, na verdade, médias observadas a partir dos valores resultantes nos testes realizados para cada uma destas vertentes (ONWUEGBUZIE; COLLINS, 2007).

Ainda neste escopo, Slavin e Smith (2009) observaram que em revisões sistemáticas da área educacional o tamanho das amostras tende a apresentar uma correlação inversamente proporcional aos efeitos sob investigação. Deste modo, concluíram que os artigos de revisões que abordam quantidades reduzidas de trabalhos sobre determinadas temáticas acabam se mostrando enviesados e relatando efeitos demasiadamente positivos do que outros trabalhos que compreendem quantidades mais abrangentes de conteúdo e de fato representam o universo pesquisado.

Tais trabalhos fortalecem a relevância de se estimar tamanhos ótimos amostrais para ser possível compreender as populações investigadas em níveis satisfatórios. Pesquisas nessa vertente, no entanto, não são frequentes no âmbito da Ciência da Computação e acredita-se que isso ocorra devido à possibilidade de algumas áreas computacionais realizarem simulações de seus estudos em quantidades elevadas para atestarem suas hipóteses. No entanto, pesquisas que envolvem elementos externos, tais como análises que observam características em fatores humanos, apresentam a necessidade de se determinar um tamanho ótimo amostral em função da impossibilidade de se trabalhar frequentemente com quantidades muito elevadas de indivíduos.

Por causa disso, esta tese se destaca dos trabalhos observados na literatura ao demonstrar uma estimativa robusta para se analisar o desempenho acadêmico de estudantes em contato com ferramentas de apoio ao ensino, tais como Ambientes Educacionais Ubíquos (AEUs) ou Ambientes Virtuais de Aprendizagem (AVAs) (Capítulo 4, pág. 55). Tendo em vista que a quantidade de discentes em turmas avançadas de cursos de graduação e pós-graduação é, muitas vezes, reduzida, mostra-se necessário estimar e trabalhar com tamanhos amostrais capazes de compreender a variabilidade das populações de alunos sob análise e permitir que resultados confiáveis sejam alcançados por pesquisadores que realizam investigações nesta vertente científica.

2.2.2 Learning Analytics para a Análise do Impacto de Plataformas de Apoio ao Ensino

A aplicabilidade da vertente de pesquisa em *Learning Analytics* (LA) tem se fortalecido na literatura como uma técnica válida para a realização de estudos acadêmicos, tanto por meio de trabalhos mais extensos e abrangentes (KIENTZ, 2012; SCHMIDT et al., 2014; DAWSON et al., 2017; MOURI; OGATA; UOSAKI, 2017; BODILY et al., 2018; ZHANG et al., 2018), quanto por pesquisas mais específicas que investigam o emprego das ferramentas de apoio ao ensino em contextos mais sucintos (SETTLE; DETTORI; DAVIDSON, 2011; DICKSON et al., 2012; FERREIRA et al., 2012; DURÁN; ÁLVAREZ; ÚNZAGA, 2014; MOURI; OGATA; UOSAKI, 2015; GELAN et al., 2018).

Na literatura, é possível encontrar trabalhos que investigam linhas de raciocínio paralelas, explorando comportamentos de usuários em contato com plataformas educacionais (SMYTH; VOLKER, 2013; KNIGHT et al., 2015; ZHANG et al., 2018) e fatores como o rendimento e a assiduidade de alunos (EUZENT et al., 2011; DICKSON et al., 2012) e as impressões de professores (PRODANOV, 2012; DANIELSON et al., 2014). Dentre os trabalhos com enfoque na análise de discentes, percebe-se que a maioria é voltada para a avaliação de estudantes em cursos de graduação (WIELING; HOFMAN, 2010; DICKSON et al., 2012; SMYTH; VOLKER, 2013; KNIGHT et al., 2015; OLIVARES; HUNDHAUSEN, 2017). Isso é justificado pelo fato das investigações nessa área precisarem considerar as diferentes particularidades dos indivíduos sob análise – como seus níveis de escolaridade e estilos de aprendizagem –, atentando-se para fatores que demonstram as nuances que os métodos educacionais possuem em lugares diferentes.

Em decorrência disso, uma prática frequentemente vista na literatura é a replicação de pesquisas modificando as variáveis investigadas ou os indivíduos em contato com as aplicações. Tais mudanças geram novos resultados devido ao emprego de fatores humanos, mas eventualmente se mostram questionáveis já que repetem falhas metodológicas e carecem de validações estatísticas robustas em seus resultados.

Especificamente na área de LA, não são encontrados trabalhos que envolvem estimativas de tamanhos amostrais adequados para se analisar o desempenho acadêmico de estudantes. A abordagem mais próxima disso é apresentada em Mertens (2014), que faz indicações de tamanhos diversos para a realização de testes estatísticos, no entanto, o trabalho foca apenas nos meios de testagem e não nas variáveis sob investigação. Além desse enfoque, não foram encontradas pesquisas que fazem cálculos acerca da quantidades de indivíduos sob análise para atestarem se seus resultados representam com fidelidade as populações das quais as amostras são retiradas. Tal carência prejudica a confiabilidade dos experimentos realizados e dos resultados obtidos, visto que os ambientes acadêmicos possuem características de heterogeneidade intrínsecas à área educacional.

Para esta tese, determinou-se uma estimativa de tamanho ótimo amostral para ser possível compreender com maior precisão o comportamento da variabilidade que o desempenho dos alunos está sujeito a apresentar. A partir desta estimativa, analisou-se o desempenho de discentes de graduação e pós-graduação com o intuito de se observar a existência de faixas de alunos que apresentam maior receptividade à inserção de novas plataformas de apoio ao ensino e, deste modo, verificar quais usuários são mais influenciados pelo seu uso no cotidiano.

Ainda acerca de trabalhos presentes na literatura, o questionário se destaca como a técnica de avaliação mais comumente empregada (DICKSON et al., 2012; SMYTH; VOLKER, 2013; DANIELSON et al., 2014; MOURI; OGATA; UOSAKI, 2015; COOPER; FERGUSON; WOLFF, 2016; MOURI; OGATA; UOSAKI, 2017). Nas investigações observadas, os usuários que têm contato com plataformas de apoio ao ensino normalmente informam suas impressões sobre as tecnologias, pontuando características positivas e negativas que foram encontradas durante o período em que fizeram uso das ferramentas, e dão sugestões de novas funcionalidades para engrandecer os sistemas. Todavia, nos trabalhos em que a análise de dados se baseia unicamente nesse tipo de resposta, é possível perceber que não são utilizados métodos de validação regulares, o que inibe a correta verificação das informações e eventualmente gera conclusões pretensiosas e refutáveis.

Em Settle, Dettori e Davidson (2011), por exemplo, fez-se o uso de questionários para avaliar as impressões de alunos, mas não foram aplicadas técnicas para legitimar essas respostas. Ao se analisar o respaldo dado pelos usuários, percebeu-se que a maioria demonstrou satisfação em acessar o sistema e que haviam compreendido melhor o conteúdo – o que ocasionaria rendimentos melhores. Em seguida, com o intuito de relacionar os resultados, empregaram-se técnicas estatísticas para analisar se os usuários que interagiram com o AEU realmente obtiveram notas melhores que os demais, mas ao final da etapa descobriu-se que não houve diferença significativa entre as turmas observadas. O trabalho então concluiu-se de maneira confusa, explicando que não foi possível afirmar que a tecnologia resultava em benefícios pois os dados se mostraram contraditórios. Todavia, percebe-se que a investigação falhou no método de pesquisa empregado, já que apenas parte das informações foi corretamente validada.

Em Mouri, Ogata e Uosaki (2017), é apresentado um software intitulado Visualization and Analysis System for COnnecting Relationships of Learning Logs (VASCORLL). Desenvolvido com o intuito específico de auxiliar na visualização e análise de *logs* de aprendizagem que são coletados em sistemas de *seamless learning* (aprendizagem contínua), o mesmo possui a intenção de preencher a lacuna atualmente existente entre os processos de aprendizagem formal e informal.

Embora o trabalho possua uma proposta de caráter relevante e o processo de implementação seja explicado em detalhes, o modo de avaliação do sistema mostrou-se carente, pois baseou-se em um questionário compreendendo perguntas sobre o sistema e essas foram respondidas em Escalas Likert de 5 pontos nas quais os usuários informaram seus níveis de satisfação. Para as análises, calcularam-se médias e desvios padrões das respostas dadas aos questionários, empregando-se os resultados obtidos com o intuito de legitimar a relevância do sistema. No entanto, além de não ter sido utilizada qualquer métrica de validação de questionários, o trabalho empregou 20 indivíduos de características muito heterogêneas, sendo de idades distintas que variaram de 21 a 36 anos, além de níveis de instrução e até mesmo nacionalidades diferentes.

No trabalho de Euzent et al. (2011) examinou-se o rendimento de diferentes turmas que assistiram à mesma disciplina, atentando-se para o emprego do mesmo professor e método de avaliação no decorrer das observações. Numa das turmas analisadas, utilizou-se uma nova abordagem de captura de conteúdo em sala de aula no semestre letivo, enquanto na outra as aulas foram lecionadas de maneira expositiva clássica.

O intuito do trabalho foi compreender se a gravação dessas aulas causava algum tipo de dano à maneira com que o conteúdo era repassado aos alunos e se a aprendizagem seria prejudicada de algum modo. Ao final da pesquisa, mesmo com os dados sendo corretamente legitimados, não se observou qualquer diferença expressiva entre as notas finais das turmas, mas percebeu-se que a faixa de rendimento do segundo bimestre foi maior na turma que utilizou o sistema, o que permitiu a conclusão de não ter havido danos ou prejuízos para os alunos, mas tampouco ganhos significativos.

Jeremic, Kumar e Graf (2017) apresentaram uma plataforma denominada MORPH, desenvolvida para facilitar a integração de módulos e ferramentas de LA em sistemas institucionais de aprendizagem. A ferramenta fornece uma arquitetura distribuída capaz de combinar processamento de dados em *batch* (lote), *stream* (fluxo) e tempo real, empregando um modelo de paralelismo para habilitar e permitir o processamento de grandes quantidades de dados. Tal arquitetura foi construída para se conectar módulos de pesquisa aos *Learning Management Systems* - Sistemas de Gestão da Aprendizagem (LMSs), fornecendo suporte para questões gerenciais dos sistemas e permitindo que pesquisadores se concentrem mais em suas análises em vez de questões de engenharia de software.

Embora essa proposta se assemelhe em partes ao módulo de gerenciamento de experimentos apresentado nesta tese, tal ferramenta não se mostra capaz de controlar grupos distintos de estudantes matriculados numa mesma turma, de modo a permitir que as subamostras tenham níveis diferentes de visibilidade para as mesmas funcionalidades. O Gerenciador de Experimentos (Capítulo 3, pág. 45) permite a realização deste tipo de controle, o que possibilita, por exemplo, a realização de análises comparativas mais confiáveis, pois os elementos investigados encontram-se em escopos experimentais idênticos e não possuem efeitos residuais que podem gerar resultados enviesados.

Além disso, nesta tese os dados de desempenho discente foram empregados para a realização de dois estudos: uma Análise de Variância Fatorial 4x5 (*Two-way ANOVA*) (Capítulo 5, pág. 65) e uma análise empregando a Correlação de Spearman (Capítulo 6, pág. 77). A ANOVA Fatorial foi utilizada para se averiguar a existência de diferenças significativas no desempenho acadêmico de estudantes em contato cotidiano com uma plataforma educacional à medida que a mesma foi aprimorada progressivamente e novas funcionalidades foram acrescentadas à sua estrutura.

Já a Correlação de Spearman foi empregada para se averiguar a relação entre o desempenho acadêmico dos estudantes e seus fluxos de interação com a plataforma CX, analisando a quantidade de interações realizadas com os módulos que foram progressivamente acrescentados ao sistema. Tais análises empregaram dados (notas) de alunos de cursos de graduação em Sistemas de Informação e graduação e pós-graduação em Ciência da Computação da Faculdade de Computação da Universidade Federal de Uberlândia (FACOM/UFU), com o intuito de se obter uma perspectiva realista da influência de ferramentas de apoio ao ensino em ambientes acadêmicos reais.

Por meio desta pesquisa, acredita-se que serão abertas margens para a realização de novos trabalhos, dado o fato que o âmbito acadêmico mostra-se como uma área em franco crescimento (AIHUA, 2010; KINSHUK; GRAF, 2012; KORTEMEYER, 2017). Por isso, investigações que avaliam o uso de tecnologias educacionais permitem explorar tendências não apenas para o futuro científico deste campo, mas também em demais vertentes da área em LA, tendo em vista que uma de suas premissas originais diz respeito à sua abrangência multidisciplinar (WILDE; BRUEGGER; HIRSBRUNNER, 2010; COOK; DAS, 2012; SMYTH; VOLKER, 2013; COOPER; FERGUSON; WOLFF, 2016; MOURI; OGATA; UOSAKI, 2017; GELAN et al., 2018; ZHANG et al., 2018).

2.3 Classroom eXperience

O emprego de dispositivos computacionais para se realizar a instrumentação de salas de aula permite a automatização de processos de ensino e aprendizagem, gerando benefícios para alunos e professores por meio da criação de Ambientes Educacionais Ubíquos (AEUs) (AIHUA, 2010; ZHAO; WAN; OKAMOTO, 2010; KINSHUK; GRAF, 2012; DURÁN; ÁLVAREZ; ÚNZAGA, 2014). Um AEU, deste modo, pode ser criado a partir da integração de diversos componentes de hardware, tais como microfones, câmeras, projetores multimídia e lousas digitais, de maneira a possibilitar a captura automática dos conteúdos abordados em sala de aula e permitir seu acesso posterior pelos usuários (TRUONG; HAYES, 2009).

O Classroom eXperience (CX)³ é um AEU desenvolvido de modo conjunto pelo grupo de pesquisa UbiMedia⁴ da Faculdade de Computação da Universidade Federal de Uber-

³ <http://cx.facom.ufu.br/>

⁴ <http://ubimedia.facom.ufu.br>

lândia (FACOM/UFU). Possui como intuito permitir o registro de fluxos de conteúdo multimídia apresentados em sala de aula e sua disponibilização posterior para professores e estudantes (FERREIRA et al., 2012; ARAÚJO et al., 2013). Por meio de sua utilização, é possível capturar, armazenar, acessar e enriquecer as mídias geradas em ambientes educacionais instrumentados e contribuir com as atividades de ensino/aprendizagem de usuários que utilizam esta tecnologia no cotidiano acadêmico (BRANT-RIBEIRO; BIASE; CATTELAN, 2015; BRANT-RIBEIRO et al., 2015).

O processo de captura do CX se baseia nas quatro fases essenciais de Captura e Acesso (C&A): 1) *pré-produção* (momento em que as aulas são preparadas), 2) *gravação ao vivo* (quando as atividades são registradas no momento em que acontecem), 3) *pós-produção* (fase em que os fluxos produzidos são armazenados) e 4) *acesso* (quando as aulas capturadas são disponibilizadas para alunos e professores) (ABOWD; MYNATT, 2000; TRUONG; HAYES, 2009).

Além destas fases, uma quinta etapa intitulada *extensão* também pode ser destacada das supracitadas (PIMENTEL et al., 2001). Tal etapa compreende o momento em que os usuários enriquecem os conteúdos capturados por meio de atividades interativas sociais e colaborativas com a plataforma educacional. Para que esses processos ocorram de maneira transparente, o CX atualmente conta com duas salas de aula equipadas com dispositivos computacionais (como microfones, lousas digitais, câmeras e projetores multimídia), localizadas na FACOM/UFU, conforme pode ser visto na Figura 2.



Figura 2 – Salas de aula instrumentadas com dispositivos computacionais, tais como: lousas digitais, microfones, câmeras e projetores multimídia (Fonte: Brant-Ribeiro, Biase e Cattelan (2015)).

Demandando modificações mínimas na maneira com que as aulas convencionais são lecionadas, o CX emprega recursos da Computação Ubíqua (UbiComp) para auxiliar tanto professores, enquanto realizam a captura das aulas, quanto alunos, durante atividades de acesso ao conteúdo capturado. Para isso, o sistema conta com componentes de hardware e software especializados para realizar a gravação das atividades educacionais, sincronizar os fluxos de mídia capturados e disponibilizar este conteúdo por meio de uma interface de acesso *Web*, cujas telas podem ser observadas nas Figuras 3 e 4.

(a)

The login page features the Classroom eXperience logo at the top center. Below it, a white box contains the text "Informe suas credenciais:". There are two input fields: "Usuário" with a person icon and "Senha" with a lock icon. A black "Entrar" button is positioned below the fields. At the bottom of the box, there are two links: "Novo usuário?" and "Esqueceu sua senha?".

(b)

The context selection page shows the Classroom eXperience logo at the top left. Below it, a section titled "Contexto do usuário" contains the text "Olá, Taffarell! Selecione seu contexto de acesso:". There are three dropdown menus: "Tempo disponível:" with "Até 15 Minutos" selected, "Local de acesso:" with "Trabalho" selected, and "Motivo:" with "Aula perdida" selected. Below these are two green buttons: "Ignorar" and "Salvar". At the bottom, there is a footer with "© Copyright 2012-2017 UbiMedia - Todos os direitos reservados", a "Português" language selector, and a "Sobre" link.

Figura 3 – Interface *Web* do Classroom eXperience (CX): (a) página de *login* e (b) tela de seleção de contexto (Fonte: Adaptado de Brant-Ribeiro e Cattelan (2015)).

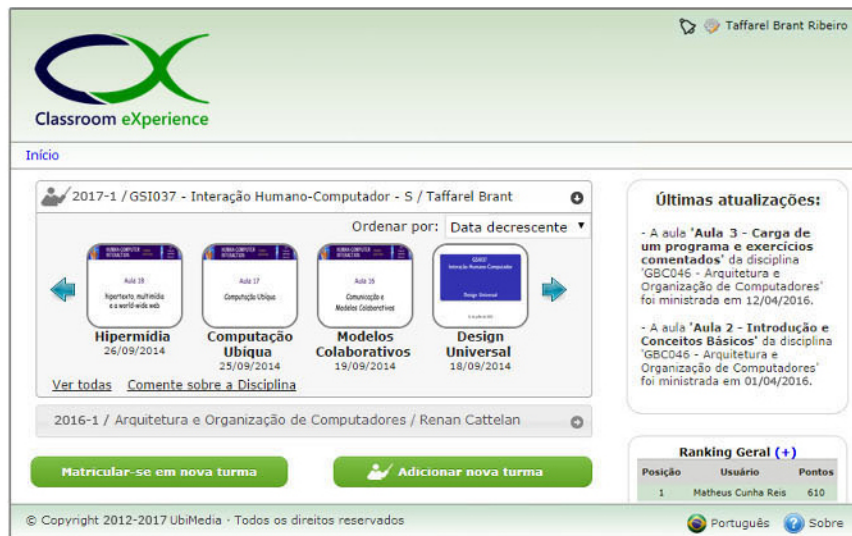
A página inicial do CX (Figura 3(a)) é uma tela de *login* que realiza a validação do nome de acesso e senha dos usuários que desejam acessar o sistema. Além disso, ela possui um *link* que direciona para a tela de cadastro de novos usuários, na qual alunos podem registrar-se livremente e professores devem informar um código de segurança para que consigam se cadastrar. Caso os usuários esqueçam suas credenciais de acesso, o sistema também dispõe de uma área para a recuperação de senha a partir da inserção de um e-mail já cadastrado na base de dados. Após fazerem *login* no CX, os usuários são direcionados para uma tela que contém um mini formulário de contexto de acesso (Figura 3(b)). Esses dados fazem parte de um conjunto padronizado de dimensões de contexto que tem como objetivo coletar informações dos usuários em tempo de execução, com o intuito de proporcionar uma melhor experiência enquanto realizam acessos e interagem com a plataforma educacional (ARAÚJO et al., 2013).

Atualmente, o CX utiliza sete dimensões de contexto: a) tipo de dispositivo, b) velocidade de conexão, c) resolução da tela, d) data e hora de acesso, e) tempo disponível, f) local de origem do acesso e g) motivo do usuário acessar o sistema. Dessas, as quatro primeiras são aferidas de maneira automática pelo sistema, enquanto as três últimas precisam ser informadas manualmente por meio do mini formulário de contexto presente na Figura 3(b). Com os dados obtidos a partir das dimensões de contexto, torna-se possível fornecer benefícios aos usuários, tais como a personalização e recomendação de conteúdos capturados e a restrição de determinadas mídias de acordo com os dispositivos empregados pelos usuários para acessar o sistema.

Após informarem as dimensões contextuais, os usuários são direcionados para a *home page* da plataforma (Figura 4(a)), na qual é possível, para um professor, cadastrar uma nova turma e, para os alunos, matricularem-se nas turmas já cadastradas no CX. Além disso, esta página também exibe as últimas atualizações que aconteceram no sistema, tais como aulas capturadas recentemente e avaliações registradas pelos professores – como provas, trabalhos, exercícios e seminários. De modo a melhorar a organização desses dados, eles são exibidos empregando-se filtros temporais e de disciplinas vinculadas aos usuários do CX, de maneira que apenas alunos matriculados em turmas nas quais essas atualizações acontecem são capazes de enxergar tais informações na *home page*.

Ao selecionarem uma aula específica, os usuários devem escolher em qual formato de apresentação o conteúdo capturado deverá ser exibido. No CX, atualmente há três tipos de formato de apresentação possíveis: somente *slides*, somente vídeo e completo. Caso

(a)



(b)

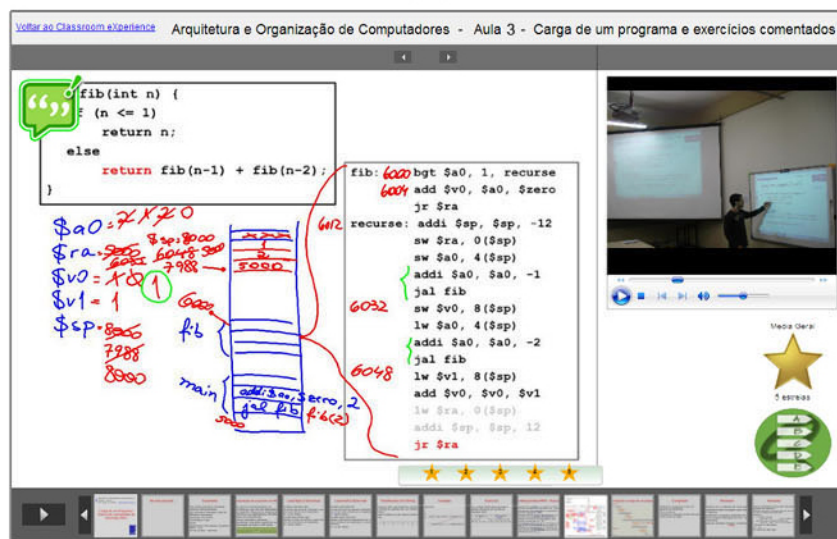


Figura 4 – Interface Web do CX: (a) *home page* da plataforma e (b) tela de aula capturada (Fonte: Brant-Ribeiro e Cattelan (2015)).

selecione a primeira opção, o usuário será direcionado para uma página contendo os *slides* utilizados pelo professor durante uma sessão de captura, já com as anotações realizadas nessas mídias. A segunda opção apresentará apenas o vídeo de uma aula capturada pelo sistema e a terceira exibirá ambos os fluxos de mídia, conforme a Figura 4(b).

Gerenciador de Experimentos

Neste capítulo será apresentado o Gerenciador de Experimentos que foi desenvolvido no decorrer desta pesquisa e integrado à plataforma de apoio ao ensino Classroom eXperience (CX) (Seção 2.3, pág. 40). Para esta finalidade, inicialmente é realizada uma visão geral da proposta na Seção 3.1; em seguida, o processo de integração do Gerenciador de Experimentos é explicado na Seção 3.2 e, por fim, a interface *web* do Gerenciador é apresentada na Seção 3.3.

3.1 Visão Geral da Proposta

A realização de um planejamento experimental adequado mostra-se como uma etapa essencial para a execução de pesquisas científicas. A isso, soma-se também a necessidade de se manter um controle de todos os tratamentos amostrais sob análise, de modo a ser possível manter a organização dos experimentos e se evitar erros durante os processos de coleta e avaliação dos dados. Deste modo, devido ao escopo de investigação desta pesquisa compreender muitas amostras com níveis distintos de tratamentos, enxergou-se a necessidade de se desenvolver um módulo de gerenciamento capaz de estruturar as etapas de análise sem a ocorrência de desorganizações ou equívocos que levariam à perda de confiabilidade ou robustez nos resultados obtidos.

Com isso em mente, foi desenvolvido um modelo de Gerenciador de Experimentos com características generalistas e capaz de ser empregado para realizar o controle da visibilidade de módulos e funcionalidades em plataformas de apoio ao ensino, tais como Ambientes Educacionais Ubíquos (AEUs) ou Ambientes Virtuais de Aprendizagem (AVAs).

O modelo aqui proposto compreendeu um mecanismo que realiza o controle de quais funcionalidades devem ser exibidas para usuários, de modo a ser possível habilitá-las em tratamentos diferentes e para grupos distintos de indivíduos simultaneamente. Para esta finalidade, considerou-se um módulo de monitoramento capaz de receber dados advindos de sistemas educacionais e empregá-los de modo a criar estruturas de controle para se definir a visibilidade de funcionalidades (Figura 5).

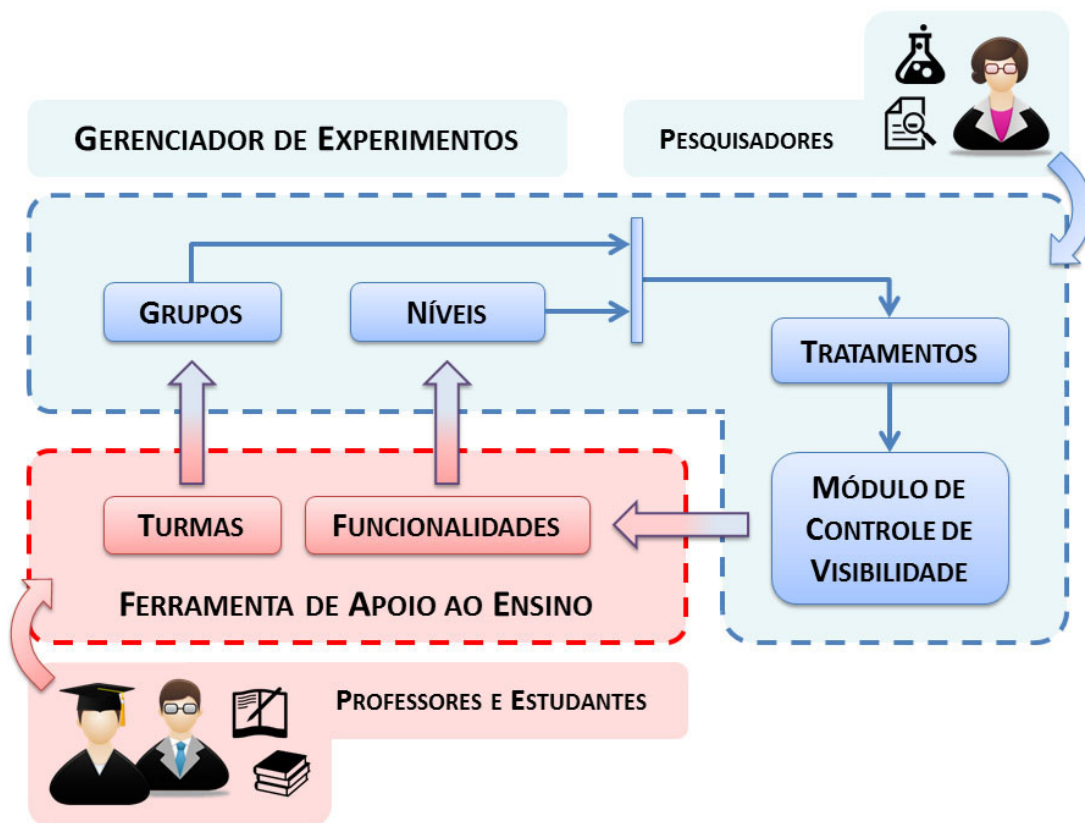


Figura 5 – Estrutura do Gerenciador de Experimentos e interface de comunicação com ferramentas de apoio ao ensino (Fonte: Brant-Ribeiro e Cattelan (2018)).

O Gerenciador de Experimentos (realçado em azul na Figura 5) compreende um módulo de organização de dados provenientes de ferramentas acadêmicas. No âmbito de plataformas educacionais, as turmas que possuem estudantes matriculados em disciplinas cadastradas nessas ferramentas são caracterizadas como grupos (ou amostras) experimentais a serem utilizados pelo gerenciador. Além disso, cada funcionalidade presente na plataforma fornece dados referentes aos níveis de acesso a essas funcionalidades, visto que as mesmas podem ser habilitadas ou desabilitadas para cada usuário ou, também, para conjuntos de usuários que fazem parte de uma mesma turma.

Empregando-se os dados dos grupos e dos níveis de acesso às funcionalidades, os pesquisadores podem montar tratamentos, que são organizações dos níveis de visibilidade das funcionalidades que serão habilitadas para cada grupo de estudantes. Esses tratamentos são organizados utilizando-se um módulo de controle de visibilidade das funcionalidades nas plataformas, capaz de regular e restringir o que cada usuário poderá acessar em cada uma das turmas que fizer parte.

Por se tratar de um elemento capaz de realizar o controle de funcionalidades e módulos que compõem as ferramentas de apoio ao ensino, apenas pesquisadores devem ter acesso ao módulo de gerenciamento. Por causa disso, o mesmo mostra-se invisível para professores e alunos que empregam tais ferramentas no decorrer dos semestres letivos.

3.2 Integração da Proposta ao CX

Com o intuito de se automatizar o planejamento experimental desta pesquisa, o Gerenciador de Experimentos proposto foi integrado à plataforma de apoio ao ensino CX. Tal módulo foi desenvolvido utilizando-se o *Integrated Development Environment* - Ambiente de Desenvolvimento Integrado (IDE) Eclipse¹ e empregando-se a biblioteca de funções JavaScript jQuery² e a tecnologia JavaServer Pages (JSP)³. De modo a ilustrar como realizou-se a integração do Gerenciador de Experimentos ao sistema CX, a Figura 6 apresenta um Diagrama de Componentes que possui as interfaces de acesso ao sistema e a dependência entre seus componentes.

O componente do Gerenciador de Experimentos (realçado em azul na Figura 6) é um elemento responsável pelo controle da visibilidade dos módulos e funcionalidades existentes no CX, tais como os quizzes, comentários e ranqueamento de *slides*. Para isso, ele possui a interface “Visibilidade” que contém três serviços:

- **GET seleçãoTurma**: serviço que requer os parâmetros “semestre letivo”, “nome da disciplina” e “professor responsável” e retorna todos os estudantes que estão matriculados naquela oferta da disciplina;

¹ <http://eclipse.org/ide>

² <http://jquery.com/>

³ <http://docs.oracle.com/javaee/5/tutorial/doc/bnagy.html/>

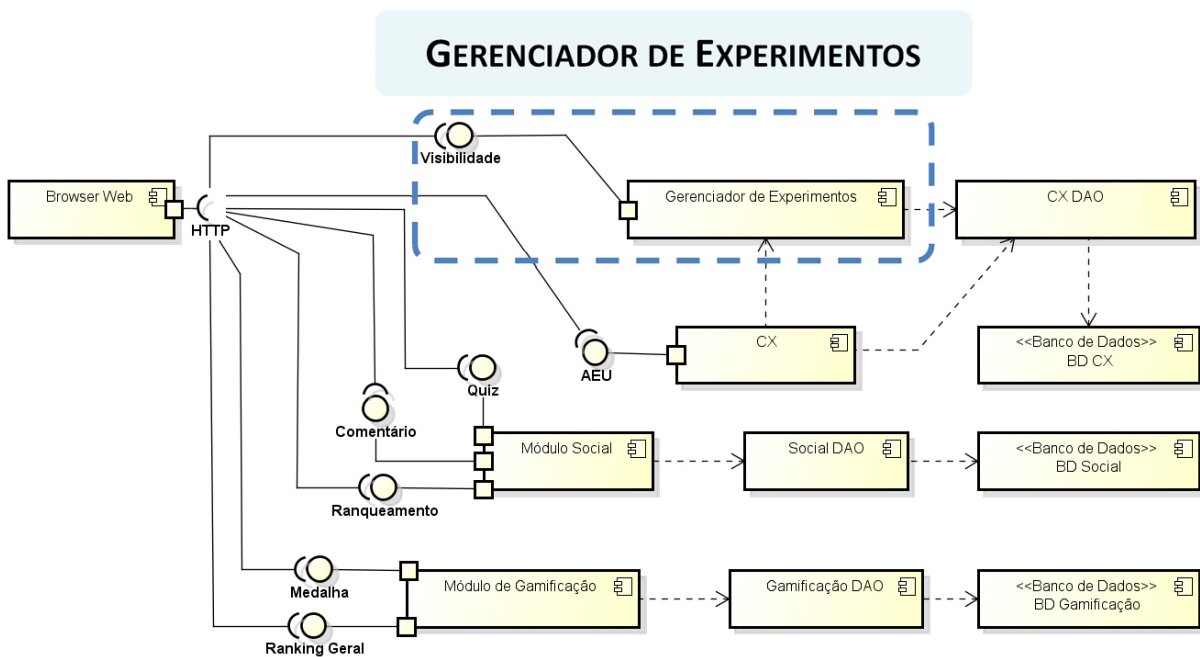


Figura 6 – Diagrama de Componentes com destaque para o componente do Gerenciador de Experimentos.

- ❑ **GET seleçãoVisibilidade:** serviço que requer o parâmetro “código de matrícula” de todos os estudantes matriculados numa disciplina e retorna o nível de visibilidade que cada discente possui das funcionalidades do CX;
- ❑ **POST atualizaVisibilidade:** serviço que requer o parâmetro “código de matrícula” de todos os estudantes matriculados numa disciplina e atualiza o nível de visibilidade que cada discente possui das funcionalidades do CX.

Para isso, foi montada uma base de dados própria para o Gerenciador de Experimentos utilizando-se o Sistema de Gerenciamento de Banco de Dados (SGBD) PostgreSQL⁴. Para esta base de dados, foi criada uma tabela denominada *Visibility* contendo dois campos: *IdEnrolled* (referente ao parâmetro “código de matrícula”) e *VisibilityLevel* (referente ao nível de visibilidade das funcionalidades do CX). A Tabela 1 apresenta alguns alunos matriculados em turmas e seus respectivos níveis de visibilidade.

Definiu-se o campo inteiro *IdEnrolled* como a chave primária da tabela *Visibility*, de modo a ele conter os códigos das matrículas de cada aluno nas disciplinas cadastradas no CX. Como um mesmo usuário pode ser discente em diversas turmas simultaneamente

⁴ <http://www.postgresql.org/>

Tabela 1 – Tabela *Visibility* da base de dados do Gerenciador de Experimentos com algumas tuplas contendo dados.

IdEnrolled	VisibilityLevel
150	SQG
151	SQG
285	S
286	S
287	S
308	QG
309	QG
310	QG

(e em cada turma o estudante pode estar em grupos de tratamentos distintos), optou-se por utilizar os códigos de matrícula em vez dos identificadores únicos dos alunos, o que também caracterizou o campo *IdEnrolled* como uma chave estrangeira que referenciou a tabela *Enrolled*, presente na base de dados do CX e que, atualmente, realiza a organização das matrículas que os estudantes possuem nas disciplinas. Essa estratégia pode ser adaptada para outras ferramentas educacionais, visto que normalmente são empregadas tabelas semelhantes, com chaves simples ou compostas, para controlarem as matrículas dos estudantes nos cursos cadastrados nessas plataformas.

O campo de caracteres variáveis *VisibilityLevel* fez o controle de quais funcionalidades devem ser visíveis para os discentes de cada turma. Deste modo, o caractere “S” representou o Módulo Social (MS), o caractere “Q” indicou os Quizzes presentes nas aulas e o caractere “G” expressou a funcionalidade de Gamificação do CX. Exemplificando o modo que esses dados foram utilizados no CX, caso um estudante possuísse apenas o caractere “S” para o controle de seu nível de visibilidade (p. ex., usuários 285, 286 e 287), apenas a funcionalidade de Socialização do CX seria apresentada. Por outro lado, alunos que foram atribuídos com os caracteres “SQG” (p. ex., usuários 150 e 151), tiveram acesso a todos os módulos presentes na ferramenta.

Com isso, foi possível atribuir níveis de visibilidades diferentes aos mesmos estudantes nos casos em que eles estiveram matriculados em disciplinas distintas durante os semes-

tres letivos. Essa característica também é passível de extensão e integração a diversas plataformas de apoio ao ensino, tais como AEU's ou AVAs, com o intuito de possibilitar o controle da visibilidade de módulos e funcionalidades nessas ferramentas.

De modo a permitir a atribuição dos níveis de visibilidade para os alunos matriculados nas disciplinas do CX, foram desenvolvidas páginas *Web* capazes de realizar consultas e inserções/atualizações na tabela *Visibility* da base de dados. Isso permitiu a criação de estruturas condicionais nas páginas do sistema que, assim que carregadas, realizam buscas no banco de dados e exibem cada uma das funcionalidades caso os caracteres supracitados se encontrem nas tuplas referentes ao *IdEnrolled* dos estudantes.

3.3 Interface *Web* do Gerenciador de Experimentos

Com o intuito de permitir a utilização do Gerenciador de Experimentos junto ao CX, desenvolveu-se uma interface *Web* para os pesquisadores que empregam esta ferramenta para a realização de experimentos. De modo a simular um fluxo comum de utilização do módulo de gerenciamento por um pesquisador, é possível acessá-lo a partir da *home page* do CX por meio de um *link* presente no menu suspenso do usuário (Figura 7(a)).

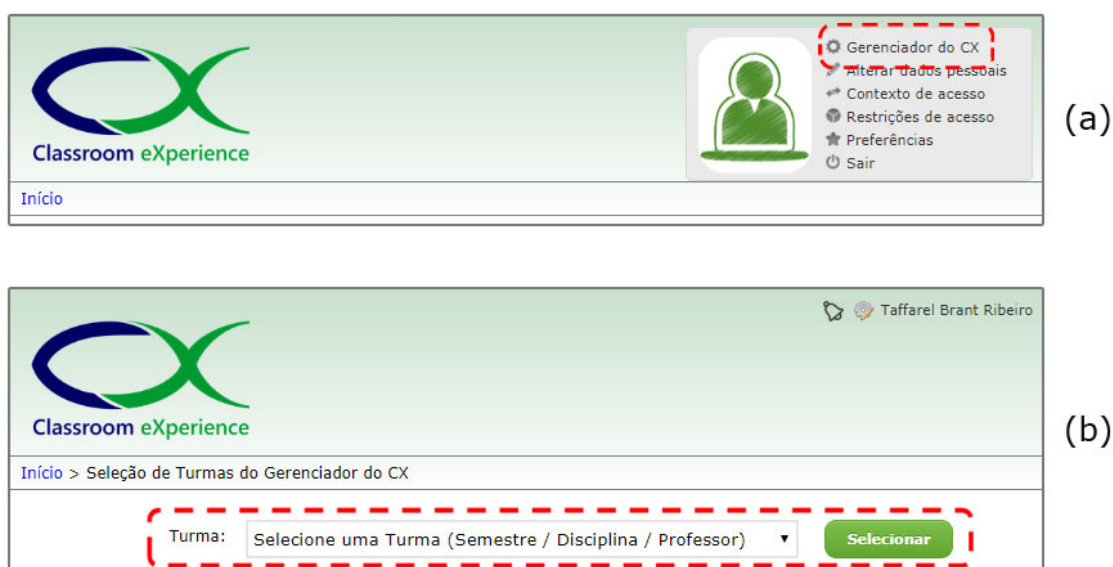


Figura 7 – Acesso ao Gerenciador de Experimentos do CX: (a) *link* no menu suspenso da *home page* de usuários pesquisadores e (b) área de seleção de turmas cadastradas no sistema.

Ao clicar no *link*, o pesquisador é direcionado para uma página na qual é possível realizar a seleção de uma turma específica para se controlar a visibilidade das funcionalidades do CX que serão exibidas para os estudantes matriculados (Figura 7(b)). Por terem havido circunstâncias de ofertas semestrais consecutivas das mesmas disciplinas (diferenciando-se, em sua maioria, apenas pelos estudantes matriculados), optou-se pela apresentação das turmas explicitando o semestre letivo, o nome da disciplina e o professor responsável. Deste modo, foi possível selecionar apenas as turmas que realmente estiveram sob investigação sem que ocorressem confusões durante o planejamento das análises semestrais.

Após selecionar uma turma específica, o pesquisador acessa uma página na qual são listados todos os estudantes matriculados naquela oferta da disciplina (Figura 8). Neste momento, são apresentadas caixas de seleção referentes às funcionalidades existentes no sistema para cada um dos alunos vinculados, o que permite controlar a visibilidade dos módulos e elementos já implementados no CX. Na Figura 8, é possível observar uma turma de quinze alunos cursando a disciplina Lógica Matemática (LM), na qual optou-se por habilitar as funcionalidades de Socialização e Quizzes, mas decidiu-se manter a Gamificação invisível para os estudantes.

As funcionalidades listadas na Figura 8 tiveram seus níveis de visibilidade controlados por meio de estruturas condicionais inseridas nas páginas do CX. Deste modo, se o usuário fosse, por exemplo, um estudante matriculado na disciplina LM e os pesquisadores optassem por realizar uma análise na qual este indivíduo não pudesse ter acesso a qualquer funcionalidade do sistema, no momento em que o aluno fosse acessar uma aula que foi capturada empregando-se o CX, ele visualizaria apenas o conteúdo daquela aula, sem outros módulos visíveis para o enriquecimento do material didático (Figura 9(a)).

Caso o indivíduo sob análise fosse um estudante que precisasse ter acesso às funcionalidades do CX, ao acessar as aulas da disciplina ele visualizaria os módulos habilitados pelo Gerenciador de Experimentos junto ao conteúdo capturado pelo sistema. Tais funcionalidades foram realçadas na Figura 9(b) e referem-se a um botão para a inserção de comentários nos *slides*, uma área de ranqueamento de *slides* empregando estrelas, uma região com um *feedback* da média de estrelas para cada *slide* e uma opção para responder aos Quizzes que foram previamente cadastrados pelos professores.

Deste modo, tornou-se possível organizar o nível de visibilidade das funcionalidades presentes no CX e realizar a automatização do planejamento semestral necessário para esta

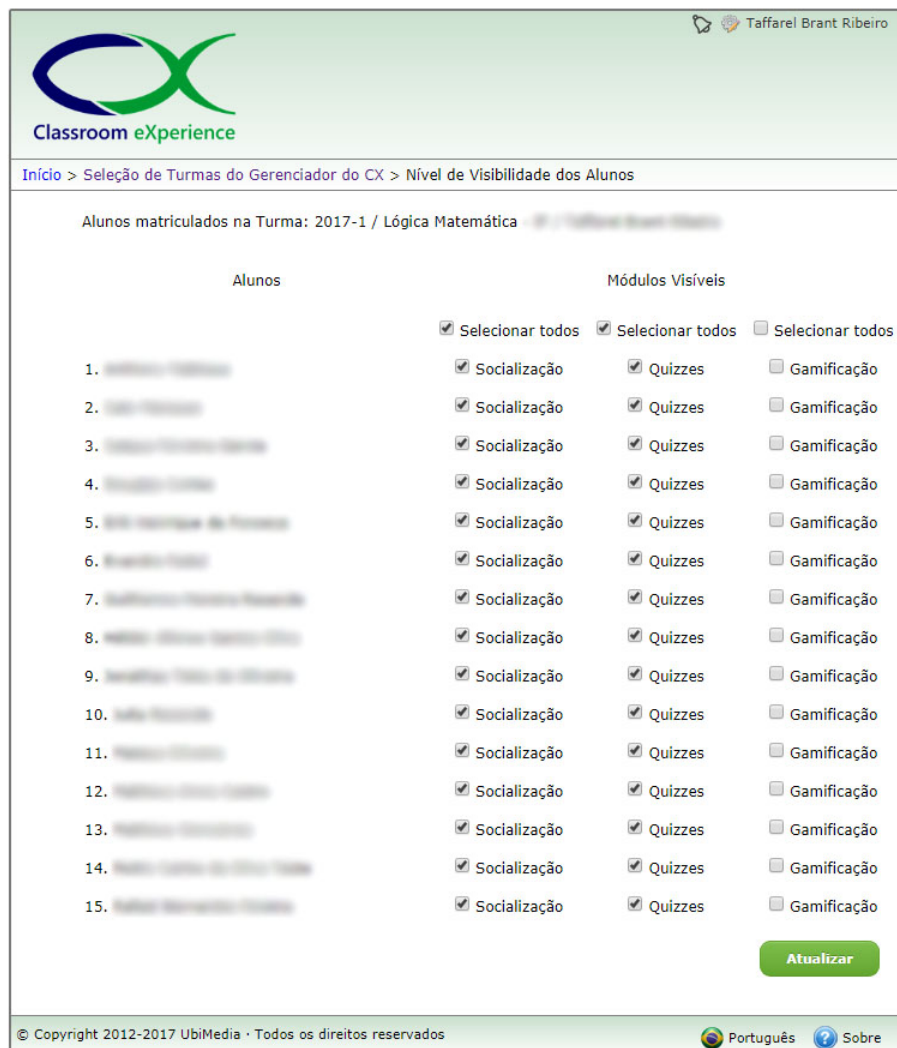


Figura 8 – Interface do Gerenciador de Experimentos do CX após a seleção de uma turma com estudantes matriculados.

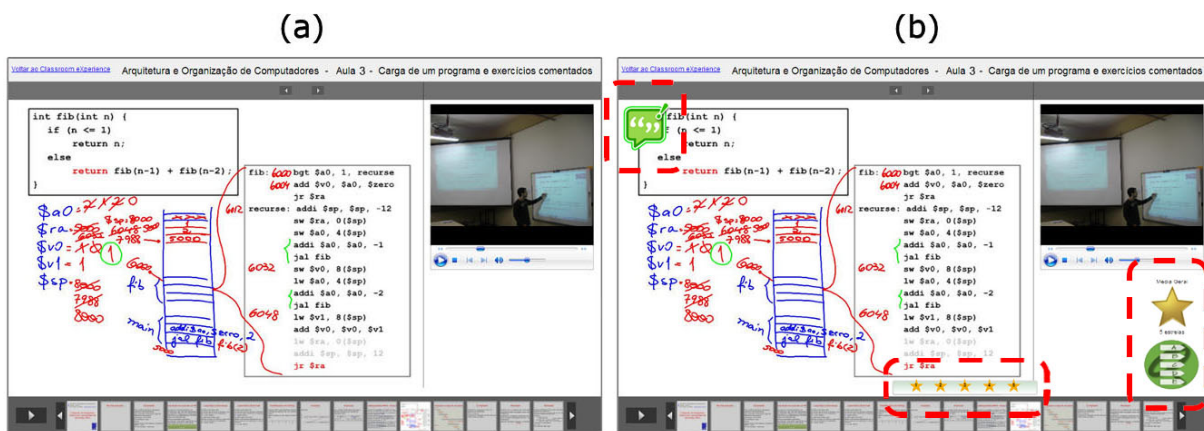


Figura 9 – Telas de aulas capturadas empregando o CX: (a) aula com os módulos desabilitados para os estudantes e (b) aula com as funcionalidades de Socialização e Quizzes habilitadas.

pesquisa. O Gerenciador de Experimentos mostrou-se capaz, inclusive, de permitir que grupos distintos de estudantes matriculados numa mesma turma pudessem ter diversos níveis de visibilidade das funcionalidades das plataformas. Isso abre margem para a realização de testes pareados (em que os mesmos indivíduos são observados em vários momentos durante pesquisa) e análises comparativas confiáveis, visto que os elementos analisados se encontram integralmente num mesmo escopo experimental.

Estimativa de Tamanho Ótimo Amostrai para a Análise do Desempenho Discente

Neste capítulo serão apresentados os resultados obtidos a partir do cálculo de tamanhos ótimos amostrais para a análise do desempenho acadêmico de estudantes em cursos de graduação e pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia (FACOM/UFU). Tais estimativas foram calculadas tanto para turmas que não utilizaram tecnologias educacionais no cotidiano quanto para alunos em contato com ferramentas acadêmicas. Deste modo, inicialmente tanto o método de pesquisa quanto os materiais empregados são descritos na Seção 4.1. A Seção 4.2 apresenta os resultados obtidos e discussões e, por fim, considerações finais são pontuadas na Seção 4.3.

4.1 Material e Métodos

Para realizar esta análise¹, foram utilizados dados provenientes de dez turmas que cursaram disciplinas da grade curricular dos cursos de graduação e pós-graduação da FACOM/UFU, totalizando 365 observações que foram coletadas no decorrer de três anos letivos (de 2013 a 2015). Dessas, cinco não tiveram qualquer contato com a plataforma de apoio ao ensino Classroom eXperience (CX) (Seção 2.3, pág. 40), ao passo que as demais utilizaram o Ambiente Educacional Ubíquo (AEU) no decorrer dos semestres le-

¹ Pesquisa aprovada pelo Comitê de Ética da UFU (CAAE 46909515.4.0000.5152).

tivos. Objetivando compreender estudantes tanto do curso de graduação em Sistemas de Informação quanto dos cursos de graduação e pós-graduação em Ciência da Computação, foram selecionadas disciplinas que todos os cursos possuem em suas respectivas grades curriculares, de modo a propositalmente haver amostras mistas com quantidades diversificadas de estudantes de cada curso.

As dez amostras selecionadas foram compostas por N observações da variável sob análise. Tendo em vista que cada amostra compreendeu uma turma de alunos matriculados numa respectiva disciplina, o tamanho das amostras não se manteve fixo e apresentou variações entre 26 e 44 estudantes. Como a variável estudada nesta pesquisa foi o desempenho acadêmico dos alunos, as notas dos estudantes foram coletadas ao final de cada semestre letivo e, para cada uma das amostras resultantes, foram realizadas combinações aleatorizadas (reamostragens) das observações para formar subamostras de diversos tamanhos a partir da amostra original.

Para a realização do processo de reamostragem, o método Bootstrapping (Subseção 2.1.2, pág. 29) foi empregado nesta fase (EFRON, 1979). Deste modo, para cada subamostra de tamanho n originada de uma amostra de tamanho N , foram realizadas 2000 reamostragens para todo $n \leq N$. Em seguida, para cada um dos conjuntos de 2000 subamostras de mesmo tamanho resultantes, realizou-se o cálculo da Média (\bar{x}), Variância (s^2), Desvio Padrão (s), Coeficiente de Variação (CV), além de um Intervalo de Confiança (IC) de 95% para essas medidas.

Para o cálculo do tamanho ótimo amostral, empregou-se o Método de Máxima Curvatura Modificado (MMCM) (Subseção 2.1.3, pág. 30) (LESSMAN; ATKINS, 1963; MEIER; LESSMAN, 1971). Deste modo, inicialmente representou-se a relação entre cada CV obtido e seu conjunto de 2000 subamostras de mesmo tamanho empregando-se a equação de regressão de potência (Fórmula 1, pág. 30). Após esta etapa, realizou-se o cálculo do Coeficiente de Determinação (R^2) para cada turma por meio da Fórmula 3:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

na qual: \hat{y}_i é o valor estimado de cada y_i e \bar{y} é a média das observações y de cada amostra.

A fórmula de cálculo do R^2 emprega, em seu numerador, a diferença entre a média das observações da amostra e o valor estimado de cada observação, realizando o somatório de seus quadrados. Em seu denominador, é calculada a soma dos quadrados das diferenças entre a média das observações e cada valor observado na amostra, resultando na soma total dos quadrados obtidos.

O Coeficiente de Determinação indica, portanto, quanto o modelo empregado mostra-se capaz de explicar as observações obtidas nas amostras, permitindo compreender o quão explicativo é determinado modelo e o quanto ele se ajusta em relação aos valores observados (OZER, 1985). Como a Fórmula 3 resulta em um valor que varia entre zero e um, este resultado foi convertido para porcentagem. Além disso, embora quanto maior o R^2 obtido, melhor seja o ajuste da função aos dados, um R^2 acima de 70% já indica um bom ajuste para os modelos utilizados.

Por fim, representou-se a relação entre os CVs calculados e os tamanhos das subamostras de maneira gráfica em eixos cartesianos, nos quais o eixo y representou o CV e o eixo x retratou o intervalo dos tamanhos de amostra. Após estabelecer as estimativas dos CVs de todas as subamostras de cada turma analisada, calculou-se o valor referente ao ponto da abcissa no qual ocorre a máxima curvatura do modelo utilizado, correspondente ao tamanho ótimo de amostra, denominado X_{MC} e obtido por meio da Fórmula 2 (Subseção 2.1.3, pág. 31).

Em razão da Fórmula 2 resultar num valor real e, para a presente pesquisa, o tamanho ótimo precisar ser um valor inteiro, realizou-se o arredondamento de cada X_{MC} obtido respeitando a regra adotada pela Associação Brasileira de Normas Técnicas (ABNT) (NBR, 1977). Por fim, conforme Meier e Lessman (1971), os tamanhos ótimos resultantes a partir do MMCM devem ser expressos numa medida denominada Unidade Básica (UB), que demanda transformações quando uma UB compreende um conjunto de observações maior que um. Nesta pesquisa, cada UB representou exatamente um estudante, de modo a não ser necessário realizar outras transformações nos tamanhos obtidos além de arredondamentos.

4.2 Resultados e Discussões

Durante a realização do método Bootstrapping, as subamostras com tamanhos muito reduzidos ($n < 5$) apresentaram dispersões de caráter desuniforme e inconclusivo. Tal

comportamento já era esperado, pois necessita-se que haja um número mínimo de observações para ser possível analisar a distribuição das amostras. Deste modo, estabeleceu-se um limite mínimo de cinco observações em cada amostra para o início da análise dos CVs nas turmas examinadas. Esta análise se estendeu até o tamanho original das dez amostras selecionadas para este estudo.

A Figura 10 apresenta os dez eixos cartesianos correspondentes às amostras analisadas, sendo os cinco dispostos na primeira coluna (gráficos de A a E) referentes às turmas que não tiveram contato com o CX e os demais eixos (gráficos de F a J) relativos às turmas que empregaram o sistema durante o semestre letivo. Os CVs exatos de cada tamanho de subamostra foram representados em seus respectivos eixos junto à equação de regressão de potência que indicou a tendência dos dados obtidos. Também calcularam-se os limites superiores e inferiores dos CVs, optando-se por representá-los já aplicados ao modelo, de modo a demonstrar os pontos máximos e mínimos que os CVs seriam capazes de alcançar em cada tamanho de subamostra sob um IC de 95%.

A partir da análise das dez turmas utilizadas, foi possível perceber que o valor do CV diminuiu ao passo que o tamanho das subamostras aumentou. Por meio do MMCM, o tamanho ótimo de amostra obtido entre as cinco turmas que não tiveram contato com o CX variou entre 15 e 22 estudantes, ao passo que nas turmas que empregaram o sistema durante os semestres letivos a estimativa foi de 9 a 19 alunos.

Com base nesses resultados, torna-se possível compreender que as turmas que utilizaram o AEU no decorrer dos semestres letivos tenderam a apresentar pontos de máxima curvatura inferiores aos obtidos naquelas que não tiveram contato com a tecnologia. Isso demonstra que quantidades de observações menores que as originalmente existentes nas amostras já teriam sido suficientes para representar o comportamento da variável sob análise. Além disso, como entre as turmas que utilizaram o CX o nível médio de variabilidade das notas foi ainda menor que naquelas que não utilizaram o sistema, é possível ratificar a hipótese de que maiores índices de nivelamento entre o desempenho dos alunos podem ocorrer em razão dos mesmos estudarem os conteúdos das disciplinas a partir de uma fonte em comum, conforme presumido em Brant-Ribeiro, Biase e Cattelan (2015).

Como o MMCM possibilita o cálculo do valor da abcissa referente ao ponto onde ocorre a máxima curvatura do modelo, e isso indica o tamanho ótimo de amostra para se analisar uma determinada variável, também é necessário propor um tamanho capaz de

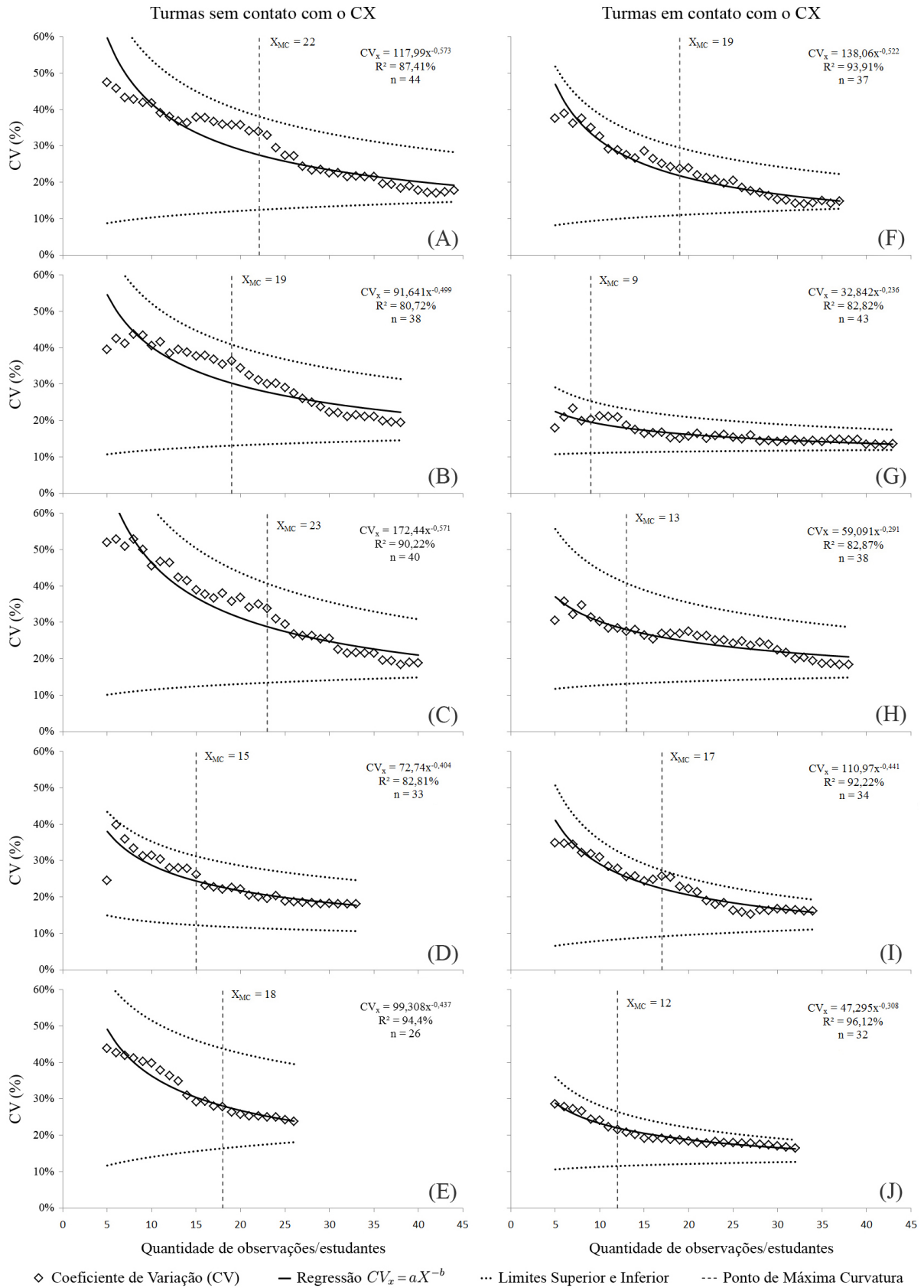


Figura 10 – Relações entre os CVs e os tamanhos de subamostras para a variável desempenho dos estudantes em dez turmas de cursos da área computacional (Fonte: Brant-Ribeiro e Cattelan (2016)).

abranger todas as amostras que contenham observações dessa variável. Em função disso, como as turmas que utilizaram o CX apresentaram tamanhos que variaram entre 9 e 19 alunos, infere-se que o tamanho ótimo de amostra que permite a análise do desempenho de alunos em contato com tecnologias educacionais seja de pelo menos 20 indivíduos. Do mesmo modo, como os tamanhos obtidos entre as turmas que não tiveram contato com o AEU variaram entre 15 e 22 alunos, propõe-se que para se investigar turmas que não utilizam sistemas educacionais, no mínimo sejam observados 25 estudantes.

Além disso, por meio da análise das linhas de tendência presentes nos eixos da Figura 10, também foi possível perceber que a diminuição do CV não apresentou um comportamento proporcional ao aumento dos tamanhos das subamostras, acentuando-se no início e tornando-se mais próxima à estabilidade em quantidades mais elevadas de observações. Deste modo, caso as amostras analisadas tivessem sido ainda maiores, pressupõe-se que se alcançaria um ponto em que as linhas de tendência dos limites superior e inferior e do CV acabariam se interceptando. Por meio deste comportamento, é possível deduzir que o aumento do tamanho das amostras apenas se manteve conveniente até determinado ponto, a partir do qual o emprego de quantidades ainda maiores de observações não foi compensatório, em função dos baixos ganhos em precisão experimental.

Por fim, o R^2 alcançado nas análises manteve-se acima de 80% em todas as amostras estudadas, o que demonstra boa precisão de ajuste das regressões e alta confiabilidade nas estimativas obtidas nesta pesquisa. O coeficiente de regressão b , capaz de representar a variabilidade dos experimentos sob análise, variou de 0,404 a 0,573 entre as cinco turmas que não tiveram contato cotidiano com o CX e de 0,236 a 0,522 nas turmas que o empregaram durante seus respectivos semestres letivos (Tabela 2). É possível observar, inclusive, que os maiores valores de b observados entre as turmas foram obtidos nas amostras A (0,573) e C (0,571), que fizeram parte do grupo de turmas que não tiveram contato com o CX e apresentaram os maiores X_{MC} obtidos (22 e 23, respectivamente) entre todas as turmas analisadas. Além disso, nenhuma dessas amostras apresentou um valor do parâmetro b menor que 0,4.

Tendo por base que o coeficiente de regressão b representa a taxa de mudança que ocorre ao longo da curva de regressão e explicita o grau de variabilidade das amostras, tais resultados reforçam a hipótese de que as turmas que não empregaram o CX no decorrer dos semestres letivos apresentaram um comportamento mais heterogêneo do que aquele

Tabela 2 – Coeficientes de Variação (CVs) unitários e de regressão obtidos a partir da análise das dez turmas utilizadas para o cálculo de tamanho ótimo amostral (Fonte: Adaptado de Brant-Ribeiro e Cattelan (2016)).

Amostra ¹	Parâmetro a	Parâmetro b
A	117,99	0,573*
B	91,641	0,499
C	172,44	0,571
D	72,74	0,404
E	99,308	0,437
F	138,06	0,522
G	32,842	0,236**
H	59,091	0,291
I	110,97	0,441
J	47,295	0,308

¹Amostras de A a E representam turmas que não tiveram contato com o CX; Amostras de F a J representam turmas que empregaram o CX no decorrer dos semestres letivos; Parâmetro a : Coeficiente de variação de amostras de tamanho unitário; Parâmetro b : Coeficiente de regressão das amostras; *: Maior valor de b obtido entre as amostras; **: Menor valor de b obtido entre as amostras.

observado nas turmas que utilizaram o AEU. De fato, os menores valores de b observados foram obtidos justamente nas amostras G (0,236), H (0,291) e J (0,308), compreendidas entre as turmas que tiveram contato com o sistema e que obtiveram os menores X_{MC} (9, 13 e 12, respectivamente) dentre todas as turmas investigadas nesta pesquisa.

4.3 Considerações Finais

Esta etapa da pesquisa teve como objetivo estimar o tamanho ótimo amostral para a análise do desempenho acadêmico de estudantes em cursos de graduação em Sistemas de Informação e graduação e pós-graduação em Ciência da Computação. Determinou-se um tamanho ótimo de 25 alunos para a análise desta variável em turmas que não utilizam

plataformas de apoio ao ensino no cotidiano e de 20 estudantes para turmas que possuem contato com tecnologias educacionais (tais como AEU's) durante os semestres letivos.

Por meio da interpretação das linhas de tendência geradas, percebeu-se que a diminuição dos CVs não apresentou um comportamento proporcional à elevação dos tamanhos das amostras, intensificando-se apenas no começo e tornando-se estável à medida que as amostras tiveram suas quantidades de observações acrescidas. Em função disso, é possível compreender que a elevação do tamanho das amostras se manteve conveniente apenas até determinado ponto, a partir do qual o emprego de quantidades maiores de observações não foi compensado devido ao ganho reduzido em precisão experimental.

Durante a realização desta análise, um fator limitante observado referiu-se à variável que foi investigada para a estimativa de tamanho ótimo. Os cálculos realizados permitiram a determinação de tamanhos amostrais confiáveis para a realização de novas pesquisas que envolvam especificamente as notas de estudantes de graduação ou pós-graduação. Para investigações que empreguem outras variáveis tais como assiduidade ou, até mesmo, o desempenho de estudantes em outros níveis de ensino, mostra-se necessária a realização de novos procedimentos experimentais para se calcular tamanhos ótimos em cada uma dessas circunstâncias, algo que não foi compreendido por este trabalho.

Além disso, o emprego do desempenho discente para a determinação de um tamanho ótimo amostral pode gerar questionamentos de natureza pedagógica, visto que se trata de um fator que não necessariamente representa com fidelidade o nível de conhecimento dos estudantes. No entanto, a nota obtida por um aluno se mantém como o elemento-chave que o sistema educacional brasileiro emprega para avaliar aprovações e reprovações. Acredita-se que encontrar maneiras de se analisar com alto índice de confiabilidade os fluxos de dados que se enquadram nesta categoria ainda se mostra como um assunto relevante para que debates acerca dos critérios de avaliação estudantis se mantenham em foco e o amadurecimento sobre esta temática continue a ocorrer.

Os resultados referentes aos tamanhos ótimos amostrais aqui obtidos reforçam a presunção de que quantidades de observações menores que as inicialmente existentes nas amostras já teriam sido suficientes para a compreensão do comportamento da variável estudada. Deste modo, pesquisas que objetivam realizar estudos acerca do desempenho de estudantes em circunstâncias nas quais novas tecnologias são introduzidas podem empregar as estimativas obtidas neste trabalho para realizarem suas investigações com confiabilidade em suas análises.

Para uma análise ainda mais robusta acerca do desempenho acadêmico dos estudantes, recomenda-se empregar amostras de pelo menos 25 observações em ambas as vertentes demonstradas. Desta maneira, mostra-se possível compreender a variabilidade de quaisquer turmas neste escopo e se realizar uma investigação capaz de abranger corretamente as populações de interesse.

Análise Fatorial do Impacto de AEU's no Desempenho Discente

Neste capítulo serão apresentados os resultados da Análise de Variância (ANOVA) (Subseção 2.1.4, pág. 31) que foi realizada para se investigar a influência dos fatores “Uso do Classroom eXperience (CX)” e “Disciplinas” no desempenho acadêmico de estudantes em cursos de graduação e pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia (FACOM/UFU). Deste modo, inicialmente o método de pesquisa e os materiais utilizados são explicados detalhadamente na Seção 5.1; na Seção 5.2 são apresentados resultados obtidos e discussões sobre a aplicação da ANOVA e, por fim, considerações finais são pontuadas na Seção 5.3.

5.1 Material e Métodos

Para a realização desta análise¹, foram utilizados dados provenientes de vinte turmas que cursaram disciplinas da grade curricular dos cursos de graduação em Sistemas de Informação e graduação e pós-graduação em Ciência da Computação da FACOM/UFU, totalizando 501 observações. A variável analisada foi o desempenho obtido pelos estudantes em disciplinas cursadas durante seis anos letivos (de 2012 a 2017) e, para isso, as notas alcançadas pelos discentes foram coletadas ao final de cada semestre letivo. De

¹ Pesquisa aprovada pelo Comitê de Ética da UFU (CAAE 46909515.4.0000.5152).

modo a assegurar a confiabilidade desses dados, objetivou-se que os mesmos professores permanecessem lecionando suas respectivas disciplinas durante os semestres analisados, sendo também orientados sobre a necessidade de empregarem uma sequência similar de exposição de conteúdos em suas aulas, objetos de avaliação semelhantes e que os mesmos fossem aplicados em datas proporcionalmente aproximadas.

Optou-se também por desconsiderar as notas de alunos que apresentaram, ao final dos semestres, rendimentos muito destoantes dos demonstrados em outros momentos dos mesmos períodos letivos. Entende-se que esses alunos apresentaram maiores possibilidades de terem se comportado de modo displicente ao final dos semestres, visto que já haviam conseguido cumprir os requisitos mínimos necessários para aprovação nas disciplinas.

Por ter se mostrado necessária a realização de um planejamento robusto acerca do modo com que as etapas desta análise ocorreriam, empregou-se o Gerenciador de Experimentos (Capítulo 3, pág. 45) para se organizar todos os níveis de análise que compuseram esta etapa da pesquisa, evitando inconsistências durante o processo de coleta de dados e, conseqüentemente, eliminando enviesamentos nos resultados aqui obtidos.

Embora o tamanho das amostras utilizadas para a presente pesquisa não tenha se mantido fixo, respeitou-se a estimativa de tamanho ótimo amostral determinada nesta tese (Capítulo 4, pág. 55) e assegurou-se que as amostras cumprissem um mínimo médio de 25 estudantes por turma – o que garantiu a confiabilidade nos resultados obtidos neste trabalho. Deste modo, inicialmente as observações coletadas (notas dos estudantes) foram submetidas ao teste de Shapiro-Wilk (Subseção 2.1.1, pág. 27) para verificar se os resíduos da variável observada seguiam a distribuição normal e ao teste de Levene para averiguar se as amostras possuíam variâncias homogêneas (TRIOLA, 2014).

Posteriormente, realizou-se uma Análise de Variância Fatorial 4x5 (*Two-way ANOVA*) (Subseção 2.1.4, pág. 31) para se atestar a existência de diferenças significativas entre os níveis de dois fatores investigados. O primeiro fator, denominado “Uso do CX”, caracterizou-se como o modo com que o CX foi empregado pelos estudantes e possuiu quatro níveis distintos: 1) CX Ausente, 2) Uso isolado do CX, 3) Uso do CX + Módulo Social (MS) e 4) Uso do CX + MS + Gamificação. O segundo fator, intitulado “Disciplinas”, compreendeu todas as disciplinas que foram cursadas pelos alunos empregando os níveis de uso do CX supracitados e possuiu cinco níveis: Lógica Matemática (LM), Otimização, Arquitetura e Organização de Computadores (AOC), Interação Humano-

Computador (IHC) (Graduação) e IHC (Pós-graduação). O delineamento fatorial utilizado e a organização dos níveis dos fatores podem ser observados na Figura 11.

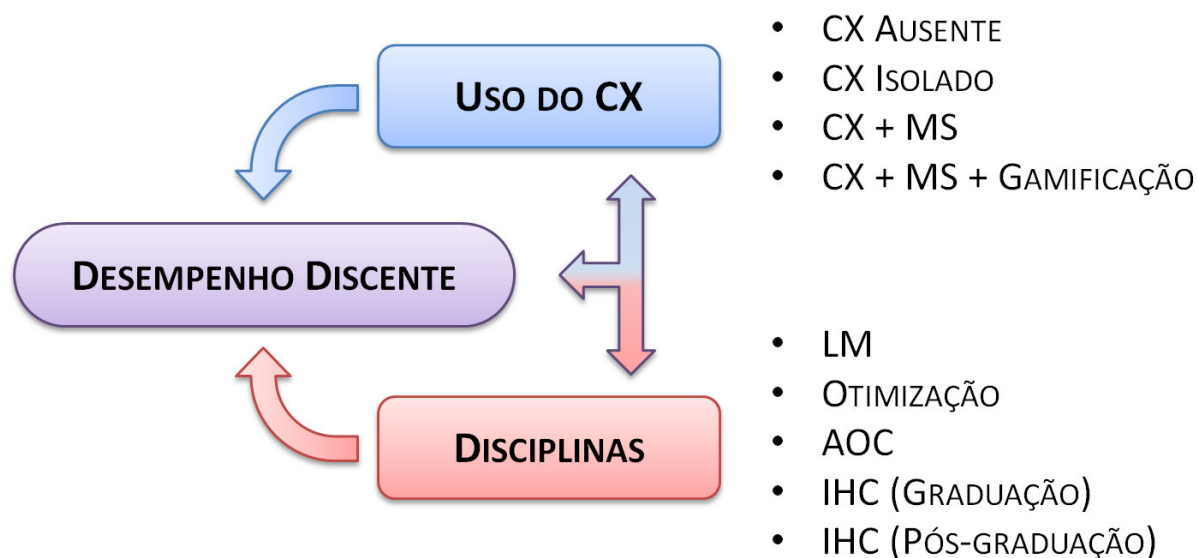


Figura 11 – Análise de Variância Fatorial 4x5 (*Two-way ANOVA*) dos fatores “Uso do CX” e “Disciplinas” sobre o desempenho discente.

Para a realização da ANOVA, inicialmente foi executado o teste *F* de Snedecor (Subseção 2.1.4, pág. 31) com o intuito de se atestar a existência de diferenças significativas entre as médias dos níveis dos fatores analisados (SNEDECOR; COCHRAN, 1980). Este teste também foi utilizado para se observar a existência de uma interação entre a utilização do CX e as disciplinas cursadas pelos estudantes, de modo a ser possível compreender se o desempenho obtido pelos alunos mostra-se influenciável tanto pelo tipo da disciplina cursada quanto pelo nível de uso do CX durante os semestres letivos. Como o teste *F* de Snedecor não é capaz de indicar especificamente quais são os níveis que possuem as médias significativamente diferentes, empregou-se o teste de Tukey (Subseção 2.1.4, pág. 32), para realizar a comparação de todas as médias amostrais executando análises consecutivas entre os pares de amostras até que todas as possibilidades combinatórias fossem realizadas (TRIOLA, 2014).

5.2 Resultados e Discussões

A realização dos testes de Shapiro-Wilk, Levene e F de Snedecor possibilitou atestar a existência de diferenças significativas entre os níveis dos fatores “Uso do CX” e “Disciplinas”. Como resultados, na Tabela 3 é possível observar que a ANOVA do desempenho de estudantes apresentou efeito significativo em três circunstâncias distintas: o uso do CX (F1), as disciplinas cursadas (F2) e também para a existência de uma interação entre os dois fatores (F1 x F2).

Tabela 3 – Análise de Variância Fatorial do desempenho discente com relação ao uso do Classroom eXperience (CX) em disciplinas de graduação e pós-graduação em cursos da área computacional (Fonte: Brant-Ribeiro e Cattelan (2018)).

Fatores de Variação ¹	Graus de Liberdade	Quadrado Médio	$F(P)$
Uso do CX (F1)	3	3263,554	24,044 (<0,01)
Disciplinas (F2)	4	3971,876	29,263 (<0,01)
Interação (F1 x F2)	12	648,836	4,780 (<0,01)
Resíduo	481	135,730	
$CV(\%)$; $W(P)$; $F'(P)$	14,92 ; 0,0961 (<0,01) ; 3,110 (0,062)		

¹ $F(P)$: Estatística e Probabilidade do teste F de Snedecor; Probabilidades menores que 0,05 indicam a existência de ao menos um tratamento (nível) diferente dos demais; CV : Coeficiente de Variação; $W(P)$ e $F'(P)$: Estatísticas dos testes de Shapiro-Wilk e Levene; Probabilidades maiores que 0,05 indicam resíduos com distribuição normal e variâncias homogêneas, respectivamente.

Devido a isso, o maior impacto deste resultado incide sobre a característica de interação entre o uso do CX e as disciplinas cursadas pelos estudantes, ou seja, o desempenho dos alunos mostrou-se de fato influenciado tanto pelo nível em que a ferramenta foi utilizada quando pelo tipo de disciplina cursada. Tal resultado permite deduzir que a natureza das disciplinas sob análise (sejam elas teóricas ou práticas, por exemplo) possui impacto sobre o desempenho obtido pelos estudantes e mostra-se como um fator que também deve ser levado em consideração ao se realizar análises que objetivam observar o impacto do uso de plataformas de apoio ao ensino. Além disso, também é possível observar que o Coeficiente de Variação (CV) experimental alcançado mostrou-se baixo (14,92%), garantindo uma variabilidade sintetizada para esta pesquisa.

Ainda acerca da Tabela 3, ao serem submetidas ao teste de Shapiro-Wilk, as observações da variável analisada indicaram resíduos que não seguiram a distribuição normal. Para o teste de Levene, as amostras do experimento apresentaram variâncias homogêneas. Conforme atestado por Scheffé (1959), a ANOVA não se mostra sensível a desvios de normalidade, empregando-se, neste caso, a homocedasticidade como pressuposição prioritária para a tomada de decisões. Deste modo, a estatística F de Snedecor, para os efeitos principais e de interação, foi aplicada sem negligenciar a pressuposição de homogeneidade de variâncias, ao passo que a ausência de normalidade não culminou em perda de poder na análise devido à robustez desta estatística.

Com o intuito de se explicar graficamente os resultados obtidos a partir do teste F de Snedecor (Tabela 3), a Figura 12 ilustra a interação entre as cinco disciplinas investigadas e os quatro níveis de uso do CX. É possível observar que a disciplina IHC (Pós-graduação) apresentou as maiores médias de desempenho dentre todas as amostras analisadas, o que acarretou numa ausência de interação entre ela e as disciplinas de graduação. De todo modo, é interessante pontuar que houve um comportamento semelhante e adjacente nos

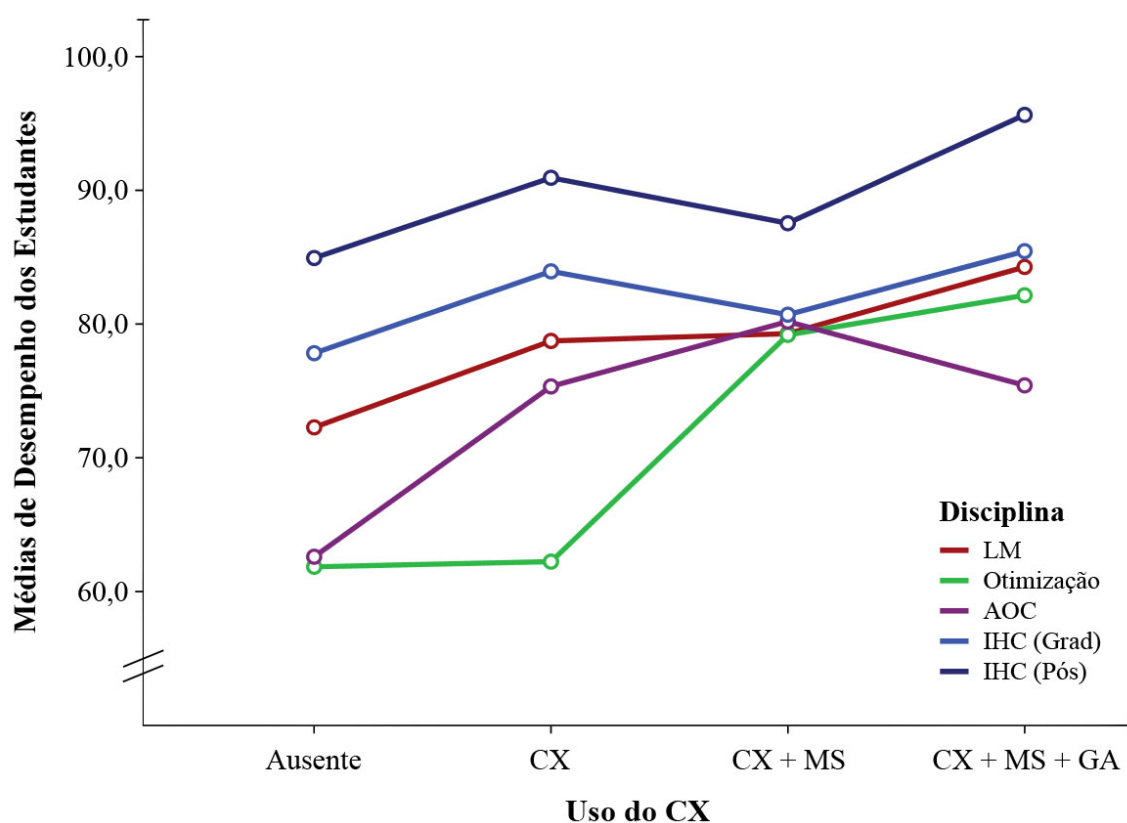


Figura 12 – Influência da interação entre os fatores “Uso do CX” e “Disciplinas” sobre o desempenho dos estudantes.

desempenhos obtidos pelos estudantes que cursaram ambas as disciplinas de IHC, tanto na graduação quanto na pós-graduação, o que denota um padrão comportamental acerca de como os estudantes são influenciados pelo CX com relação aos tópicos abordados nesta temática e como ele se difere das demais disciplinas cursadas.

Também é possível observar na Figura 12 que houve uma interação explícita entre todas as disciplinas de graduação analisadas em razão delas terem obtido acréscimos e decréscimos de diferentes magnitudes e terem se interceptado em determinado ponto do gráfico. Isso acontece em razão de existirem disciplinas com temáticas que possuem maior tendência a serem favorecidas pela utilização do CX e seus módulos interativos, tal como ocorre, por exemplo, em circunstâncias de resolução de exercícios propostos – em que os alunos podem socializar e colaborar entre si apresentando resoluções e enriquecendo o conteúdo apresentado pelos professores em sala de aula.

Por ter havido um efeito significativo de interação entre os fatores “Uso do CX” e “Disciplinas” (conforme apresentado na Tabela 3), também realizaram-se os desdobramentos deste resultado comparando-se os pares das médias amostrais do desempenho dos estudantes pelo teste de Tukey (Tabelas 4 e 5).

Tabela 4 – Desempenho discente com relação ao uso do CX em disciplinas de graduação e pós-graduação em cursos da área computacional com enfoque nas disciplinas cursadas pelos estudantes (Fonte: Brant-Ribeiro e Cattelan (2018)).

Uso do CX ¹	Disciplinas					Média
	LM	Otimização	AOC	IHC (Grad)	IHC (Pós)	
A	72,28 B	61,86 B	62,62 B	77,83 B	84,94 A	71,11
B	78,74 AB	62,25 B	75,34 A	83,94 AB	90,94 A	77,88
C	79,29 AB	79,20 A	80,19 A	80,70 AB	87,54 A	80,77
D	84,27 A	82,15 A	75,42 A	85,45 A	95,64 A	82,76
Média	78,88	71,73	73,40	81,88	89,28	

¹Médias seguidas por letras distintas maiúsculas nas colunas diferem-se entre si pelo teste de Tukey a 0,05 de significância; LM: Lógica Matemática; AOC: Arquitetura e Organização de Computadores; IHC: Interação Humano-Computador; Grad: Turmas de graduação; Pós: Turmas de pós-graduação; A: CX Ausente; B: Uso isolado do CX; C: Uso do CX junto ao Módulo Social; D: Uso do CX junto aos Módulos Social e de Gamificação.

Ao se analisar o desdobramento da interação com enfoque nas disciplinas cursadas pelos estudantes (Tabela 4), é possível perceber que o uso do CX, de modo geral, resultou em acréscimos gradativos no desempenho dos alunos. As disciplinas Lógica Matemática (LM) e Interação Humano-Computador (IHC) (Graduação) apresentaram resultados equivalentes, alcançando aumentos significativos nas médias de desempenho logo após a inclusão do módulo de Gamificação (nível D) no CX. No entanto, as turmas dessas disciplinas que empregaram o CX isoladamente (nível B) ou com apenas a inclusão do Módulo Social (MS) (nível C) não apresentaram acréscimos em suas notas, mostrando-se similares às turmas que cursaram tais disciplinas sem a utilização o CX (nível A).

Na disciplina Otimização, a turma que empregou o CX isoladamente (nível B) não apresentou um desempenho superior ao daquela que não o utilizou, fato similar ao ocorrido em LM e IHC (Graduação). No entanto, em Otimização o acréscimo do MS (nível C) provocou um aumento significativo das notas, gerando uma elevação no desempenho tanto dos alunos que utilizaram o módulo social isoladamente quanto daqueles que o empregaram com a Gamificação. Ainda neste escopo, a partir da inserção do CX isolado na disciplina AOC (nível B), já foi possível observar um aumento do desempenho dos alunos e essa característica se manteve em todos os demais níveis de uso do CX.

Em suma, para todas as turmas de graduação analisadas nesta pesquisa, ocorreram aumentos significativos no desempenho dos alunos num comparativo direto entre a ausência do CX (nível A) e a utilização desta ferramenta junto aos módulos de Socialização e Gamificação (nível D). Analisando as disciplinas de Otimização e AOC, é possível perceber que este aumento ocorreu ainda antes, pois em Otimização a inclusão do MS já gerou aumentos robustos e em AOC a inserção isolada do CX já garantiu a elevação do desempenho acadêmico discente.

Para IHC (Pós-graduação), não se observou diferença no desempenho dos alunos entre quaisquer níveis de uso do CX. Pressupõe-se que, conforme indicado por Brant-Ribeiro, Biase e Cattelan (2015), alunos de pós-graduação normalmente já possuem seus próprios métodos de estudo e tendem a ser esforçados – algo que faz com que seus níveis de desempenho sejam naturalmente elevados com ou sem a utilização de tecnologias. A inserção de novas tecnologias neste escopo acaba se configurando como mais uma opção, diluindo sua importância e conveniência de uso do sistema.

Tal característica pode ser confirmada ao se analisar o desdobramento da ANOVA com enfoque nos níveis de utilização do CX (Tabela 5), pois a disciplina IHC (Pós-graduação) se manteve com os maiores desempenhos em todas as circunstâncias observadas. Acerca das turmas que não empregaram o CX durante os semestres letivos (nível A), o desempenho das disciplinas de IHC (tanto na graduação quanto na pós-graduação) mostrou-se similar, ao passo que Otimização e AOC apresentaram desempenhos menores e LM obteve um resultado intermediário e superior a essas, mas que não se diferiu significativamente de IHC (Graduação).

Tabela 5 – Desempenho discente com relação ao uso do CX em disciplinas de graduação e pós-graduação em cursos da área computacional com enfoque nos níveis de utilização do CX (Fonte: Brant-Ribeiro e Cattelan (2018)).

Uso do CX ¹	Disciplinas					Média
	LM	Otimização	AOC	IHC (Grad)	IHC (Pós)	
A	72,28 <i>b</i>	61,86 <i>c</i>	62,62 <i>c</i>	77,83 <i>ab</i>	84,94 <i>a</i>	71,11
B	78,74 <i>b</i>	62,25 <i>c</i>	75,34 <i>b</i>	83,94 <i>ab</i>	90,94 <i>a</i>	77,88
C	79,29 <i>a</i>	79,20 <i>a</i>	80,19 <i>a</i>	80,70 <i>a</i>	87,54 <i>a</i>	80,77
D	84,27 <i>bc</i>	82,15 <i>bc</i>	75,42 <i>c</i>	85,45 <i>b</i>	95,64 <i>a</i>	82,76
Média	78,88	71,73	73,40	81,88	89,28	

¹Médias seguidas por letras distintas minúsculas nas linhas diferem-se entre si pelo teste de Tukey a 0,05 de significância; LM: Lógica Matemática; AOC: Arquitetura e Organização de Computadores; IHC: Interação Humano-Computador; Grad: Turmas de graduação; Pós: Turmas de pós-graduação; A: CX Ausente; B: Uso isolado do CX; C: Uso do CX junto ao Módulo Social; D: Uso do CX junto aos Módulos Social e de Gamificação.

Para o uso isolado do CX (nível B), as disciplinas de IHC mantiveram-se semelhantes. AOC e LM apresentaram desempenhos similares aos de IHC (Graduação), enquanto Otimização manteve-se com o menor desempenho dentre as disciplinas nesta circunstância. Com a inserção do MS (nível C), todas as disciplinas mostraram-se semelhantes, ao passo que com a inclusão da Gamificação (nível D), a pós-graduação destacou-se novamente com um desempenho maior que os demais.

Analisando-se apenas as disciplinas de graduação no último nível de uso do CX, é possível verificar que AOC apresentou o menor desempenho dentre as disciplinas analisadas.

De todo modo, este não deve ser considerado um desempenho ruim, devido à disciplina ter apresentado uma média de 75,42. Já as demais disciplinas alcançaram desempenhos superiores, ainda que LM e Otimização não tenham se diferenciado significativamente de AOC. A disciplina IHC (Graduação) obteve o melhor resultado entre as disciplinas de graduação com o uso da Gamificação, alcançando um desempenho médio de 85,45.

Por fim, retomando a análise da Figura 12, é possível observar que todas as disciplinas apresentaram acréscimos gradativos nos desempenhos obtidos nesta pesquisa. A disciplina de IHC (Pós-graduação) manteve-se acima das demais para todos os níveis de uso do CX, ao passo que IHC (Graduação) e LM mantiveram-se intermediárias e permaneceram próximas entre si. Otimização e AOC também tiveram aumentos consideráveis de desempenho e tenderam a se aproximar das demais disciplinas a partir das inserções dos módulos de Socialização e Gamificação no CX.

5.3 Considerações Finais

O emprego de um método bem definido para se organizar e evitar inconsistências durante as etapas de pesquisa tem se mostrado como uma atividade indispensável para a realização de investigações confiáveis. Com isso em mente, esta pesquisa teve como objetivo realizar uma ANOVA Fatorial 4x5 para se atestar a existência de diferenças significativas entre os fatores “Uso do CX” e “Disciplinas” sobre o desempenho acadêmico de estudantes em cursos de graduação em Sistemas de Informação e graduação e pós-graduação em Ciência da Computação.

Inicialmente, foi projetado e desenvolvido um Gerenciador de Experimentos (Capítulo 3, pág. 45) com o intuito de se organizar o nível de visibilidade das funcionalidades presentes em ferramentas de apoio ao ensino e se automatizar o planejamento experimental de pesquisas. O Gerenciador de Experimentos mostrou-se capaz de automatizar a execução deste experimento ao permitir a habilitação e desabilitação de funcionalidades específicas do CX para grupos distintos de estudantes que cursaram disciplinas dos cursos de graduação e pós-graduação da FACOM/UFU. Deste modo, seu emprego se mostrou como uma prática vantajosa, tendo em vista que permitiu controlar o nível de acesso que grupos distintos de indivíduos tiveram ao utilizarem a plataforma CX como uma ferramenta de apoio ao ensino no cotidiano acadêmico.

Por meio da utilização do Gerenciador de Experimentos, realizou-se uma ANOVA que empregou 501 observações referentes aos desempenhos obtidos pelos estudantes ao final dos semestres letivos. A partir desta análise, percebeu-se preliminarmente a existência de efeitos significativos em três casos: na influência do uso do CX sobre as notas dos estudantes, no efeito que disciplinas diferentes cursadas pelos alunos causaram em seus desempenhos e, também, na interação existente entre estes dois fatores. Em razão disso, o maior impacto deste resultado reflete-se diretamente sobre as características de interação entre a utilização do CX e as disciplinas cursadas pelos estudantes, ou seja, o desempenho dos alunos é uma variável influenciada tanto pelas funcionalidades da ferramenta quanto pela natureza da disciplina cursada pelos estudantes, ambas em conjunto.

Para todas as turmas de graduação investigadas neste trabalho, houve aumentos significativos no desempenho dos estudantes num comparativo direto entre a ausência do CX e a utilização desta ferramenta junto aos módulos de Socialização e Gamificação. Numa análise específica acerca das disciplinas de Otimização e AOC, foi possível concluir que este aumento ocorreu ainda antes, pois em Otimização a inclusão do MS já foi capaz de gerar aumentos significativos nas notas e em AOC a inserção isolada do CX já garantiu um aumento robusto no desempenho discente.

No entanto, para a disciplina de IHC (Pós-graduação) não se observou diferença significativa no desempenho dos alunos entre quaisquer níveis de uso do CX. Isso permite deduzir que estudantes de pós-graduação já possuem seus próprios hábitos e métodos de estudo, conseguindo obter notas naturalmente mais elevadas que os alunos de graduação. Conclui-se, portanto, que propostas de novas tecnologias educacionais para pós-graduandos acabam se mostrando como mais uma opção elencável, o que dilui a importância e conveniência de utilização da plataforma para este público.

Além disso, durante a realização desta pesquisa também foi observado um fator limitante referente às disciplinas que foram avaliadas empregando-se a ANOVA. Das cinco disciplinas cursadas pelas vinte turmas utilizadas nesta análise, apenas uma delas foi cursada por estudantes de pós-graduação. Os resultados obtidos e as discussões feitas com os dados das turmas de pós-graduação refletem, portanto, o comportamento que os alunos tiveram especificamente na disciplina de Interação Humano-Computador (IHC).

Embora esta característica não afete a integridade do tamanho ótimo amostral determinado no Capítulo 4 (visto que, para essa estimativa, turmas de disciplinas distintas de pós-graduação foram utilizadas), não é possível generalizar os resultados aqui obtidos

para todas as disciplinas existentes em cursos de pós-graduação em Ciência da Computação. Enxerga-se, portanto, a necessidade de se realizar uma ANOVA englobando mais turmas com estudantes pós-graduandos, de modo a ser possível analisar adequadamente o impacto que AEU's (e outras ferramentas de apoio ao ensino) causam no desempenho de turmas de pós-graduação em ambientes acadêmicos reais.

De todo modo, todas as disciplinas analisadas apresentaram acréscimos gradativos nos desempenhos obtidos durante esta pesquisa. A disciplina de IHC (Pós-graduação) manteve-se acima das demais para todos os níveis de utilização do CX, ao passo que, na graduação, IHC e LM mantiveram-se intermediárias e permaneceram próximas entre si. Otimização e AOC também apresentaram aumentos em suas médias de desempenho e tenderam a se aproximar das demais disciplinas a partir das inserções de funcionalidades do sistema, tais como os módulos de Socialização e Gamificação do CX.

Análise Correlacional entre o Desempenho Discente e o Nível de Interações com o CX

Neste capítulo serão apresentados os resultados obtidos a partir da Análise Correlacional (Subseção 2.1.5, pág. 32) realizada para verificar a existência de relações entre o desempenho acadêmico de estudantes em cursos de graduação e pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia (FACOM/UFU) e suas interações com a plataforma educacional Classroom eXperience (CX). Deste modo, inicialmente o método de pesquisa e os materiais utilizados são explicados na Seção 6.1; na Seção 6.2 são apresentados os resultados obtidos a partir da correlação realizada e, por fim, considerações finais são pontuadas na Seção 6.3.

6.1 Material e Métodos

Para a realização desta análise¹, foram utilizados dados advindos de quinze turmas que cursaram disciplinas da grade curricular dos cursos de graduação em Sistemas de Informação e graduação e pós-graduação em Ciência da Computação da FACOM/UFU, totalizando 378 observações. Tais dados compreenderam um conjunto amostral semelhante ao empregado para a realização da Análise de Variância (ANOVA) (Capítulo 5,

¹ Pesquisa aprovada pelo Comitê de Ética da UFU (CAAE 46909515.4.0000.5152).

pág. 65), no entanto, nesta etapa foram desconsideradas as turmas que não tiveram contato com o CX durante os períodos letivos. Enxergou-se essa necessidade devido ao objetivo desta análise ter sido observar a relação entre a nota alcançada pelos estudantes e as interações que os mesmos tiveram com a plataforma CX, sendo, portanto, dispensável o emprego de dados de turmas que não utilizaram a ferramenta.

A principal variável analisada foi o desempenho acadêmico obtido pelos estudantes em disciplinas cursadas durante seis anos letivos (de 2012 a 2017) e, para isso, suas notas foram coletadas ao final de cada semestre letivo. As disciplinas empregadas nesta análise foram: Lógica Matemática (LM), Otimização, Arquitetura e Organização de Computadores (AOC), Interação Humano-Computador (IHC) (Graduação) e IHC (Pós-graduação). Para cada disciplina, foram coletadas as notas de turmas distintas que tiveram três níveis de contato com a plataforma CX: A) Uso isolado do CX, B) Uso do CX junto ao Módulo Social (MS) e C) Uso do CX junto ao MS e à Gamificação.

De modo a assegurar a confiabilidade desses dados, objetivou-se que os mesmos professores permanecessem lecionando suas respectivas disciplinas durante os semestres analisados, sendo também orientados sobre a necessidade de empregarem uma sequência similar de exposição de conteúdos em suas aulas, objetos de avaliação semelhantes e que os mesmos fossem aplicados em datas proporcionalmente aproximadas. Optou-se também por desconsiderar as notas de alunos que apresentaram, ao final dos semestres letivos, rendimentos muito destoantes dos demonstrados em outros momentos dos mesmos períodos. Entende-se que esses alunos apresentaram maiores possibilidades de terem se comportado de modo displicente ao final dos semestres, visto que já haviam conseguido cumprir os requisitos mínimos necessários para aprovação nas disciplinas.

Por também ter se mostrado necessária a realização de um planejamento adequado sobre o modo com que esta análise ocorreria, o Gerenciador de Experimentos (Capítulo 3, pag. 45) foi utilizado para se estruturar os níveis e etapas que compuseram esta fase da pesquisa, evitando inconsistências durante a coleta de dados e, deste modo, eliminando enviesamentos nos resultados obtidos. Embora o tamanho das amostras não tenha se mantido constante, respeitou-se a estimativa de tamanho ótimo amostral determinada nesta tese (Capítulo 4, pag. 55) e assegurou-se que as amostras cumprissem um mínimo de 25 estudantes por turma – o que garantiu a confiabilidade dos resultados.

Além do desempenho acadêmico, também foram coletadas quatro variáveis referentes aos fluxos interativos que os estudantes tiveram com a plataforma CX: 1) *Login*, 2) Acesso

às Aulas (AA), 3) Socializações e Colaborações (SC) e 4) Respostas aos Quizzes (RQ), conforme detalhado na Tabela 6. Neste contexto, a variável *Login* compreendeu todas as vezes em que cada um dos estudantes inseriu suas credenciais de acesso e fez *login* no CX; a variável AA referiu-se a todos os acessos que cada aluno realizou às aulas capturadas pelos professores; a variável SC representou todas as atividades de socialização e colaboração realizadas pelos estudantes e a variável RQ compreendeu todas as tentativas de respostas que os alunos deram aos Quizzes cadastrados pelos professores.

Tabela 6 – Variáveis referentes aos fluxos interativos que os estudantes tiveram com a plataforma CX.

Variáveis	Descrição
<i>Login</i>	Quantidade de <i>logins</i> realizados no CX.
Acesso às Aulas (AA)	Quantidade de vezes que as aulas foram abertas e acessadas pelos estudantes matriculados.
Socializações e Colaborações (SC)	Quantidade de atividades sociais e colaborativas realizadas, tais como ranqueamentos de <i>slides</i> , comentários em <i>slides</i> e respostas/curtidas em comentários feitos por outros usuários.
Respostas aos Quizzes (RQ)	Quantidade de respostas dadas aos Quizzes cadastrados pelos professores das disciplinas.

As amostras compostas pelas notas dos estudantes foram inicialmente submetidas ao teste de Shapiro-Wilk para verificar se seus resíduos seguiam a distribuição normal (Subseção 2.1.1, pág. 27) (TRIOLA, 2014). Por não ter sido alcançada a normalidade em alguns casos, optou-se por empregar a Correlação de Spearman (Subseção 2.1.5, pág. 32) para se observar a relação entre o desempenho obtido pelos estudantes e a quantidade de interações realizadas com o CX.

Inicialmente, realizou-se uma análise correlacional entre o desempenho discente e a quantidade de *logins* realizados na plataforma CX. Para esta finalidade, os níveis de uso do CX foram organizados em: A) Uso isolado do CX, B) Uso do CX junto ao Módulo Social (MS), C) Uso do CX junto ao MS e à Gamificação e D) Agrupamento desses três níveis para investigar se a relação entre o desempenho dos alunos e a quantidade de *logins* realizados mostrava-se como um fenômeno concentrado ou espalhado entre as diferentes turmas que cursaram a mesma disciplina.

Com base nos resultados obtidos, investigou-se também a relação entre as notas alcançadas pelos estudantes e as variáveis AA, SC e RQ (Tabela 6) – que compreenderam módulos que foram acrescentados progressivamente ao CX. Deste modo, foi possível analisar se os alunos que obtiveram as maiores notas foram realmente aqueles que mais utilizaram a ferramenta e observar quais funcionalidades da plataforma foram as mais empregadas por esses usuários. A organização integral das variáveis e níveis utilizados na análise correlacional pode ser vista na Figura 13.

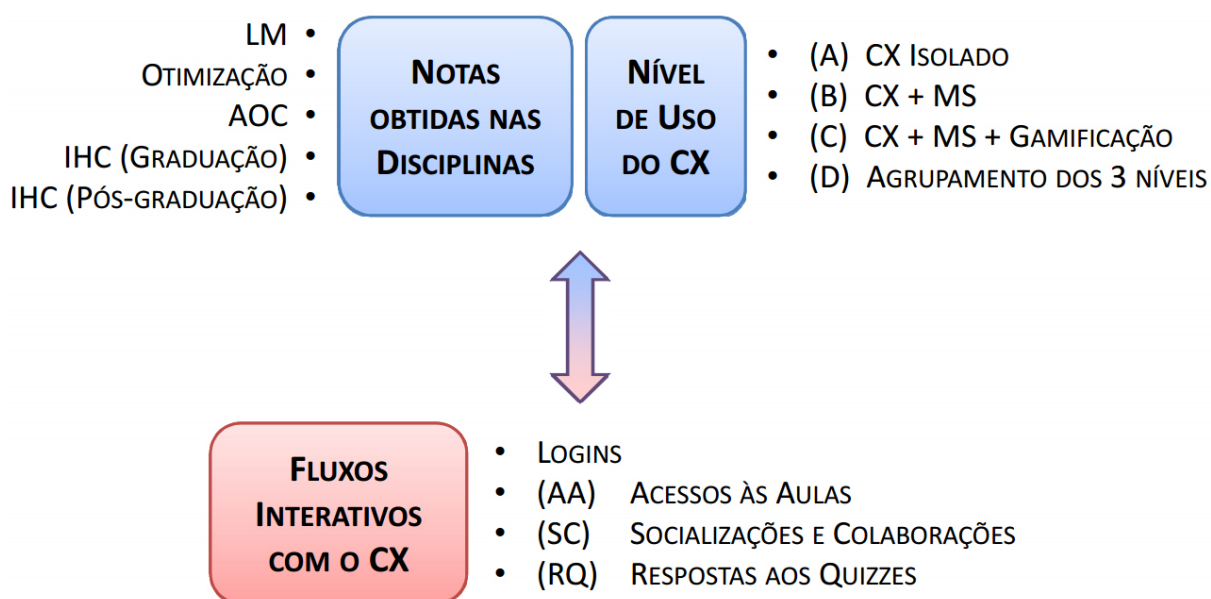


Figura 13 – Análise Correlacional entre o desempenho obtido pelos estudantes e seus respectivos níveis de interação com o CX.

6.2 Resultados e Discussões

A realização do teste de Shapiro-Wilk possibilitou identificar a existência de amostras normais e não-normais dentre as quinze turmas analisadas (Tabela 7). Ao todo, cinco amostras apresentaram resíduos que não seguiram a distribuição normal: duas turmas de Lógica Matemática (LM) que tiveram contato com o CX nos níveis de uso A e B, uma turma de Otimização que empregou o CX no nível A e duas turmas de Arquitetura e Organização de Computadores (AOC) que utilizaram o CX nos níveis A e C. Para as demais amostras analisadas, foram encontrados resíduos normais.

Tabela 7 – Análise da normalidade residual das quinze turmas que empregaram o Classroom eXperience (CX) durante os semestres letivos.

Disciplinas ¹	Níveis de uso do CX		
	A	B	C
LM	0,0921 (0,046)	0,0916 (0,024)	0,0989 (0,177)*
Otimização	0,0906 (<0,01)	0,0924 (0,098)*	0,0941 (0,223)*
AOC	0,0914 (0,021)	0,0941 (0,223)*	0,0933 (0,038)
IHC (Grad)	0,0978 (0,783)*	0,0934 (0,661)*	0,0976 (0,691)*
IHC (Pós)	0,0981 (0,792)*	0,0917 (0,648)*	0,0993 (0,795)*

¹ $W(P)$: Estatística e Probabilidade do teste de Shapiro-Wilk; *:Probabilidades maiores que 0,05 indicam resíduos com distribuição normal; A: Uso isolado do CX; B: Uso do CX junto ao Módulo Social; C: Uso do CX junto aos Módulos Social e de Gamificação; LM: Lógica Matemática; AOC: Arquitetura e Organização de Computadores; IHC: Interação Humano-Computador; Grad: Turmas de graduação; Pós: Turmas de pós-graduação.

Devido a isso, utilizou-se a Correlação de Spearman para averiguar se os estudantes que obtiveram as maiores notas das quinze turmas analisadas também foram aqueles que realizaram a maior quantidade de *logins* na plataforma durante os semestres letivos. Os resultados obtidos podem ser observados na Tabela 8.

Tabela 8 – Correlação de Spearman entre o desempenho acadêmico obtido pelos estudantes e seus respectivos *logins* na plataforma CX.

Uso do CX ¹	Disciplinas				
	LM	Otimização	AOC	IHC (Grad)	IHC (Pós)
A	0,187 <i>b</i>	0,228 <i>b</i>	0,074	0,068	0,035
B	0,241 <i>b</i>	0,193 <i>b</i>	0,232 <i>b</i>	0,204 <i>b</i>	0,069
C	0,326 <i>m</i>	0,447 <i>m</i>	0,269 <i>b</i>	0,127 <i>b</i>	0,287 <i>b</i>
D	0,072	0,095	0,051	0,024	0,016

¹ ρ : Coeficiente da Correlação de Spearman; *b*: Correlação baixa; *m*: Correlação moderada; LM: Lógica Matemática; AOC: Arquitetura e Organização de Computadores; IHC: Interação Humano-Computador; Grad: Turmas de graduação; Pós: Turmas de pós-graduação; A: Uso isolado do CX; B: Uso do CX junto ao Módulo Social; C: Uso do CX junto aos Módulos Social e de Gamificação; D: Agrupamento dos três níveis de uso do CX numa única amostra.

A partir da análise inicial da Tabela 8, é possível perceber que não houve nenhuma correlação para as disciplinas na circunstância em que as amostras tiveram as observações agrupadas (nível D). Isso indica que a relação entre as variáveis realmente possui efeitos mais concentrados e que os diferentes módulos acrescentados ao CX ocasionaram quantidades diversificadas de acessos à plataforma. Tal indício pode ser confirmado ao se analisar que em todas as turmas que fizeram uso do CX no nível C houve algum tipo de correlação, ao passo que entre as amostras que empregaram o sistema sem o acréscimo de módulos (nível A), apenas duas turmas apresentaram correlação baixa.

De modo mais específico, as turmas que cursaram LM apresentaram duas correlações baixas nos níveis A e B, já no nível C ocorreu uma correlação moderada. O mesmo aconteceu entre as turmas que cursaram Otimização, sendo relevante destacar que nesta disciplina a turma que empregou o CX no nível C apresentou o maior coeficiente de correlação dentre os demais. As turmas que cursaram as disciplinas AOC e IHC (Graduação) também apresentaram resultados semelhantes e ambas alcançaram duas correlações baixas nos níveis B e C. Para IHC (Pós-graduação), apenas uma correlação baixa foi obtida no nível C e, fora este resultado, os demais coeficientes alcançados nesta disciplina foram os mais baixos dentre todos os obtidos.

Por terem ocorrido circunstâncias de correlações baixas e moderadas entre o desempenho dos estudantes e a quantidade de *logins* realizados no CX, também foram investigadas as possíveis correlações entre as notas obtidas pelos alunos e as três variáveis restantes: Acesso às Aulas (AA), Socializações e Colaborações (SC) e Respostas aos Quizzes (RQ) (Figura 13). Os resultados obtidos a partir de todas as correlações realizadas podem ser observados na Tabela 9.

De modo a permitir um melhor entendimento da organização dos coeficientes de correlação apresentados na Tabela 9: para a variável AA, os resultados das quinze turmas testadas nos três níveis de uso do CX (A, B e C) se encontram no primeiro agrupamento de resultados. Para a variável SC, devido às turmas que utilizaram o CX no nível A não terem utilizado elementos de socialização e colaboração na ferramenta, não foi possível testar a correlação para este nível e, deste modo, são apresentados os resultados obtidos nos níveis B e C. Por fim, para a variável RQ, as turmas que utilizaram o CX nos níveis A e B não tiveram contato com elementos de Gamificação no sistema e, portanto, apenas os resultados obtidos no nível C foram apresentados.

Tabela 9 – Correlação de Spearman entre o desempenho acadêmico obtido por cada turma e as variáveis: Acesso às Aulas (AA), Socializações e Colaborações (SC) e Respostas aos Quizzes (RQ).

Variáveis ¹	Uso do CX	Disciplinas				
		LM	Otimização	AOC	IHC (Grad)	IHC (Pós)
AA	A	0,212 <i>b</i>	0,198 <i>b</i>	0,286 <i>b</i>	0,092	0,085
	B	0,163 <i>b</i>	0,173 <i>b</i>	0,255 <i>b</i>	0,085	0,037
	C	0,309 <i>m</i>	0,412 <i>m</i>	0,084	0,201 <i>b</i>	0,129 <i>b</i>
SC	A	-	-	-	-	-
	B	0,015	0,416 <i>m</i>	0,251 <i>b</i>	0,079	0,046
	C	0,194 <i>b</i>	0,353 <i>m</i>	0,066	0,214 <i>b</i>	0,135 <i>b</i>
RQ	A	-	-	-	-	-
	B	-	-	-	-	-
	C	0,393 <i>m</i>	<0,01	0,081	0,237 <i>b</i>	0,031

¹ ρ : Coeficiente da Correlação de Spearman; *b*: Correlação baixa; *m*: Correlação moderada; LM: Lógica Matemática; AOC: Arquitetura e Organização de Computadores; IHC: Interação Humano-Computador; Grad: Turmas de graduação; Pós: Turmas de pós-graduação; AA: Acesso às Aulas; SC: Socializações e Colaborações; RQ: Respostas aos Quizzes; A: Uso isolado do CX; B: Uso do CX junto ao Módulo Social; C: Uso do CX junto aos Módulos Social e de Gamificação.

A partir da análise da Tabela 9, é possível perceber que a disciplina Otimização apresentou um dos melhores resultados dentre os obtidos. Nela, ocorreram três correlações moderadas e duas baixas. Com relação à turma que utilizou o CX isoladamente na disciplina (nível A), ocorreu uma correlação baixa com a quantidade de acessos realizados às aulas capturadas (variável AA). A turma que empregou a ferramenta junto ao MS (nível B) obteve uma correlação baixa com o acesso às aulas (variável AA) e uma correlação moderada com os elementos de socialização e colaboração (variável SC). Já a turma que fez uso do CX junto a todos os módulos disponíveis (nível C) obteve correlações moderadas tanto com a quantidade de acessos realizados às aulas (variável AA) quanto com as funcionalidades sociais e colaborativas (variável SC).

Outro resultado relevante foi percebido na disciplina LM, na qual houve duas correlações moderadas e três baixas. A turma que empregou o CX de modo isolado (nível A)

apresentou uma correlação baixa com o acesso às aulas (variável AA), algo que também aconteceu com a turma que empregou o sistema utilizando o nível B de acesso nesta disciplina. A turma que utilizou o CX junto a todos os módulos disponíveis (nível C) obteve uma correlação baixa com as ferramentas de socialização e colaboração (variável SC) e também alcançou mais duas correlações moderadas: com a quantidade de acessos às aulas (variável AA) e as respostas dadas aos Quizzes (variável RQ).

Tais resultados mostram-se importantes devido às disciplinas LM e Otimização terem sido lecionadas de modo semelhante. Nelas, com exceção de poucas ocasiões durante os semestres letivos, os professores optaram por não fazer o *upload* prévio de *slides* com os conteúdos específicos antes das aulas. Por se tratarem de disciplinas teóricas e matemáticas de cursos da área computacional, os professores preferiram abrir slides em branco no início das aulas e repassarem todo o conteúdo normalmente, aplicando exercícios e fazendo suas correções utilizando o CX.

Isso permitiu, por exemplo, que os professores retornassem a *slides* com conteúdos e exercícios já apresentados nas aulas, o que gerou maior dinamismo a essas disciplinas – que normalmente eram lecionadas apenas com giz ou pincel sobre a lousa. Além disso, a ocorrência de diversas correlações moderadas em Otimização e LM permite o entendimento de que os usuários do CX que realizaram os maiores índices de interação com a plataforma foram também os estudantes que obtiveram as melhores notas em suas respectivas turmas, ou seja, o uso da plataforma durante os semestres letivos realmente causou alguma influência nas notas dos alunos e gerou sua elevação.

Ainda acerca da Tabela 9, as disciplinas AOC e IHC (Graduação) apresentaram resultados similares, tendo ocorrido três correlações baixas em cada uma. Em AOC, a turma que utilizou o CX isoladamente (nível A) obteve uma correlação baixa com a quantidade de acessos realizados às aulas capturadas (variável AA). Já a turma que empregou a plataforma junto ao MS (nível B) apresentou duas correlações baixas: uma com o acesso às aulas (variável AA) e outra com os elementos de socialização e colaboração do CX (variável SC). Com relação à disciplina IHC (Graduação), a turma que utilizou a plataforma com todos os módulos disponíveis (nível C) apresentou correlações baixas para as três circunstâncias testadas (variáveis AA, SC e RQ).

Por fim, a disciplina IHC (Pós-graduação) apresentou apenas duas correlações baixas a partir da turma que utilizou o CX com todos os refinamentos desenvolvidos (nível C): com a quantidade de acessos realizados às aulas (variável AA) e o uso da socialização

e da colaboração na plataforma (variável SC). Tais resultados ratificam as informações obtidas a partir do teste de Tukey (Tabela 4, pág. 70), em que foram percebidos aumentos significativos no desempenho dos alunos em todas as turmas de graduação analisadas, mas nada ocorreu com relação ao desempenho dos alunos da pós-graduação.

Conforme pontuado por Brant-Ribeiro, Biase e Cattelan (2015), a baixa ocorrência de correlações neste nível de ensino pode ser explicada devido aos estudantes de pós-graduação já possuírem suas próprias técnicas de estudo e serem naturalmente esforçados, o que faz com que seus desempenhos já sejam elevados com ou sem a utilização de plataformas de apoio ao ensino. Infere-se, portanto, que a utilização de propostas tecnológicas na pós-graduação acaba se mostrando apenas como mais uma possibilidade dentre as existentes, diluindo sua importância e eficácia de uso.

De todo modo, como ocorreram correlações baixas e moderadas em todas as disciplinas de graduação aqui analisadas, seja em menor escala – como em AOC e IHC (Graduação) – ou na maior parte dos casos – como em LM e Otimização –, é possível inferir que os estudantes de cursos de graduação que obtiveram os desempenhos mais elevados também foram aqueles que mais empregaram o CX no decorrer dos semestres letivos e se beneficiaram do uso da ferramenta para seus estudos. O uso de uma plataforma de apoio ao ensino, neste caso, evidencia-se como uma abordagem relevante para auxiliar os estudantes durante os processos de aprendizagem que ocorrem em cursos de graduação.

6.3 Considerações Finais

Esta etapa da pesquisa teve como objetivo realizar uma análise correlacional entre o desempenho acadêmico de estudantes de graduação em Sistemas de Informação e graduação e pós-graduação em Ciência da Computação e seus fluxos de interação com a plataforma de apoio ao ensino CX. Para esta finalidade, foram empregadas 378 observações referentes às notas obtidas pelos estudantes ao final dos semestres letivos e contabilizou-se a quantidade de interações que esses respectivos alunos tiveram com a ferramenta para as variáveis de uso: *Login*, Acesso às Aulas (AA), Socializações e Colaborações (SC) e Respostas aos Quizzes (RQ).

Inicialmente, as amostras compostas pelas notas dos estudantes foram submetidas ao teste de Shapiro-Wilk para verificar se os resíduos da variável observada seguiam a

distribuição normal. Das quinze turmas observadas, cinco não apresentaram normalidade e, em função disso, optou-se pela realização da Correlação de Spearman para se investigar a relação entre as variáveis utilizadas nesta análise.

O teste correlacional entre as notas dos alunos e a quantidade de *logins* realizados possibilitou a compreensão de que não houve correlações para as disciplinas na circunstância em que as amostras tiveram suas observações agrupadas. Tal agrupamento ocorreu para analisar se a relação entre o desempenho e o *login* no CX mostrava-se como uma característica concentrada e específica de cada turma ou se ocorreria de modo espalhado entre as diversas turmas de uma mesma disciplina.

Como não foram alcançadas correlações entre as amostras agrupadas, presumiu-se que o nível de uso do CX apresentava alguma influência sobre a maneira com que os usuários interagem com a plataforma. Tal indício foi confirmado a partir da análise de que todas as turmas que empregaram o sistema com todos os módulos habilitados apresentaram algum nível de correlação, ao passo que as amostras que utilizaram o CX de modo isolado não alcançaram resultados similares.

Por causa disso, realizou-se também uma Correlação de Spearman entre o desempenho acadêmico dos estudantes e as variáveis AA, SC e RQ. Nesta circunstância, para todas as turmas de graduação investigadas, ocorreram correlações baixas e moderadas. As disciplinas Otimização e LM apresentaram os melhores resultados dentre os demais, visto que juntas alcançaram cinco correlações moderadas e cinco baixas. Acredita-se que isso tenha ocorrido devido aos professores dessas disciplinas terem empregado o CX de modo cotidiano em suas aulas, substituindo o quadro negro/branco pela lousa digital do Ambiente Educacional Ubíquo (AEU).

Em LM e Otimização, ocorreram correlações moderadas principalmente acerca do uso dos elementos de socialização e colaboração e, também, com relação à utilização dos Quizzes cadastrados pelos professores. Isso permite inferir que os usuários do CX que realizaram os maiores índices de interação com a plataforma também foram os estudantes que obtiveram as melhores notas em suas turmas. Deste modo, o emprego da plataforma realmente influenciou de modo positivo as notas dos alunos, elevando-as.

As disciplinas AOC e IHC (Graduação) apresentaram resultados similares entre si, tendo ocorrido três correlações baixas em cada uma. Os resultados obtidos nesses casos também são relevantes, visto que ocorreram correlações espalhadas entre as turmas. No entanto, acredita-se que o uso do CX se mostrou mais diluído neste caso, seja pelo modo

com que os professores utilizaram a ferramenta (fazendo o *upload* prévio de conteúdos diversos antes das aulas) ou pela natureza das disciplinas (que não se mostram tão teóricas e matemáticas como LM e Otimização). Como ocorreram correlações baixas e moderadas em todas as disciplinas de graduação analisadas, infere-se que o uso de uma ferramenta educacional se mostra válido para esta circunstância. Alunos de cursos de graduação tendem a ser mais receptivos para empregarem ferramentas educacionais e as utilizarem no cotidiano acadêmico.

Para a disciplina IHC (Pós-graduação), ocorreram apenas duas correlações baixas com a turma que utilizou o CX com todos os refinamentos desenvolvidos. Tal resultado gera indícios de que o uso de plataformas de apoio ao ensino não se mostra como uma proposta muito relevante para alunos de pós-graduação. Acredita-se que os estudantes de disciplinas de pós-graduação não se mostram diretamente influenciados pelo uso de tais ferramentas pois já demonstram certa maturidade e possuem rotinas de estudo que elevam naturalmente seus desempenhos.

De todo modo, das cinco disciplinas cursadas pelas quinze turmas utilizadas nesta análise, apenas uma delas foi cursada por estudantes em nível de pós-graduação. Os resultados aqui obtidos refletem especificamente o comportamento que os estudantes tiveram na disciplina de IHC, não sendo possível generalizá-los para todas as disciplinas existentes em cursos de pós-graduação em Ciência da Computação. Enxerga-se, portanto, a necessidade de se realizar uma análise englobando mais turmas de pós-graduandos, de modo a ser possível analisar adequadamente o impacto que AEU's (assim como outras plataformas de apoio ao ensino) causam no desempenho de turmas de pós-graduação em ambientes acadêmicos reais.

Conclusão

De modo a complementar as discussões realizadas junto aos resultados obtidos nesta tese, este capítulo apresenta conclusões e considerações finais referentes às atividades que foram desenvolvidas. Inicialmente, resultados e contribuições são apresentados na Seção 7.1; a produção bibliográfica resultante é listada na Seção 7.2 e, por fim, limitações encontradas e trabalhos futuros são apresentadas na Seção 7.3.

7.1 Resultados e Contribuições

Com o intuito de apresentar novas contribuições abordando aspectos de robustez e confiabilidade experimentais para a linha de pesquisa em *Learning Analytics* (LA), a presente tese almejou inicialmente desenvolver um módulo de Gerenciamento de Experimentos (Capítulo 3, pág. 45) que realizou a organização da visibilidade das funcionalidades de um AEU intitulado Classroom eXperience (CX) e forneceu subsídios para o planejamento e realização deste trabalho. Deste modo, empregou-se um método adequado para se analisar o impacto progressivo que o CX ocasionou no desempenho acadêmico de estudantes que cursaram disciplinas em cursos graduação e pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia (FACOM/UFU) no decorrer dos semestres letivos. A utilização do gerenciador se mostrou como uma prática vantajosa para a análise que foi realizada nesta pesquisa e o emprego deste módulo também pode ser expandido para outras investigações que demandem organizações detalhadas do nível

de acesso que grupos distintos de indivíduos devem ter ao utilizarem ferramentas de apoio ao ensino no cotidiano acadêmico.

Também almejou-se determinar e explorar um tamanho ótimo de amostra para a análise do desempenho de estudantes em contato com ferramentas de apoio ao ensino, tais como Ambientes Educacionais Ubíquos (AEUs) e Ambientes Virtuais de Aprendizagem (AVAs) (Capítulo 4, pág. 55). No decorrer desta pesquisa, percebeu-se também a necessidade de se determinar um tamanho ótimo amostral para analisar alunos de graduação e pós-graduação que não estão em contato com tecnologias educacionais, com o intuito de se traçar um paralelo entre os resultados obtidos e se recomendar uma estimativa capaz de compreender ambas as circunstâncias.

Determinou-se, portanto, um tamanho ótimo amostral de 25 alunos para a análise do desempenho acadêmico em turmas que não utilizam tecnologias educacionais no cotidiano e de 20 estudantes para turmas que possuem contato com tecnologias educacionais durante os semestres letivos. Percebeu-se que a diminuição do Coeficiente de Variação (CV) das amostras não apresentou um comportamento proporcional à elevação da quantidade de observações, havendo uma redução drástica na variabilidade quando as amostras adquiriram tamanhos aceitáveis, mas que tendeu à estabilidade ao passo que elas tiveram quantidades maiores de observações acrescidas. Em função disso, foi possível concluir que a elevação do tamanho amostral se manteve conveniente apenas até determinado momento, a partir do qual o emprego de quantidades maiores de observações não foi compensado devido ao ganho reduzido em precisão experimental.

Esses resultados reforçam a pressuposição de que quantidades de observações menores que as originalmente existentes nas amostras analisadas já teriam sido suficientes para se compreender o comportamento da variável estudada. Deste modo, tais estimativas mostram-se relevantes para pesquisas em LA que realizam estudos acerca do desempenho discente em circunstâncias nas quais novas tecnologias são introduzidas, de modo que o emprego destes tamanhos amostrais consiga garantir uma maior precisão e confiabilidade experimentais aos resultados obtidos.

Empregando-se o Gerenciador de Experimentos e o Tamanho Ótimo Amostral, foram realizados dois estudos: uma Análise de Variância Fatorial 4x5 (*Two-way ANOVA*) (Capítulo 5, pág. 65) e uma análise empregando a Correlação de Spearman (Capítulo 6, pág. 77). Inicialmente, a ANOVA Fatorial empregou 501 observações referentes aos desempenhos obtidos pelos estudantes ao final dos semestres letivos para observar se o acréscimo

gradativo de novas funcionalidades ao CX trouxe aumentos significativos em suas notas.

A partir desta análise, percebeu-se que o desempenho dos alunos foi diretamente influenciado pela interação entre as funcionalidades habilitadas no CX e a disciplina cursada pelos estudantes. Tal resultado evidencia que apenas o uso de uma ferramenta de apoio ao ensino não deve ser caracterizado como o único fator responsável por elevações significativas no desempenho de estudantes sob análise. É preciso, também, observar a natureza das disciplinas que estão sob investigação, visto que disciplinas que fazem parte de eixos de formação diferentes possuem seus próprios fluxos de apresentação dos conteúdos, compreensão das temáticas e avaliação do processo de aprendizagem.

De todo modo, num comparativo entre as amostras que não utilizaram o CX e aquelas que empregaram a plataforma com o Módulo Social (MS) e Gamificação, houve elevações significativas no desempenho discente em todas as turmas de graduação investigadas. Na disciplina de pós-graduação analisada, não se observou diferença significativa no desempenho discente entre os níveis de uso do CX, embora também tenham havido acréscimos gradativos no desempenho obtido pelos estudantes desta faixa.

O estudo que empregou a Correlação de Spearman foi realizado para se averiguar a relação entre o desempenho acadêmico dos estudantes e seus fluxos de interação com a plataforma CX. Para isso, foram analisadas 378 observações referentes às notas dos estudantes e a quantidades de *logins* que cada aluno realizou no sistema. A partir deste teste, percebeu-se que o nível de utilização do CX apresentou influência sobre a maneira com que os usuários interagiram com a plataforma. Tal indício foi confirmado a partir da análise de que todas as turmas que empregaram o sistema com todos os módulos habilitados apresentaram algum nível de correlação, ao passo que as amostras que utilizaram o CX de modo isolado não alcançaram resultados similares.

Devido a isso, realizou-se também uma Correlação de Spearman entre o desempenho acadêmico dos estudantes e variáveis que compreenderam os diferentes níveis de visibilidade e uso do CX: Acesso às Aulas (AA), Socializações e Colaborações (SC) e Respostas aos Quizzes (RQ). Nesta circunstância, para todas as turmas de graduação investigadas ocorreram correlações baixas e moderadas. As disciplinas teóricas e matemáticas apresentaram os melhores resultados dentre os demais e acredita-se que isso tenha ocorrido devido aos professores dessas disciplinas terem empregado o CX de modo cotidiano em suas aulas, substituindo o quadro negro/branco pela lousa digital do AEU.

A disciplina de pós-graduação investigada apresentou poucas correlações baixas, o que gerou mais indícios de que o uso de plataformas de apoio ao ensino não se mostra como uma proposta muito relevante para alunos de pós-graduação. Concluiu-se, portanto, que o emprego de novas propostas educacionais de apoio ao ensino acaba se evidenciando como uma opção mais trivial para estudantes de pós-graduação, visto que eles possuem seus próprios métodos de estudo e aprendizagem, além de tenderem a ser mais esforçados e obterem níveis de desempenho naturalmente elevados. Para turmas de graduação, a utilização destas tecnologias se mantém como um tópico de alta relevância para pesquisas científicas, visto que foram observados benefícios significativos para todas as disciplinas analisadas neste nível de ensino.

7.2 Produção Bibliográfica

Produção relacionada ao Gerenciador de Experimentos (Capítulo 3):

- ❑ BRANT-RIBEIRO, Taffarel; ARAÚJO, Rafael D.; MENDONÇA, Igor E.; SOARES, Michel S.; CATTELAN, Renan G. **A User-centered Approach for Modeling Web Interactions Using Colored Petri Nets**. Trabalho completo publicado nos Anais da 17th International Conference on Enterprise Information Systems. p. 37-48, 2015.
- ❑ BRANT-RIBEIRO, Taffarel; ARAÚJO, Rafael D.; MENDONÇA, Igor E.; SOARES, Michel S.; CATTELAN, Renan G. **Interactive Web Interfaces Modeling, Simulation and Analysis Using Colored Petri Nets**. Software & Systems Modeling, v.1, p. 1-17, 2017.

Produção relacionada ao cálculo de tamanhos ótimos amostrais (Capítulo 4):

- ❑ BRANT-RIBEIRO, Taffarel; CATTELAN, Renan G. **Tamanho Ótimo de Amostra para Análise do Desempenho de Estudantes em Ambientes Educacionais Ubíquos**. Trabalho completo publicado nos Anais do XXVI Simpósio Bra-

sileiro de Informática na Educação. p. 31-40, 2015. **Trabalho premiado como *best-paper*.**

- ❑ BRANT-RIBEIRO, Taffarel; CATTELAN, Renan G. **Determinação de um Tamanho Ótimo Amostral para a Análise do Desempenho de Estudantes em Ambientes Educacionais Ubíquos.** Revista Brasileira de Informática na Educação, v. 24, n. 3, p. 86-96, 2016.

Produção relacionada à análise do impacto da plataforma computacional de apoio ao ensino (Capítulos 5 e 6):

- ❑ BRANT-RIBEIRO, Taffarel; MENDONÇA, Igor E.; ARAÚJO, Rafael D.; MENDES, Miller M.; DORÇA, Fabiano A.; CATTELAN, Renan G. **Um Modelo Social e Colaborativo para Extensão de Conteúdo em Ambientes Educacionais Ubíquos.** Tecnologias, Sociedade e Conhecimento, v. 2, p. 105-130, 2014.
- ❑ BRANT-RIBEIRO, Taffarel; BIASE, Nádia G.; CATTELAN, Renan G. **Impacto de Aplicações de Captura e Acesso em Estudantes de Cursos de Computação.** Revista Brasileira de Informática na Educação, v. 23, p. 111-126, 2015.
- ❑ ARAÚJO, Rafael D.; BRANT-RIBEIRO, Taffarel; MENDONÇA, Igor E.; MENDES, Miller M.; DORÇA, Fabiano A.; CATTELAN, Renan G. **Social and Collaborative Interactions for Educational Content Enrichment in ULEs.** Educational Technology & Society, v. 20, p. 133-144, 2017.
- ❑ BRANT-RIBEIRO, Taffarel; CATTELAN, Renan G. **Análise de Variância Fatorial do Desempenho de Estudantes sob a Influência de Aprimoramentos Progressivos em Plataformas de Apoio ao Ensino.** Trabalho completo publicado nos Anais do XXIX Simpósio Brasileiro de Informática na Educação. p. 1233-1242, 2018.

Outras publicações que possuem relação com esta tese:

- ❑ MENDONÇA, Igor E.; ARAÚJO, Rafael D.; MENDES, Miller M.; BRANT-RIBEIRO, Taffarel; DORÇA, Fabiano A.; CATTELAN, Renan G. **Explorando Funcionalidades Sociais e Colaborativas em Ambientes Educacionais Ubíquos**. Trabalho completo publicado nos Anais do XXV Simpósio Brasileiro de Informática na Educação. p. 486-495, 2014.
- ❑ ARAÚJO, Rafael D.; BRANT-RIBEIRO, Taffarel; FREITAS, Rafaela S.; DORÇA, Fabiano A.; CATTELAN, Renan G. **Autoria automática de objetos de aprendizagem a partir de captura multimídia e associação a estilos de aprendizagem**. Trabalho completo publicado nos Anais do XXV Simpósio Brasileiro de Informática na Educação. p. 229-238, 2014.
- ❑ FERREIRA, Hiran N. M.; BRANT-RIBEIRO, Taffarel; ARAÚJO, Rafael D.; DORÇA, Fabiano A.; CATTELAN, Renan G. **An Automatic and Dynamic Student Modeling Approach for Adaptive and Intelligent Educational Systems Using Ontologies and Bayesian Networks**. Trabalho completo publicado nos Anais da IEEE 28th International Conference on Tools with Artificial Intelligence. p. 738-745, 2016.
- ❑ ARAÚJO, Rafael D.; BRANT-RIBEIRO, Taffarel; FERREIRA, Hiran N. M.; DORÇA, Fabiano A.; CATTELAN, Renan G. **Segmentação Colaborativa de Objetos de Aprendizagem Utilizando Bookmarks em Ambientes Educacionais Ubíquos**. Trabalho completo publicado nos Anais do XXVII Simpósio Brasileiro de Informática na Educação. p. 1205-1214, 2016.
- ❑ FERREIRA, Hiran N. M.; BRANT-RIBEIRO, Taffarel; ARAÚJO, Rafael D.; DORÇA, Fabiano A.; CATTELAN, Renan G. **An Automatic and Dynamic Knowledge Assessment Module for Adaptive Educational Systems**. Trabalho completo publicado nos Anais da IEEE 17th International Conference on Advanced Learning Technologies. p. 517-521, 2017.

- ARAÚJO, Rafael D.; BRANT-RIBEIRO, Taffarel; FERREIRA, Hiran N. M.; DORÇA, Fabiano A.; CATTELAN, Renan G. **A Semiautomatic and Probabilistic Approach for Student Modeling in Ubiquitous Learning Environments**. Trabalho completo publicado nos Anais do XXIX Simpósio Brasileiro de Informática na Educação. p. 1313-1322, 2018. **Trabalho premiado como *best-paper***.

7.3 Limitações e Trabalhos Futuros

Durante a realização desta tese, um fator limitante observado referiu-se à variável que foi investigada para a estimativa de tamanho ótimo amostral. Os cálculos realizados nesta pesquisa permitiram a determinação de tamanhos amostrais confiáveis para a realização de novas pesquisas que envolvam especificamente as notas de estudantes de graduação ou pós-graduação. Para investigações que empreguem outras variáveis, tais como assiduidade ou, até mesmo, o desempenho de estudantes em outros níveis de ensino, mostra-se necessária a realização de novos procedimentos experimentais para se calcular tamanhos ótimos em cada uma dessas circunstâncias, algo que não foi contemplado nesta pesquisa.

Além disso, o emprego do desempenho discente para a determinação de um tamanho ótimo amostral pode gerar questionamentos de natureza pedagógica, visto que se trata de um fator que não necessariamente representa com fidelidade o nível de conhecimento dos estudantes. No entanto, a nota obtida por um aluno se mantém como o elemento-chave que o sistema educacional brasileiro emprega para avaliar aprovações e reprovações. Acredita-se que encontrar maneiras de se analisar com alto índice de confiabilidade os fluxos de dados que se enquadram nesta categoria ainda se mostra como um assunto relevante para que o debate acerca dos critérios de avaliação estudantis se mantenha em foco e o amadurecimento sobre esta temática continue a ocorrer.

Deste modo, estudos que investigam a heterogeneidade de variáveis e o cálculo de tamanhos amostrais mantêm-se como tópicos relevantes para pesquisas que fazem uso de fatores humanos, pois demonstram que amostras com quantidades reduzidas de observações são capazes de compreender a maior parte do comportamento das variáveis sob análise. O uso de tamanhos amostrais ainda maiores que os estabelecidos como ótimos também podem garantir níveis de precisão experimental mais robustos, no entanto o ga-

nho de confiabilidade em amostras com quantidades de observações acima das ótimas pode se mostrar como algo reduzido ou, até mesmo, insignificante. Devido aos custos e obstáculos existentes para se trabalhar com um elevado número de indivíduos em contextos diversificados, o cálculo de um tamanho ótimo amostral se revela como uma alternativa robusta e válida para se obter alto índice de confiabilidade em inferências obtidas a partir de análises em contextos mais reduzidos.

Outra limitação observada referiu-se às disciplinas que foram avaliadas empregando-se tanto a ANOVA Fatorial quanto a Correlação de Spearman nesta tese. Das cinco disciplinas cursadas por todas as turmas utilizadas, apenas uma delas foi cursada por estudantes de pós-graduação. Os resultados obtidos e as discussões feitas com os dados das turmas de pós-graduação refletem, portanto, o comportamento que os alunos tiveram especificamente na disciplina de Interação Humano-Computador (IHC).

Embora esta característica não afete a confiabilidade do tamanho ótimo amostral calculado (visto que, para esta estimativa, turmas de diversas disciplinas de pós-graduação foram utilizadas), não é possível generalizar os resultados aqui obtidos para todas as disciplinas existentes em cursos de pós-graduação em Ciência da Computação. Enxerga-se, portanto, a necessidade de se realizar novas análises englobando mais turmas com estudantes pós-graduandos, de modo a ser possível investigar adequadamente o impacto que plataformas de apoio ao ensino causam no desempenho de turmas de pós-graduação em ambientes acadêmicos reais.

Como trabalhos futuros, enxerga-se como relevante o aumento do escopo de investigação desta pesquisa, de modo a compreender dados referentes a todos os semestres letivos de cursos de graduação em Sistemas de Informação e Ciência da Computação – e, também, de outros cursos da área computacional. Dessa maneira, será possível discernir quais disciplinas e momentos (períodos) dos cursos tendem a apresentar maiores níveis de heterogeneidade no desempenho dos estudantes, além de ser possível, também, compreender o processo de amadurecimento dos estudantes no decorrer do curso e como essa variável pode influenciar no desempenho obtido nas disciplinas.

Tais informações se mostram válidas para a identificação prévia de circunstâncias que apresentam maiores riscos de reprovação dos alunos e, conseqüentemente, podem levar à desistência em disciplinas matriculadas e, até mesmo, à evasão dos cursos. Acredita-se, portanto, que a compreensão de circunstâncias neste escopo permitirá a criação e o estudo

de técnicas e propostas para a prevenção dessas situações, tais como o desenvolvimento e o refinamento de ferramentas educacionais com o intuito de melhorar os processos de ensino/aprendizagem.

Referências

- ABOWD, G. D.; MYNATT, E. D. Charting Past, Present, and Future Research in Ubiquitous Computing. **ACM Transactions on Computer-Human Interaction (TOCHI)**, ACM, v. 7, n. 1, p. 29–58, 2000. <https://doi.org/10.1145/344949.344988>
- AIHUA, Z. Study of Ubiquitous Learning Environment Based on Ubiquitous Computing. In: **Proceedings of the 3rd IEEE International Conference on Ubi-Media Computing**. Jinhua, CN: IEEE. p. 136–138, 2010. <https://doi.org/10.1109/UMEDIA.2010.5544482>
- ARAÚJO, R. D. et al. Personalization of Interactive Digital Media in Ubiquitous Educational Environments. In: **Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics**. Manchester, UK: IEEE, p. 3955–3960, 2013. <https://doi.org/10.1109/SMC.2013.675>
- ASARTA, C. J.; SCHMIDT, J. R. Comparing student performance in blended and traditional courses: Does prior academic achievement matter? **The Internet and Higher Education**, Elsevier, v. 32, p. 29–38, 2017. <https://doi.org/10.1016/j.iheduc.2016.08.002>
- BARNEVELD, A. V.; ARNOLD, K. E.; CAMPBELL, J. P. Analytics in Higher Education: Establishing a Common Language. **EDUCAUSE Learning Initiative**, v. 1, p. 1–11, 2012.
- BARROS, I. d.; TAVARES, M. Estimativa do Tamanho Ótimo de Parcelas Experimentais Através de Cálculos Algébricos. **Bragantia**, v. 54, n. 1, p. 209–215, 1995. <https://dx.doi.org/10.1590/S0006-87051995000100024>
- BODILY, R. et al. Open learner models and learning analytics dashboards: a systematic review. In: ACM. **Proceedings of the 8th International Learning Analytics & Knowledge Conference**. p. 1–10, 2018. <https://doi.org/10.1145/3170358.3170409>

- BOX, G. E. P. Non-normality and tests on variances. **Biometrika**, Oxford University Press, Biometrika Trust, v. 40, n. 3/4, p. 318–335, 1953. <https://dx.doi.org/10.2307/2333350>
- BRANT-RIBEIRO, T.; BIASE, N. G.; CATTELAN, R. G. Impacto de Aplicações de Captura e Acesso em Estudantes de Cursos de Computação. **Revista Brasileira de Informática na Educação**, v. 23, n. 2, p. 111–126, 2015. <https://dx.doi.org/10.5753/rbie.2015.23.02.111>
- BRANT-RIBEIRO, T.; CATTELAN, R. Tamanho Ótimo de Amostra para Análise do Desempenho de Estudantes em Ambientes Educacionais Ubíquos. In: **Anais do Simpósio Brasileiro de Informática na Educação**. Maceió, BR: SBIE, v. 26, n. 1, p. 31–40, 2015. <https://dx.doi.org/10.5753/cbie.sbie.2015.31>
- BRANT-RIBEIRO, T.; CATTELAN, R. G. Determinação de um Tamanho Ótimo Amostral para a Análise do Desempenho de Estudantes em Ambientes Educacionais Ubíquos. **Revista Brasileira de Informática na Educação**, v. 24, n. 3, p. 86–96, 2016. <https://dx.doi.org/10.5753/rbie.2016.24.3.86>
- BRANT-RIBEIRO, T.; CATTELAN, R. Análise de Variância Fatorial do Desempenho de Estudantes sob a Influência de Aprimoramentos Progressivos em Plataformas de Apoio ao Ensino. In: **Anais do Simpósio Brasileiro de Informática na Educação**. Fortaleza, BR: SBIE, v. 1, n. 1, p. 1233–1242, 2018. <https://dx.doi.org/10.5753/cbie.sbie.2018.1233>
- BRANT-RIBEIRO, T. et al. Um Modelo Social e Colaborativo para Extensão de Conteúdo em Ambientes Educacionais Ubíquos. **Tecnologias, Sociedade e Conhecimento**, v. 2, n. 1, p. 105–130, 2015.
- CAI, S. Case studies of augmented reality applications for authentic learning. In: **Authentic Learning Through Advances in Technologies**. Springer, p. 115–134, 2018. https://doi.org/10.1007/978-981-10-5930-8_8
- CHATTI, M. A. et al. A Reference Model for Learning Analytics. **International Journal of Technology Enhanced Learning**, Inderscience Publishers, Geneva, SUI, v. 4, n. 5/6, p. 318–331, 2012. <https://doi.org/10.1504/IJTEL.2012.051815>
- CHATTI, M. A. et al. Learning Analytics: Challenges and Future Research Directions. **E-Learning & Education**, v. 10, n. 1, 2014.
- CONFALONIERI, R. et al. Resampling-based Software for Estimating Optimal Sample Size. **Environmental Modelling & Software**, Elsevier, v. 22, n. 12, p. 1796–1800, 2007. <https://doi.org/10.1016/j.envsoft.2007.02.006>

- COOK, D. J.; DAS, S. K. Review: Pervasive Computing at Scale: Transforming the State of the Art. **Pervasive and Mobile Computing**, Elsevier Science Publishers B. V., Amsterdam, NL, v. 8, n. 1, p. 22–35, 2012. <https://doi.org/10.1016/j.pmcj.2011.10.004>
- COOPER, M.; FERGUSON, R.; WOLFF, A. What can analytics contribute to accessibility in e-learning systems and to disabled students' learning? In: **Proceedings of the 6th International Conference on Learning Analytics & Knowledge**. NY, USA: ACM, p. 99–103, 2016. <https://doi.org/10.1145/2883851.2883946>
- CUI, Y.-J. et al. A Study of Sample Size with Neural Network. In: IEEE. **Proceedings of 2004 International Conference on Machine Learning and Cybernetics**. v. 6, p. 3444–3448, 2004. <https://doi.org/10.1109/ICMLC.2004.1380382>
- DANIELSON, J. et al. Is the Effectiveness of Lecture Capture Related to Teaching Approach or Content Type? **Computers & Education**, v. 72, p. 121–131, 2014. <https://doi.org/10.1016/j.compedu.2013.10.016>
- DAVIS, J. A. **Elementary survey analysis**. Englewood: Prentice-Hall, 1971. <https://doi.org/10.1093/sf/51.3.376>
- DAWSON, S. et al. Current State and Future Trends: A Citation Network Analysis of the Learning Analytics Field. In: **Proceedings of the 4th International Conference on Learning Analytics And Knowledge**. ACM, p. 231–240, 2014. <https://doi.org/10.1145/2567574.2567585>
- DAWSON, S. et al. From prediction to impact: Evaluation of a learning analytics retention program. In: **Proceedings of the 7th International Learning Analytics & Knowledge Conference**. NY, USA: ACM, p. 474–478, 2017. <https://doi.org/10.1145/3027385.3027405>
- DICKSON, P. E. et al. Student Reactions to Classroom Lecture Capture. In: **Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education**. ACM, p. 144–149, 2012. <https://doi.org/10.1145/2325296.2325334>
- DURÁN, E. B.; ÁLVAREZ, M. M.; ÚNZAGA, S. I. Ontological model-driven architecture for ubiquitous learning applications. In: **Proceedings of the 7th Euro American Conference on Telematics and Information Systems**. NY, USA: ACM, p. 1–6, 2014. <https://doi.org/10.1145/2590651.2590776>
- EFRON, B. Bootstrap Methods: Another Look at the Jackknife. **The Annals of Statistics**, JSTOR, v. 7, n. 1, p. 1–26, 1979. <https://doi.org/10.1214/aos/1176344552>

- EU, J. H. A Sampling Approach to Real-Time Performance Monitoring of Digital Transmission Systems. In: IEEE. **8th Annual International Phoenix Conference on Computers and Communications**. p. 207–211, 1989. <https://doi.org/10.1109/PCCC.1989.37388>
- EUZENT, P. J. et al. **Teaching Principles to the Masses: Assessing Student Performance in Lecture Capture vs. Face-to-Face Course Delivery**. 2011. URL: <<https://ssrn.com/abstract=1868945>>. Acesso em Novembro/2018.
- FERGUSON, R. Learning Analytics: Drivers, Developments and Challenges. **International Journal of Technology Enhanced Learning**, Inderscience, v. 4, n. 5, p. 304–317, 2012. <https://doi.org/10.1504/IJTEL.2012.051816>
- FERREIRA, H. N. M. et al. Classroom Experience: A Platform for Multimedia Capture and Access in Instrumented Educational Environments. In: **Proceedings of the 2012 Brazilian Symposium on Collaborative Systems**. IEEE Computer Society, p. 59–64, 2012. <https://doi.org/10.1109/SBSC.2012.20>
- FIELD, A. **Discovering Statistics Using IBM SPSS Statistics**. 4^a. ed. Thousand Oaks, CA: Sage Publications Ltd., 2013.
- GASEVIC, D. et al. Effects of instructional conditions and experience on the adoption of a learning tool. **Computers in Human Behavior**, v. 67, p. 207–220, 2017. <https://doi.org/10.1016/j.chb.2016.10.026>
- GEISSER, S.; JOHNSON, W. O. **Modes of parametric statistical inference**. John Wiley & Sons, 2006. <https://doi.org/10.1002/0471743135>
- GELAN, A. et al. Affordances and limitations of learning analytics for computer-assisted language learning: a case study of the vital project. **Computer Assisted Language Learning**, Routledge, v. 31, n. 3, p. 294–319, 2018. <https://doi.org/10.1080/09588221.2017.1418382>
- GOMEZ, K. A.; GOMEZ, A. A. **Statistical Procedures for Agricultural Research**. John Wiley & Sons, 1984. <https://doi.org/10.1017/S0014479700014496>
- GOOD, P. **Permutation tests: a practical guide to resampling methods for testing hypotheses**. Springer Science & Business Media, 2013. <https://doi.org/10.1007/978-1-4757-2346-5>
- HAUKE, J.; KOSSOWSKI, T. Comparison of values of pearson's and spearman's correlation coefficients on the same sets of data. **Quaestiones geographicae**, v. 30, n. 2, p. 87–93, 2011. <https://doi.org/10.2478/v10117-011-0021-1>

- HWANG, W.; SALVENDY, G. Number of People Required for Usability Evaluation: the 10 ± 2 Rule. **Communications of the ACM**, ACM, v. 53, n. 5, p. 130–133, 2010. <https://doi.org/10.1145/1735223.1735255>
- INTERNET WORLD STATS. **World Internet User Statistics and World Population Stats**. 2018. URL: <<https://www.internetworldstats.com/stats.htm>>. Acesso em Novembro/2018.
- JEREMIC, Z.; KUMAR, V.; GRAF, S. Morph: Supporting the integration of learning analytics at institutional level. In: **Proceedings of the 7th International Learning Analytics & Knowledge Conference**. NY, USA: ACM, p. 596–597, 2017. <https://doi.org/10.1145/3027385.3029478>
- KIENTZ, J. A. Embedded capture and access: Encouraging recording and reviewing of data in the caregiving domain. **Personal and Ubiquitous Computing**, Springer-Verlag, London, UK, v. 16, n. 2, p. 209–221, 2012. <https://doi.org/10.1007/s00779-011-0380-6>
- KINSHUK; GRAF, S. Ubiquitous Learning. In: **Encyclopedia of the Sciences of Learning**. Springer USA, p. 3361–3363, 2012. https://doi.org/10.1007/978-1-4419-1428-6_224
- KNIGHT, D. B. et al. Developing a learning analytics dashboard for undergraduate engineering using participatory design. In: **2015 ASEE Annual Conference & Exposition**. Seattle, Washington: ASEE Conferences, p. 1–11, 2015. <https://doi.org/10.18260/p.23824>
- KORTEMAYER, G. The Spectrum of Learning Analytics. **E-Learning & Education**, v. 12, n. 1, 2017.
- LAST, M. Improving Data Mining Utility with Projective Sampling. In: **ACM. Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining**. p. 487–496, 2009. <https://doi.org/10.1145/1557019.1557076>
- LESSMAN, K. J.; ATKINS, R. E. Optimum Plot Size and Relative Efficiency of Lattice Designs for Grain Sorghum Yield Tests. **Crop Science**, Crop Science Society of America, v. 3, n. 6, p. 477–481, 1963. <https://doi.org/10.2135/cropsci1963.0011183X000300060006x>
- MARKOWSKI, C. A.; MARKOWSKI, E. P. Conditions for the effectiveness of a preliminary test of variance. **The American Statistician**, American Statistical Association, Taylor & Francis, Ltd., v. 44, n. 4, p. 322–326, 1990. <https://doi.org/10.2307/2684360>

- MEIER, V. D.; LESSMAN, K. J. Estimation of Optimum Field Plot Shape and Size for Testing Yield in *Crambe abyssinica* Hochst. **Crop Science**, Crop Science Society of America, v. 11, n. 5, p. 648–650, 1971.
<https://doi.org/10.2135/cropsci1971.0011183X001100050013x>
- MERTENS, D. M. **Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods**. SAGE Publications, 2014.
- MOURI, K.; OGATA, H.; UOSAKI, N. Analysis of ubiquitous learning logs using social network analysis. **International Journal of Mobile Learning and Organisation**, Inderscience Publishers, Geneva, SUI, v. 9, n. 2, p. 101–123, 2015.
<https://doi.org/10.1504/IJMLO.2015.070702>
- MOURI, K.; OGATA, H.; UOSAKI, N. Learning analytics in a seamless learning environment. In: **Proceedings of the 7th International Learning Analytics & Knowledge Conference**. NY, USA: ACM, p. 348–357, 2017.
<https://doi.org/10.1145/3027385.3027408>
- MÜLLER, P. et al. Optimal Sample Size for Multiple Testing: the Case of Gene Expression Microarrays. **Journal of the American Statistical Association**, Taylor & Francis, v. 99, n. 468, p. 990–1001, 2004. <https://doi.org/10.1198/016214504000001646>
- NBR, A. 5891: Regras de Arredondamento na Numeração Decimal. **NBR 4: Norma Brasileira Probatória**, 1977.
- NIELSEN, J. **Why You Only Need to Test With 5 Users**. Jakob Nielsen's Alertbox, 2000. URL: <<https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users>>. Acesso em Novembro/2018.
- OLIVARES, D. M.; HUNDHAUSEN, C. D. Supporting learning analytics in computing education. In: **Proceedings of the 7th International Learning Analytics & Knowledge Conference**. NY, USA: ACM, p. 584–585, 2017.
<https://doi.org/10.1145/3027385.3029472>
- ONWUEGBUZIE, A. J.; COLLINS, K. M. A Typology of Mixed Methods Sampling Designs in Social Science Research. **The Qualitative Report**, v. 12, n. 2, p. 281–316, 2007.
- OZER, D. J. Correlation and the Coefficient of Determination. **Psychological Bulletin**, American Psychological Association, v. 97, n. 2, p. 307–315, 1985.
<https://psycnet.apa.org/doi/10.1037/0033-2909.97.2.307>

- PELKOWITZ, L.; SCHWARTS, S. Asymptotically optimum sample size for quickest detection. **IEEE Transactions on Aerospace and Electronic Systems**, IEEE, v. 23, n. 2, p. 263–272, 1987. <https://doi.org/10.1109/TAES.1987.313381>
- PHAM, H. Optimal Design of Life Testing for ULSI Circuit Manufacturing. **IEEE Transactions on Semiconductor Manufacturing**, IEEE, v. 5, n. 1, p. 68–70, 1992. <https://doi.org/10.1109/66.121982>
- PIMENTEL, M. da G. et al. Supporting Educational Activities through Dynamic Web Interfaces. **Interacting with computers**, Elsevier, v. 13, n. 3, p. 353–374, 2001. [https://doi.org/10.1016/S0953-5438\(00\)00042-4](https://doi.org/10.1016/S0953-5438(00)00042-4)
- PRODANOV, V. I. In-Class Lecture Recording: What Lecture Capture has to Offer to the Instructor. In: **Proceedings of the American Society for Engineering Education-Pacific South West Section Conference**, p. 1–12, 2012.
- PUTJORN, P. et al. Designing a ubiquitous sensor-based platform to facilitate learning for young children in thailand. In: **Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services**. ACM, p. 1–13, 2017. <https://doi.org/10.1145/3098279.3098525>
- RAZALI, N.; WAH, Y. B. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. **Journal of Statistical Modeling and Analytics**, v. 2, n. 1, p. 21–33, 2011.
- RIBOLDI, J. et al. Precisão e poder de testes de homocedasticidade paramétricos e não-paramétricos avaliados por simulação. **Revista Brasileira de Biomedicina**, v. 32, n. 3, p. 334–344, 2014.
- SCHEFFÉ, H. **The Analysis of Variance**. New York: Wiley, 1959. <https://doi.org/10.1002/bimj.19610030206>
- SCHMETTOW, M. Sample Size in Usability Studies. **Communications of the ACM**, ACM, v. 55, n. 4, p. 64–70, 2012. <https://doi.org/10.1145/2133806.2133824>
- SCHMIDT, A. et al. From photography to ubiquitous capture systems. **Pervasive Computing, IEEE**, v. 13, n. 1, p. 10–13, 2014. <https://doi.org/10.1109/MPRV.2014.8>
- SETTLE, A.; DETTORI, L.; DAVIDSON, M. J. Does Lecture Capture Make a Difference for Students in Traditional Classrooms. In: **Proceedings of the 16th annual joint conference on Innovation and technology in computer science education**. ACM, p. 78–82, 2011. <https://doi.org/10.1145/1999747.1999772>
- SILVA, R. L. d. et al. Determinação do Tamanho Ótimo da Parcela Experimental pelos Métodos da Máxima Curvatura Modificado, do Coeficiente de Correlação Intraclasse

e da Análise Visual em Testes Clonais de Eucalipto. **Revista Árvore**, Sociedade de Investigações Florestais, v. 27, n. 5, p. 669–676, 2003. <https://dx.doi.org/10.1590/S0100-67622003000500009>

SLAVIN, R.; SMITH, D. The Relationship Between Sample Sizes and Effect Sizes in Systematic Reviews in Education. **Educational Evaluation and Policy Analysis**, SAGE Publications, v. 31, n. 4, p. 500–506, 2009. <https://dx.doi.org/10.3102/0162373709352369>

SMYTH, E.; VOLKER, J. **Enhancing Instruction with Visual Media: Utilizing Video and Lecture Capture**. Harrisburg, PA: IGI Global, 2013. <https://dx.doi.org/10.4018/978-1-4666-3962-1>

SNEDECOR, G. W.; COCHRAN, W. G. **Statistical Methods**. 7th. Iowa State University Press, 1980. <https://doi.org/10.1017/S0021859600074104>

STEEL, R. G. D.; TORRIE, J. H.; DICKEY, D. A. **Principles and Procedures of Statistics: a Biometrical Approach**. McGraw-Hill, New York, USA, 1980. [https://doi.org/10.1016/0020-7101\(81\)90057-X](https://doi.org/10.1016/0020-7101(81)90057-X)

TRIOLA, M. F. **Introdução à Estatística - Atualização da Tecnologia**. 11^a. ed. São Paulo, SP: LTC, 2014.

TRUONG, K. N.; HAYES, G. R. Ubiquitous Computing for Capture and Access. **Foundations and Trends in Human-Computer Interaction**, Now Publishers Inc., Hanover, MA, v. 2, n. 2, p. 95–171, 2009. <https://doi.org/10.1561/11000000014>

TUKEY, J. W. Comparing individual means in the analysis of variance. **Biometrics**, JSTOR, p. 99–114, 1949. <https://doi.org/10.2307/3001913>

WEISER, M. The Computer for the 21st Century. **Scientific American**, v. 265, n. 3, p. 66–75, 1991. <https://doi.org/10.1145/329124.329126>

WIELING, M. B.; HOFMAN, W. H. A. The Impact of Online Video Lecture Recordings and Automated Feedback on Student Performance. **Computers & Education**, Elsevier Science Ltd., Oxford, UK, v. 54, n. 4, p. 992–998, 2010. <https://doi.org/10.1016/j.compedu.2009.10.002>

WILDE, A.; BRUEGGER, P.; HIRSBRUNNER, B. An Overview of Human-Computer Interaction Patterns in Pervasive Systems. In: **Proceedings of the International Conference on User Science and Engineering**, p. 145–150, 2010. <https://doi.org/10.1109/IUSER.2010.5716740>

YASSINE, S.; KADRY, S.; SICILIA, M. A. Learning analytics and learning objects repositories: Overview and future directions. In: **Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy**.

Springer International Publishing. p. 1–29, 2016. https://doi.org/10.1007/978-3-319-17727-4_13-1

YUAN, F.; LU, J.; BATALDEN, B. M. Sample Size Determination and Confidence Interval Derivation for Exponential Distribution. In: International Offshore and Polar Engineering Conference. **The 26th International Ocean and Polar Engineering Conference**, p. 664–669, 2016.

ZHANG, J.-H. et al. What learning analytics tells us: Group behavior analysis and individual learning diagnosis based on long-term and large-scale data. **Journal of Educational Technology & Society**, International Forum of Educational Technology & Society, v. 21, n. 2, p. 245–258, 2018.

ZHAO, X.; WAN, X.; OKAMOTO, T. Adaptive Content Delivery in Ubiquitous Learning Environment. In: **Proceedings of the 6th IEEE International Conference on Wireless, Mobile and Ubiquitous Technologies in Education (WMUTE)**, p. 19–26, 2010. <https://doi.org/10.1109/WMUTE.2010.10>