



HAL
open science

Formação do raciocínio estatístico na conceptualização da estimação estatística: estudo exploratório de um dispositivo pedagógico no ensino superior

Leonardo Dalla Porta

► To cite this version:

Leonardo Dalla Porta. Formação do raciocínio estatístico na conceptualização da estimação estatística: estudo exploratório de um dispositivo pedagógico no ensino superior. Education. Université de Lyon; universidade franciscana rio grande do sul, 2019. Portuguese. NNT: 2019LYSE2064 . tel-02476133

HAL Id: tel-02476133

<https://theses.hal.science/tel-02476133v1>

Submitted on 12 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° d'ordre NNT : 2019LYSE2064

THESE de DOCTORAT DE L'UNIVERSITÉ DE LYON
en COTUTELLE avec LE CENTRE UNIVERSITAIRE FRANCISCAIN

Opérée au sein de

L'UNIVERSITÉ LUMIÈRE LYON 2

École Doctorale : ED 485

Education Psychologie Information Communication

Discipline : Sciences de l'Éducation

Soutenue publiquement le 23 Août 2019, par :

Leonardo DALLA PORTA

**Formação do raciocínio estatístico na
conceptualização da estimação estatística.**

*estudo exploratório de um dispositivo pedagógico no ensino
superior.*

Devant le jury composé de :

Nadja ACIOLY-REGNIER, Professeure des universités, Université Claude Bernard Lyon 1, Présidente

Saddo AG ALMOULOU, Professeur d'université, Universidade São Paulo, Examineur

Eleni BISOGNIN, Professeure, Centre universitaire Franciscain, Examinatrice

Luciane FLORES JACOBI, Professeure, Universidade Santa Maria, Examinatrice

Vladimir LIRA VERAS XAVIER DE ANDRADE, Professeur d'université, Université Federale rural do
Pernambuco, Examineur

Adriano MENDOÇA SOUZA, Professeur Universidade Santa Maria, Examineur

André D. ROBERT, Professeur des universités, Université Lumière Lyon 2, Co-Directeur de thèse

Silvia Maria DE AGUIAR ISAIA, Professeure, Centre universitaire Franciscain, Co-Directrice de thèse

Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité – pas d'utilisation commerciale – pas de modification](#) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer, l'adapter ni l'utiliser à des fins commerciales.



UNIVERSIDADE FRANCISCANA

SANTA MARIA - RIO GRANDE DO SUL -
BRASIL

Programa de Pós-Graduação em Ensino de Ciências
e Matemática - PPGECCIMAT



UNIVERSITÉ LUMIÈRE – LYON 2

LYON - RHÔNE ALPES - FRANCE

École Doctorale ED 485
EPIC [Éducation, Psychologie, Information et
Communication]
Unité Mixte de Recherche – UMR 5191 ICAR

LEONARDO DALLA PORTA

**FORMAÇÃO DO RACIOCÍNIO
ESTATÍSTICO NA
CONCEPTUALIZAÇÃO DA
ESTIMAÇÃO ESTATÍSTICA:
ESTUDO EXPLORATÓRIO DE UM
DISPOSITIVO PEDAGÓGICO NO
ENSINO SUPERIOR**

**LA FORMATION DU
RAISONNEMENT STATISTIQUE À
TRAVERS LA
CONCEPTUALISATION DE
L'ESTIMATION STATISTIQUE:
ÉTUDE EXPLORATOIRE À PARTIR
D'UN DISPOSITIF PÉDAGOGIQUE
DANS L'ENSEIGNEMENT
SUPÉRIEUR**

Santa Maria, 2019

LEONARDO DALLA PORTA

**FORMAÇÃO DO RACIOCÍNIO
ESTATÍSTICO NA
CONCEPTUALIZAÇÃO DA
ESTIMAÇÃO ESTATÍSTICA:
ESTUDO EXPLORATÓRIO DE UM
DISPOSITIVO PEDAGÓGICO NO
ENSINO SUPERIOR**

**LA FORMATION DU
RAISONNEMENT STATISTIQUE À
TRAVERS LA
CONCEPTUALISATION DE
L'ESTIMATION STATISTIQUE:
ÉTUDE EXPLORATOIRE À PARTIR
D'UN DISPOSITIF PÉDAGOGIQUE
DANS L'ENSEIGNEMENT
SUPÉRIEUR.**

Tese em cotutela como parte dos requisitos para obtenção do título de doutor em Ensino de Ciências e Matemática pela Universidade Franciscana.

Thèse en cotutelle dans le cadre des conditions requises pour l'obtention du titre de docteur en Sciences de l'Éducation à l'Université Lumière – Lyon 2.

Orientadora (Directrice de thèse) da
Universidade Franciscana:

Prof^ª Dr. Silvia Maria de Aguiar Isaia

Directeur de thèse (Orientador) de
l'Université Lumière – Lyon 2:

Prof^º Dr. Jean-Claude Régnier

Santa Maria, 2019

D144f	<p>Dalla Porta, Leonardo</p> <p>Formação do raciocínio estatístico na conceptualização da estimação estatística: estudo exploratório de um dispositivo pedagógico no ensino superior / Leonardo Dalla Porta ; orientadores Silvia Maria de Aguiar Isaia, Jean-Claude Régnier – Santa Maria : Universidade Franciscana – UFN, 2019.</p> <p>263 <u>f.</u>: il.</p> <p>Tese (Doutorado em Ensino de Ciências e Matemática) – Programa de Pós-graduação em Ensino de Ciências e Matemática – Universidade Franciscana – UFN</p> <p>1. Competências estatísticas 2. Estimação estatística 3. Teoria dos Campos Conceituais 4. Análise Estatística Implicativa – A.S.I. I. Isaia, Silvia Maria de Aguiar II. Régnier, Jean-Claude III. Título.</p> <p style="text-align: right;">CDU 519.22</p>
-------	---

ÁREA DE CIÊNCIAS TECNOLÓGICAS

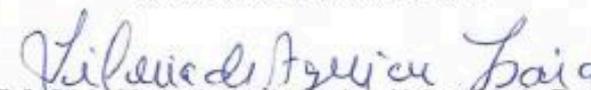
Programa de Pós-Graduação em Ensino de Ciências e Matemática

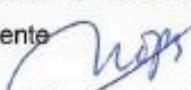
A COMISSÃO EXAMINADORA, ABAIXO-ASSINADA, APROVA A TESE:

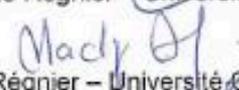
FORMAÇÃO DO RACIOCÍNIO ESTATÍSTICO NA CONCEPTUALIZAÇÃO DA
ESTIMAÇÃO ESTATÍSTICA. ESTUDO EXPLORATÓRIO DE UM DISPOSITIVO
PEDAGÓGICO NO ENSINO SUPERIOR

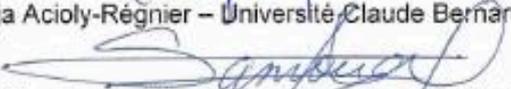
Elaborada por
Leonardo Dalla Porta

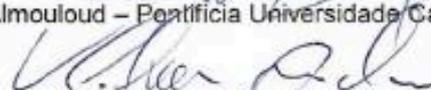
COMISSÃO EXAMINADORA:


Prof.ª Dr.ª Silvia Maria de Aguiar Isaia – Universidade Franciscana
Presidente

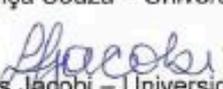

Prof. Dr. Jean-Claude Régnier – Université Lumière - Lyon 2

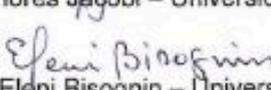

Prof.ª Dr.ª Nadja Acioly-Régnier – Université Claude Bernard – Lyon 1


Prof. Dr. Saddo Ag Almouloud – Pontifícia Universidade Católica de São Paulo


Prof. Dr. Vladimir Lira Xavier de Andrade - Universidade Federal Rural de Pernambuco


Prof. Adriano Mendonça Souza – Universidade Federal de Santa Maria


Prof.ª Luciane Flores Jacobi – Universidade Federal de Santa Maria


Prof.ª Eleni Bisognin – Universidade Franciscana

Santa Maria, 23 de agosto de 2019.

Dedico este trabalho aos meus pais, Octacílio e
Auza Carmen...com todo meu amor.

AGRADECIMENTOS

A escrita de uma tese é uma longa viagem que inclui uma trajetória permeada por inúmeros desafios, angústias, incertezas, alegrias e muitos percalços pelo caminho, mas, apesar do processo solitário a que qualquer investigador está destinado, reúne contributos de várias pessoas, indispensáveis para encontrar o melhor rumo em cada momento da caminhada. Não existem testes estatísticos que meçam a importância dessas pessoas, mas minha gratidão é desmedida. Assim, coloco alguns dos muitos agradecimentos que devo.

Aos meus pais, Octacílio e Auza. De todos os amores que podemos receber durante a vida, o amor dos nossos pais é o mais especial. Não existe outro que consiga ser incondicional, e demonstrar ao mesmo tempo um interesse tão grande e genuíno na nossa felicidade. Apenas posso agradecer por tudo que vocês me deram, pois nunca conseguirei compensar devidamente a dedicação que sempre manifestaram. Vocês são uma bênção sem comparação. Esse trabalho é nosso!

Ao meu irmão Luciano, à minha cunhada Juliana e aos meus sobrinhos Guilherme e Bernardo. Obrigado por fazerem parte da minha vida e por todo apoio e incentivo. Amo muito vocês.

Aos meus orientadores. Professor Jean-Claude Régnier, orientador, atento e amigo, que me estimulou e me fez crescer muito enquanto pessoa e profissional. O mesmo terei de dizer da minha orientadora, a Professora Silvia Maria de Aguiar Isaia, que sempre me recebeu com todo carinho. Os ensinamentos de vocês serão levados por toda minha vida, com sentimento de admiração e respeito.

À Professora Eleni Bisognin e à Professora Vanilde Bisognin. Não existem palavras que expressem minha gratidão por todos os conhecimentos e oportunidades que vocês sempre me proporcionaram. Meu carinho, minha admiração e minha gratidão eterna.

Ao Programa de Pós-Graduação em Ensino de Ciências e Matemática (PPGECIMAT), em especial à Professora Thais Scotti do Canto-Dorow, pela ternura e por estar sempre disposta em ajudar.

À Universidade Franciscana, representada pela admirável Irmã Iraní Rupolo, Reitora. Obrigado pela confiança. Espero contribuir muito com a nossa universidade.

Ao Frei Valdir Preto por toda cordialidade recebida. Seus conselhos sempre ajudaram durante a minha caminhada.

Aos meus colegas de doutorado, em especial à Janice Rachelli e à Vânia Denardi. Obrigado pela convivência.

Aos meus queridos colegas e amigos Letícia Stefenon, Letícia Fogaça e Rodrigo Pereira. Vocês são muito especiais para mim, agradeço imensamente por todo apoio e carinho de sempre.

À minha amada amiga Livia Da Cás Pereira, por todo carinho e atenção.

A cotutela proporcionou inúmeras emoções. Não é fácil sair da zona de conforto e vivenciar uma cultura diferente. Mas a vida quis me presentear com pessoas especiais que amenizaram as angústias encontradas no caminho. Professora Nadja Acioly-Régnier, obrigado pelo acolhimento e por todos os conselhos dados, minha eterna gratidão. Patrícia Gomes, minha querida amiga, obrigado por me receber de braços abertos na França, jamais esquecerei de toda tua ajuda. Luciana Silva dos Santos, minha querida amiga, quantos momentos de aprendizados que passamos juntos, obrigado por fazer parte dessa trajetória. Renata Jatobá, minha amada, que encontro maravilhoso e que honra poder participar da tua vida. Seguiremos juntos!

Meus respeitosos agradecimentos pela participação e contribuição da banca examinadora da defesa.

Não deixo de agradecer também ao imprescindível apoio da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela oportunidade de fazer a cotutela na França.

À Mãe, Rainha e Vencedora Três Vezes Admirável de Schoenstatt, obrigado pela proteção.

“Aqueles que passam por nós, não vão sós, não nos deixam sós. Deixam um pouco de si, levam um pouco de nós” Antoine de Saint-Exupéry.

RESUMO

Os dados estatísticos estão presentes em diversos meios da sociedade e são indispensáveis nos campos científicos e na formação do cidadão moderno. Por essa razão, a estatística é abordada no Brasil a partir da educação básica e permanece presente em outros níveis de ensino. Nosso estudo está situado no campo da Educação Estatística no Ensino Superior e se insere na linha de investigação voltada ao ensino e aprendizagem de conceitos da estatística, mais particularmente relacionados aos conceitos de estimação estatística. O trabalho, desenvolvido no contexto de uma pesquisa qualitativa e quantitativa, teve como objetivo identificar como se dá a formação do raciocínio estatístico por meio da conceptualização da estimação estatística. Observamos uma amostra de 81 estudantes oriundos de cursos de graduação das áreas de Ciências Sociais e Ciências Tecnológicas, de uma instituição comunitária da cidade de Santa Maria, estado do Rio Grande do Sul, Brasil. Como referencial teórico utilizamos a Teoria dos Campos Conceituais e a exploração de diversas pesquisas que tratam sobre as competências estatísticas: letramento, pensamento e raciocínio estatísticos. Baseamos nossa exploração em dados construídos por meio de dois questionários que foram aplicados em disciplinas de estatística, antes e depois do conteúdo de estimação estatística ser explorado em sala de aula. Os questionários foram organizados com questões baseadas em situações-problema relacionadas com a estimação estatística e por questões pertinentes quanto a relação do estudante com a estatística. Para o tratamento, análise e interpretação dos dados, ferramentas computacionais estatísticas foram utilizadas no quadro teórico da análise estatística implicativa (ASI), com o apoio do software CHIC e, na estatística clássica, para análises estatísticas descritivas e inferenciais, o software SPAD. Os resultados obtidos mostram que os estudantes que possuem melhores relações com a estatística utilizam os conceitos aprendidos de forma mais satisfatória. A orientação dessas relações com a estatística pode influenciar na formação do raciocínio estatístico, principalmente em razão de problemas matemáticos.

Palavras-chave: Competências estatísticas. Estimação estatística. Teoria dos Campos Conceituais. Análise Estatística Implicativa – A.S.I.

ABSTRACT

Statistical data are present in various media of society and are indispensable in scientific fields and in the training of the modern citizen. For this reason, statistics are discussed from basic education in Brazil and remain present at other levels of education. Our study is in the field of statistics education at the level of higher education and is part of the line of research focused on the teaching and learning of statistical concepts, especially related to the concepts of estimation. statistical. The work, developed as part of a qualitative and quantitative research, aimed to identify how the formation of statistical reasoning is achieved by the conceptualization of statistical estimation. From a sample of 81 undergraduate university students in the social sciences and technology sciences at a higher education institution in the city of Santa Maria, State of Rio Grande do Sul, in Brazil. As theoretical references, we use Conceptual Field Theory and the exploitation of several researches on statistical skills: literacy, thinking and statistical reasoning. We base our exploration on data constructed using two questionnaires applied in statistical disciplines, before and after having addressed, in progress, the contents concerning the statistical estimation. The questionnaires were organized with questions based on problematic situations related to statistical estimation and questions related to the student's

report to statistics and statistics. For data processing, analysis and interpretation, statistical computer tools have been used in the theoretical framework of implicative statistical analysis (ASI), with the help of CHIC software, and in classical statistics for descriptive and inferential statistical analysis of SPAD software. The results show that students who have better relations with statistics use the concepts learned in a more satisfactory way. The orientation of these relationships with statistics can influence the formation of statistical reasoning, mainly because of mathematical problems.

Keywords: Statistical skills. Statistical Estimation. Theory of Conceptual Fields. Statistical Implicative Analysis – A.S.I.

RÉSUMÉ

Les données statistiques sont présentes dans divers médias de la société et sont indispensables dans les domaines scientifiques et dans la formation du citoyen moderne. Pour cette raison, la statistique est abordée dès l'éducation de base au Brésil et reste présente à d'autres niveaux d'enseignement. Notre étude se situe dans le domaine de l'enseignement de la statistique au niveau de l'enseignement supérieur et fait partie de la ligne de recherche axée sur l'enseignement et l'apprentissage de concepts statistiques, plus particulièrement liés aux concepts d'estimation statistique. Les travaux, développés dans le cadre d'une recherche qualitative et quantitative, ont visé à identifier comment se réalise la formation du raisonnement statistique par la conceptualisation de l'estimation statistique. A partir d'un échantillon de 81 étudiants des cours de premier cycle universitaire dans les domaines des sciences sociales et des sciences technologiques d'un établissement d'enseignement supérieur de la ville de Santa Maria, dans l'état de Rio Grande do Sul, au Brésil. Comme références théoriques, nous utilisons la Théorie des Champs Conceptuels et l'exploitation de plusieurs recherches portant sur les compétences statistiques: littéracie, pensée et raisonnement statistique. Nous basons notre exploration sur des données construites à l'aide de deux questionnaires appliqués dans des disciplines statistiques, avant et après avoir abordé, en cours, le contenu portant sur l'estimation statistique. Les questionnaires ont été organisés avec des questions basées sur des situations problématiques liées à l'estimation statistique et des questions liées au rapport de l'étudiant à la statistique et aux statistiques. Pour le traitement, l'analyse et l'interprétation des données, des outils informatiques statistiques ont été utilisés dans le cadre théorique de l'analyse statistique implicative (ASI), avec l'aide du logiciel CHIC et, dans celui de la statistique classique, pour les analyses statistiques descriptives et inférentielles, du logiciel SPAD. Les résultats montrent que les étudiants ayant de meilleures relations avec la statistique et les statistiques mobilisent les concepts appris de manière plus satisfaisante. L'orientation de ces relations avec la statistique et les statistiques peut influencer la formation du raisonnement statistique, principalement en raison de problèmes mathématiques.

Mots-clés: Compétences statistiques. Estimation statistique. Théorie des champs conceptuels. Analyse Statistique implicative – A.S.I.

LISTA DE FIGURAS

Figura 1 - Etapas para a construção de um Estado da Arte	28
Figura 2 - Eixo: letramento, pensamento e raciocínio estatísticos	31
Figura 3 - Eixo: Teoria dos Campos Conceituais.....	36
Figura 4 - Eixo: Inferência Estatística	37
Figura 5 - Ações observadas nos trabalhos analisados.....	39
Figura 6 - Esquema sobre a definição da Estatística	43
Figura 7 - Esquema das operações e dos objetivos atribuídos para a Estatística observando sua divisão.....	45
Figura 8 - Tríade da Educação Estatística	46
Figura 9 - Relação entre a Educação Matemática e a Educação Estatística.....	48
Figura 10 - Dificuldades na formação de usuários da Estatística.....	51
Figura 11 - Composição do Letramento estatístico por Gal (2002)	54
Figura 12 - Tópicos do letramento estatístico para a estimação estatística	55
Figura 13 - Hábitos mentais de Chance (2002) adaptados para a Estimação Estatística	58
Figura 14 - Raciocínios para a Estatística indicados por Garfield e Gal (1999), adaptados para a Estimação Estatística	61
Figura 15 - Raciocínio estatístico integrador.....	62
Figura 16 - Estados para o desenvolvimento do espírito científico segundo Bachelard	64
Figura 17 - Etapas para a formação do espírito estatístico	65
Figura 18 - Elementos para a formação do espírito estatístico.....	66
Figura 19 - Competências estatísticas	67
Figura 20 - Triângulo das situações didáticas	69
Figura 21 - Triângulo das situações didáticas para a Estatística	70
Figura 22 - Contexto do Dispositivo Pedagógico.....	71
Figura 23 - Contexto do questionário Q1	72
Figura 24 - Contexto do questionário Q2	73
Figura 25 - Domínios da Estatística	77
Figura 26 - Grandes áreas da Estatística.....	77
Figura 27 - Conceitos de probabilidade.....	82
Figura 28 - População e amostra	88
Figura 29 - Tipos de variáveis aleatórias.....	88
Figura 30 - Representação de amostragem.....	89

Figura 31 - Dados agrupados sem intervalo de classes	100
Figura 32 - Dados agrupados com intervalos de classes	101
Figura 33 - Processo da estimação	102
Figura 34 - Inferência sobre o parâmetro populacional θ por meio da distribuição	109
Figura 35 - Distribuição amostral da média: com e sem reposição	110
Figura 36 - Processo da proporção amostral em relação à proporção populacional	113
Figura 37 - Tipos de estimação.....	117
Figura 38 - Resultado do teste de tiro ao alvo	118
Figura 39 - Propriedades para um bom estimador.....	119
Figura 40 - Significado de um intervalo de confiança para μ com $1 - \alpha = 0,90$ e σ conhecido	125
Figura 41 - Curva Normal Padrão para o intervalo de confiança da proporção	128
Figura 42 - Estimação de parâmetros populacionais	129
Figura 43 - Mapa conceitual da teoria dos campos conceituais	133
Figura 44 - Conjunto do conceito	133
Figura 45 - Elementos de um esquema.....	134
Figura 46 - Conjunto do conceito da média	136
Figura 47 - Processo iterativo da construção do conhecimento	139
Figura 48 - Identificação dos estudantes	141
Figura 49 - Exemplo de uma árvore de similaridade.....	144
Figura 50 - Exemplo de um grafo implicativo	145
Figura 51 - Exemplo de árvore coesiva	146
Figura 52 - Interface do software SPAD 9.1	147
Figura 53 - Q1: informações sobre a formação do estudante	149
Figura 54 - Q1: V06	150
Figura 55 - Q1: V07 e V08.....	151
Figura 56 - Q1: V09	151
Figura 57 - Q1: V10 e V11.....	152
Figura 58 - Q1: V12	152
Figura 59 - Q1: V13	153
Figura 60 - Q1: V14	153
Figura 61 - Q1: V15	154
Figura 62 - Índice de positividade	155
Figura 63 - Zonas de comportamento em relação aos índices de positividade	155

Figura 64 - Q1: V16	156
Figura 65 - Q2: variáveis da primeira situação.....	158
Figura 66 - Q2: variáveis da segunda situação	159
Figura 67 - Q2: variáveis da terceira situação	160
Figura 68 - Critério para analisar as respostas de Q2.....	160
Figura 69 - Planilha com as variáveis binárias	161
Figura 70 – Síntese das ferramentas de construção, tratamento e análise dos dados.....	161
Figura 71 - aV06 e dV06: hoje, o que lhe evoca a Estatística – Palavras	166
Figura 72 - aV06: Palavras por níveis de Positividade "Antes"	167
Figura 73 - Frases "Antes".....	168
Figura 74 - dV06: Palavras por níveis de positividade “Depois”	168
Figura 75 - Frases "Depois".....	169
Figura 76 - V09: Atualmente o que significa para você os termos a seguir?.....	173
Figura 77 - aV09c e dV09c: Estimação.....	173
Figura 78 - aV09c: Estimação	174
Figura 79 - dV09c: Estimação	174
Figura 80 - Sondagem de opinião: aV10 e dV10	175
Figura 81 - aV11 e dV11: Nos dias de hoje, qual utilidade você atribui para a Estatística?..	175
Figura 82 - aV12 e dV12: Para você, quais são as três maiores dificuldades encontradas numa disciplina de Estatística?.....	176
Figura 83 - aV12 X Zonas de comportamento “Antes”	177
Figura 84 - dV12 X Zonas de comportamento "Depois"	177
Figura 85 - aV13 e dV13: Para você, quais são as três maiores contribuições encontradas numa disciplina de Estatística?.....	178
Figura 86 - aV16 e dV016: Atualmente, qual sua definição de Estatística?.....	185
Figura 87 - Árvore de similaridades "Antes"	188
Figura 88 - Resolução das variáveis aS1_1 e aS2_2 entre os estudantes com os menores índices de positividade.....	189
Figura 89 - Respostas encontradas na primeira situação entre os estudantes com maiores índices de positividade.....	190
Figura 90 - Grafo implicativo "Antes"	191
Figura 91 - Árvore coesiva "Antes"	192
Figura 92 - Árvore de similaridades "Depois"	193
Figura 93 - Invariantes operatórios observados sobre a estimação	195

Figura 94 - Grafo implicativo "Depois"	196
Figura 95 - Grafo implicativo em relação a variável dS3_2.....	196
Figura 96 - Formação de um possível campo conceitual para a estimação.....	197
Figura 97 - H1	197
Figura 98 - H2	198
Figura 99 - H3	199
Figura 100 - Referências para o desenvolvimento do raciocínio estatístico	201

LISTA DE GRÁFICOS

Gráfico 1 - Distribuição Normal.....	106
Gráfico 2 - Simetria da Distribuição Normal	106
Gráfico 3 - Curva Normal Padrão para o intervalo de confiança da média populacional com σ conhecido.....	123
Gráfico 4 - Curva Normal Padrão para o intervalo de confiança da média com σ desconhecido	126
Gráfico 5 - Número de estudantes nos cursos de graduação participantes da pesquisa	141
Gráfico 6 - Gênero dos estudantes participantes da pesquisa.....	142
Gráfico 7 - Distribuição dos alunos por semestre.....	142
Gráfico 8 - V03: Você já enfrentou situações que exigiram conhecimentos estatísticos?	164
Gráfico 9 - V04: Quando os conhecimentos em Estatística são requisitados, isso lhe causa problemas?.....	165
Gráfico 10 - V07: Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação?.....	170
Gráfico 11 - V08: Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação?.....	171
Gráfico 12 - aV14: Você já obteve conhecimentos em Estatística?.....	179
Gráfico 13 - dV14: Você já obteve conhecimentos em Estatística?	179
Gráfico 14 - aV15 e dV15: dispersão dos índices	180

LISTA DE QUADROS

Quadro 1 - Dados de X e de Y para a propriedade 4 da média	94
Quadro 2 - Dados de X e de Y para o exemplo da propriedade 4 da média	95
Quadro 3 - Variância Populacional	98
Quadro 4 - Variância amostral alterada	98
Quadro 5 - Desvio Padrão populacional e amostral	99
Quadro 6 - Expressão geral da distribuição Binomial	104
Quadro 7 - Principais características da distribuição Binomial	104
Quadro 8 - Função densidade de probabilidade da distribuição Normal	105
Quadro 9 - Transformação de $N\mu, \sigma^2$ para $N(0,1)$	107
Quadro 10 - Amostras extraídas com reposição e suas respectivas médias	110
Quadro 11 - Amostras extraídas sem reposição e suas respectivas médias	112
Quadro 12 - Amostras extraídas com reposição e suas respectivas proporções	114
Quadro 13 - Amostras extraídas com reposição e suas respectivas proporções	115
Quadro 14 - Intervalo de confiança para a média populacional com σ conhecido em populações infinitas ou com reposição na amostra	124
Quadro 15 - Principais níveis de interesse	124
Quadro 16 - Intervalo de confiança para a média populacional com σ conhecido em populações finitas ou sem reposição na amostra	125
Quadro 17 - Principais níveis de interesse	126
Quadro 18 - Intervalo de confiança para a média populacional com σ desconhecido em populações finitas ou sem reposição na amostra	127
Quadro 19 - Intervalo de confiança para a média populacional com σ desconhecido em populações finitas ou sem reposição na amostra	127
Quadro 20 - Intervalo de confiança para a proporção de populações infinitas ou com reposição na amostra	128
Quadro 21 - Intervalo de confiança para a proporção de populações finitas ou sem reposição na amostra	129
Quadro 22 - Conjunto do conceito da estimação pontual da média populacional	137
Quadro 23 - Conjunto do conceito da estimação intervalar da média populacional	137
Quadro 24 - Q2: Situação 1	157
Quadro 25 - Q2: Situação 2	158
Quadro 26 - Q2: Situação 3	159

Quadro 27 - Frequência dos estudantes por semestre.....	166
Quadro 28 - aV06 e dV06: cinco palavras com maior frequência	167
Quadro 29 - Relação V07 X índice de positividade "Antes e Depois"	171
Quadro 30 - Relação V08 X índice de positividade "Antes e Depois"	172
Quadro 31 - aV12 e dV12: cinco palavras com maior frequência	176
Quadro 32 - aV13 e dV13: cinco palavras com maior frequência	178
Quadro 33 - Atitudes características: a_Positividade_1	181
Quadro 34 - Atitudes características: a_Positividade_2	181
Quadro 35 - Atitudes características: a_Positividade_3	182
Quadro 36 - Atitudes características: a_Positividade_4	182
Quadro 37 - Atitudes características: d_Positividade_1	183
Quadro 38 - Atitudes características: d_Positividade_2	184
Quadro 39 - Atitudes características: d_Positividade_3	184
Quadro 40 - Atitudes características: d_Positividade_4	185

LISTA DE TABELAS

Tabela 1 - Produções científicas selecionadas para análise.....	29
Tabela 2 - Ações e demandas das competências estatísticas.....	68
Tabela 3 - Distribuição de frequências sem intervalos de classes.....	91
Tabela 4 - Distribuição de frequências com intervalos de classes	92
Tabela 5 - Desvios dos Grupos A e B	96
Tabela 6 - Medidas estatísticas para distribuição de frequências	100
Tabela 7 - Principais parâmetros e seus respectivos estimadores	102
Tabela 8 - Distribuição amostral das médias extraídas das amostras com reposição.....	111
Tabela 9 - Distribuição amostral das médias extraídas das amostras sem reposição	112
Tabela 10 - Propriedades da distribuição da média amostral	113
Tabela 11 - Distribuição amostral das proporções extraídas das amostras com reposição	114
Tabela 12 - Distribuição amostral das proporções extraídas das amostras sem reposição.....	115
Tabela 13 - Propriedades da distribuição amostral.....	116
Tabela 14 - Variáveis com maiores contribuições aos nós com níveis superiores a 0,70.	189
Tabela 15 - Variáveis com maiores contribuições aos nós com níveis superiores a 0,70	194

LISTA DE ABREVIATURAS E SIGLAS

ABES	Agence Bibliographique de l'Enseignement Supérieur
ASI	Análise Estatística Implicativa
BOLEMA	Boletim de Educação Matemática
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CFIES	Colloque Francophone International sur l'Enseignement de la Statistique
CHIC	Classification Hierarchique Implicative Cohésitive
EPIC	Éducation, Psychologie, Information et Communication
PPGECIMAT	Programa de Pós-Graduação em Ensino de Ciências e Matemática
SILASI	Simpósio Internacional Lusófono de Análise Estatística Implicativa
SPAD	Système Portable d'Analyse des Données
TCC	Teoria dos Campos Conceituais
UFN	Universidade Franciscana

SUMÁRIO

INTRODUÇÃO	21
1.1. Objetivos da pesquisa	24
1.2.1. Objetivo geral	24
1.2.2. Objetivos específicos	24
1.2. Hipóteses	25
1.3. Estrutura da tese	25
PARTE 1: FUNDAMENTOS TEÓRICOS E ELEMENTOS DE PESQUISA	27
1. Utilizando os passos do estado da arte: um breve panorama bibliográfico das pesquisas realizadas no campo da Educação Estatística	27
1.1. Apreciação das tendências das produções selecionadas	30
1.2. Conclusões sobre as tendências observadas nas produções selecionadas	38
2. A Estatística: visões conceituais e educacionais	41
2.1. A Estatística enquanto ciência	42
2.2. Educação Estatística.....	46
2.3. Educação Estatística no Ensino Superior: a Estatística como uma disciplina de serviço	49
2.4. Competências para a Educação Estatística	52
2.4.1. Letramento estatístico.....	52
2.4.2. O pensamento estatístico	56
2.4.3. O raciocínio estatístico	59
2.4.4. Compreendendo o espírito estatístico.....	63
2.4.5. Reflexões finais sobre o letramento, o pensamento e o raciocínio estatístico	67
2.5. Compreensão de um dispositivo pedagógico para a Estatística.....	68
3. Inferência Estatística: um olhar sobre a estimação estatística	74
3.1. Alguns elementos sobre a história da Estatística	74
3.2. Alguns elementos sobre a história da estimação estatística.....	77
3.3. Alguns elementos sobre a história da probabilidade.....	79
3.3.1. A definição Clássica de probabilidade	82
3.3.2. A definição frequentista de probabilidade.....	84
3.3.3. A definição axiomática de probabilidade	85

3.4. Breve estudo sobre alguns conceitos fundamentais da Estatística para situar a estimação.....	87
3.4.1. População e amostra	87
3.4.2. Variáveis aleatórias	88
3.4.3. Uma síntese sobre Amostragem	89
3.4.4. Distribuição de frequências	90
3.4.5. Média.....	92
3.4.6. Variância.....	96
3.4.7. Desvio padrão	99
3.4.8. Cálculo das medidas estatísticas a partir de uma distribuição de frequências.....	100
3.4.9. Parâmetros, estimadores e estimativas	101
3.4.10. Modelos de distribuição de probabilidades.....	103
3.4.10.1. Distribuição binomial.....	103
3.4.10.2. Distribuição normal	105
3.4.11. Teorema Central Limite	108
3.4.12. Distribuições Amostrais	108
3.4.12.1. Distribuição amostral da média: com e sem reposição	109
3.4.12.2. Distribuição amostral da proporção: com e sem reposição ..	113
3.5. Estimação de parâmetros populacionais	116
3.6. Propriedades dos estimadores	117
3.6.1. Estimador não tendencioso.....	119
3.6.2. Estimador consistente	121
3.7. Intervalo de Confiança	121
3.7.1. Estimação intervalar da média populacional.....	123
3.7.1.1. Desvio Padrão populacional σ conhecido.....	123
3.7.1.2. Desvio Padrão populacional σ desconhecido	125
3.7.2. Intervalo de confiança para a proporção	127
3.8. Fechamento sobre a estimação de parâmetros populacionais.....	129
4. Alguns elementos sobre a Teoria dos Campos Conceituais	130
4.1. A estimação de parâmetros populacionais enquanto campo conceitual	136
PARTE 2: METODOLOGIA DA CONSTRUÇÃO E TRATAMENTO DOS DADOS.....	139
1. Contexto e sujeitos da pesquisa	140

2.	Ferramentas computacionais	143
2.1.	Software CHIC.....	143
2.2.	Software SPAD	146
3.	Construção dos dados: questionários da pesquisa	148
3.1.	Questionário Q1	149
3.2.	Questionário Q2	156
PARTE 3: RESULTADOS, DISCUSSÕES E PROLONGAMENTOS.....		163
1.	Resultados e discussões do Q1	163
2.	Resultados e discussões do Q2	187
2.1.	Resultados das situações antes da formação sobre estimação	188
2.2.	Resultados das situações depois da formação sobre estimação	193
3.	Considerações finais das hipóteses levantadas e prolongamentos.....	197
3.1.	Prolongamentos da pesquisa	199
CONCLUSÃO GERAL		201
REFERÊNCIAS		203
ANEXO A - QUESTIONÁRIO Q1.....		213
ANEXO B - QUESTIONÁRIO Q2.....		216
ANEXO C - CURVA NORMAL REDUZIDA		218
ANEXO D - ÍNDICE DE POSITIVIDADE		219
ANEXO E - V03 X POSITIVIDADE.....		220
ANEXO F - V04 X POSITIVIDADE.....		221
ANEXO G - SEMESTRE X NÍVEL DE POSITIVIDADE		222
ANEXO H - FRASES X POSITIVIDADE.....		226
ANEXO I - FRASES X POSITIVIDADE DEPOIS		228
ANEXO J - aV07 X dV07		230
ANEXO K - aV08 X dV08		231
ANEXO L - CORRELAÇÃO ENTRE V07 E V08.....		232
ANEXO M - V09 FRASES		233
ANEXO N - FRASES SOBRE ESTIMAÇÃO ANTES.....		234
ANEXO O - FRASES SOBRE ESTIMAÇÃO DEPOIS.....		236
ANEXO P - V011 FRASES SOBRE UTILIDADE DE ESTATÍSTICA ANTES E DEPOIS		238
ANEXO Q - aV14 E dV14 X POSITIVIDADE		242
ANEXO R - TODAS AS ATITUDES		243

ANEXO S - TESTE T - aV15 E dV15.....	250
ANEXO T - POSITIVIDADE X SEXO.....	251
ANEXO U - V16 DEFINIÇÃO PARA ESTATÍSTICA.....	252
ANEXO V - GRUPOS X POSITIVIDADE.....	255
ANEXO W - Q2 ANTES - TABELA BINÁRIA.....	256
ANEXO X - Q2 DEPOIS - TABELA BINÁRIA	257
ANEXO Y - DADOS ÁRVORE ANTES	258
ANEXO Z - DADOS ÁRVORE DEPOIS	261

INTRODUÇÃO

Durante toda minha formação na graduação de bacharelado em matemática e, após licenciatura em matemática, preoquei-me em entender e descobrir qual seria a melhor forma de trabalhar todos os conhecimentos que estavam sendo por mim adquiridos para usá-los em sala de aula como professor. Na graduação de bacharelado obtive pouco contato com pesquisas na área da Educação Matemática, mas, por outro lado, consolidei uma boa base matemática que se tornaria ferramenta muito importante para a compreensão e desenvolvimento de minhas futuras pesquisas. Foi durante o curso de licenciatura em matemática que passei a conhecer obras de pesquisadores na área da Educação Matemática e, a partir da leitura de artigos, periódicos e relatos é que pude chegar à conclusão de estar no caminho certo na busca do que precisava saber como professor.

Iniciei minha carreira, como professor, no ano de 2004, em uma escola do ensino particular de Santa Maria, RS, assumindo duas turmas da antiga oitava série, hoje sétimo ano. Depois de um ano, assumi todas as turmas do Ensino Fundamental da mesma escola e pude, durante dois anos, acompanhar o nível de desenvolvimento dos estudantes na construção de suas bases matemáticas. Logo após esses dois anos, tomei regência das turmas de Ensino Médio da mesma escola, ao mesmo tempo em que ministrava aulas de matemática em uma escola de Educação de Jovens e Adultos, logicamente em seus respectivos turnos e horários.

Na busca de novos e mais apropriados conhecimentos, ingressei no Mestrado Profissionalizante em Ensino de Física e de Matemática do então Centro Universitário Franciscano, hoje Universidade Franciscana. O ingresso no mestrado me possibilitou lecionar no Ensino Superior onde comecei ministrando aulas de Matemática Financeira e Estatística. No término do mestrado, passei a pertencer ao corpo docente da Universidade Franciscana onde atualmente atuo regendo várias disciplinas, entre elas a Estatística. Foi durante esse período que comecei a perceber melhor a dinâmica do ensino da Estatística, sua importância para uma boa formação acadêmica além de algumas dificuldades que os estudantes apresentavam.

No intuito de investigar o processo de ensino e aprendizagem da estatística no Ensino Superior, ingressei no doutorado, onde passei a ter um maior contato com pesquisas referentes ao tema da Educação Estatística, motivado em contribuir ainda mais na formação acadêmica dos estudantes. Foi durante uma palestra proferida na Universidade Franciscana pelo professor Doutor Jean-Claude Régnier, que recebi a sugestão de desenvolver uma pesquisa sobre a formação do raciocínio estatístico por meio da organização de conceitos da estimação estatística. Dessa forma, aceitando o desafio, iniciei o trabalho de pesquisa sob o regime de

cotutela que envolveu duas instituições de ensino, uma brasileira, que é a Universidade Franciscana, onde a orientação da tese foi conduzida pela professora Doutora Silvia Maria de Aguiar Isaia, e uma outra, que é a Universidade Lumière Lyon 2, da França, onde o professor Doutor Jean-Claude Régnier foi o orientador da tese. Assim sendo, a partir dessa união, o trabalho começou a ser escrito na primeira pessoa do plural. A pesquisa que desenvolvemos e que aqui estamos apresentando está posicionada na direção da área da Educação Estatística.

A estatística está presente em diversos meios de informações da sociedade moderna, sendo indispensável no âmbito científico e na formação do cidadão. Por essa razão, a estatística é apresentada no ensino básico e difundida durante os demais níveis de ensino. Contudo, no ensino básico a estatística pouco aparece, sendo atribuída ao conteúdo de matemática. No Ensino Superior, como aponta Carzola (2002), ela aparece como disciplina obrigatória da maioria dos cursos de graduação, fazendo parte da formação acadêmica de diversos profissionais. Ainda, segundo a autora, o ensino dessa disciplina enfrenta alguns problemas, havendo a necessidade de pesquisas que se ocupem desse tema.

A Educação Estatística desempenha um papel fundamental na investigação dos processos educacionais referentes a estatística, em todos os níveis de escolaridade, observando em suas discussões três competências estatísticas que são utilizadas como aporte teórico das pesquisas na área: o letramento estatístico, o pensamento estatístico e o raciocínio estatístico. Diante disso, a fim de desenvolver essa tese, propomos a seguinte questão norteadora para a investigação:

Como podemos entender o processo de formação do raciocínio estatístico durante a aprendizagem da estimação estatística no Ensino Superior?

Com base nessa pergunta, conduzimos nossa pesquisa procurando ao longo do trabalho responder, ou pelo menos, discutir as possibilidades junto ao questionamento levantado.

Para o levantamento de dados, dentro do contexto de uma pesquisa qualitativa e quantitativa, a investigação contou com a aplicação de dois questionários respondidos por estudantes de graduação que estavam cursando uma disciplina de estatística. Essa aplicação contou com dois momentos, o primeiro foi antes da disciplina tratar sobre o conteúdo de estimação e, o segundo, após os estudantes receberem o conteúdo por meio de aulas expositivas. Os questionários abordaram questões sobre a relação dos estudantes com a estatística e situações-problema envolvendo a estimação estatística.

Para subsidiar nossa investigação, analisamos as competências estatísticas observando diversas pesquisas, a exemplo de Carzola (2002), Gal (2002), Chance (2002), Merino (2003), Soares (2004), Campos (2007), Gitirana *et al.* (2010) e outros, mostrando que uma boa

formação estatística é consolidada quando o letramento, o pensamento e o raciocínio estatísticos atuam em conjunto, atendendo as demandas operacionais e interpretativas exigidas em situações onde a estatística se faz presente. Realizamos, também, um breve panorama bibliográfico dos trabalhos ocorridos nos últimos anos no campo da Educação Estatística, observando tendências e procurando elementos que pudessem contribuir para o desenvolvimento da nossa investigação. Trazemos uma discussão sobre a estatística do ponto de vista conceitual e educacional, observando seu papel científico e fazendo uma distinção entre a Educação Matemática e a Educação Estatística.

A estimação estatística, enquanto objeto de saber científico, foi examinada em um capítulo onde trouxemos elementos históricos e um breve estudo sobre alguns conceitos fundamentais da estatística. Além disso, observamos na Teoria dos Campos Conceituais, a formação de um campo conceitual para a estimação, que nos ajudou nas análises dos dados construídos pelos questionários.

Durante as análises, levantamos algumas hipóteses que poderiam estar envolvidas com o desenvolvimento do raciocínio estatístico, como a relação do estudante com a estatística e com a matemática. Essas análises foram possíveis por meio do tratamento descritivo dos dados, onde utilizamos o software SPAD.9.1¹ e o tratamento através do quadro teórico da Análise Estatística Implicativa – A.S.I., com o apoio do software CHIC².7.

Em relação ao regime da cotutela, o autor da tese, com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), permaneceu durante 12 meses (abril de 2017 a março de 2018) estudando na Escola Doutoral ED 485 EPIC (Éducation, Psychologie, Information et Communication), do laboratório de pesquisa UMR 5191 ICAR, da Universidade Lumière Lyon 2, na cidade de Lyon, França. Esse período, além de ter sido uma experiência de crescimento pessoal, contribuiu para o desenvolvimento da pesquisa. Foi participando dos seminários do professor Doutor Jean-Claude Régnier que fortalecemos nossos conhecimentos com a Análise Estatística Implicativa (ASI), utilizando o software CHIC. Conhecemos e trabalhamos, também, com o software SPAD. Esses conhecimentos foram de grande valia para impulsionar a pesquisa, além, é claro, das orientações dos professores orientadores. Nesse período tivemos a oportunidade de levarmos nossa pesquisa em três eventos. O primeiro foi o Colloque Francophone International sur l'Enseignement de la Statistique (CFIES), da

¹ Système Portable d'Analyse des Données (SPAD). COHERIS – www.coheris.com - 4, Rue du Port aux Vins – 92150 Suresnes – France.

² Classification Hierarchique Implicative Cohésitive (CHIC). Disponível em: <https://ardm.eu/partenaires/logiciel-danalyse-de-donnees-c-h-i-c/>.

Sociedade Francesa de Estatística, na cidade de Grenoble, onde apresentamos um artigo intitulado “Question autour de l’enseignement-apprentissage du raisonnement statistique à travers l’estimation statistique”. Participamos da nona edição do Colloque d’Analyse Statistique Implicative, na cidade de Belfort, onde levamos o artigo intitulado “Construção do raciocínio estatístico numa graduação em Ciências Sociais no Brasil: uma análise preliminar no quadro A.S.I.”. Na ocasião, estavam presentes os professores Régis Gras, fundador da Análise Estatística Implicativa, e Raphaël Couturier que continua desenvolvendo o software CHIC, além de Gérard Vergnaud, autor da Teoria dos Campos Conceituais, o que foi um desafio e uma honra imensa em poder apresentar nosso trabalho. O terceiro foi o primeiro Simpósio Internacional Lusófono de Análise Estatística Implicativa (SILASI), que ocorreu em Recife, Brasil. Apresentamos por videoconferência o trabalho intitulado “Campos Conceituais da Estimção Estatística: uma análise no quadro A.S.I.”.

Contudo os resultados da pesquisa se tornam uma produção para o ensino, especificamente voltada às experiências em sala de aula, vindo ao encontro do seu principal objetivo que é, justamente, contribuir para a melhoria do ensino.

1.1. Objetivos da pesquisa

Como forma de delinear nossas ações para o desenvolvimento da pesquisa, apresentamos o objetivo geral desta pesquisa junto com os objetivos específicos.

1.2.1. Objetivo geral

Entender melhor o processo de formação do raciocínio estatístico na aprendizagem da estimção estatística no Ensino Superior.

1.2.2. Objetivos específicos

- a) Construir dados aplicando questionários com situações-problema que possam contribuir para a observância da formação do raciocínio estatístico;
- b) Compreender melhor como se processa a formação do raciocínio estatístico à luz da Teoria dos Campos Conceituais;
- c) Tratar os dados construídos com uma ferramenta estatística no quadro da ASI e para a estatística descritiva o apoio do software SPAD.

1.2. Hipóteses

- a) H1: Existem limitações no processo de formação do raciocínio estatístico, na estimação estatística, ligadas aos procedimentos e conceitos básicos da Estatística Descritiva;
- b) H2: A relação afetiva do estudante com a Estatística pode influenciar na formação do raciocínio estatístico;
- c) H3: A relação afetiva do estudante com a Matemática pode influenciar na formação do raciocínio estatístico.

1.3. Estrutura da tese

A tese está dividida em três partes.

- a) Parte 1: Fundamentos teóricos e elementos da pesquisa;
- b) Parte 2: Metodologia da construção e tratamento dos dados;
- c) Parte 3: Resultados, discussões e prolongamentos.

Na primeira parte iniciamos com a utilização dos passos do estado da arte para apurar um breve panorama bibliográfico das pesquisas realizadas no campo da Educação Estatística. O segundo capítulo tratou da estatística por meio de discussões do ponto de vista conceitual e educacional. Observamos a estatística enquanto ciência e o seu papel na educação superior. Expomos as competências estatísticas, considerando o letramento, o pensamento e o raciocínio estatísticos. Abordamos, ainda, a percepção de um dispositivo pedagógico para a estatística. Desenvolvemos no terceiro capítulo desta parte, o levantamento do saber científico, envolvendo elementos históricos e um estudo de alguns conceitos fundamentais da estatística para situar a estimação estatística. No quarto capítulo, tratamos da Teoria dos Campos Conceituais, observando um campo conceitual para a estimação estatística.

A segunda parte desta tese aborda a metodologia da construção e tratamento dos dados. Introduzimos com a apresentação do contexto e os sujeitos da pesquisa. No segundo capítulo apresentamos as ferramentas computacionais que foram empregadas na pesquisa, expondo o software SPAD e a utilização da Análise Estatística Implicativa – A.S.I. por meio do software CHIC. O terceiro capítulo revela os questionários utilizados na pesquisa e todas as variáveis que foram envolvidas nas análises.

Na terceira parte apresentamos os resultados, discussões e prolongamentos. Iniciamos com as discussões do primeiro questionário que tratou das atitudes dos estudantes em relação a

estatística, fazendo uma análise descritiva dos dados apurados e procurando informações para a validação das hipóteses da pesquisa. O segundo capítulo discorre sobre as análises do segundo questionário que tratou sobre as situações-problema relacionadas com a estimação estatística, fazendo uma análise no quadro A.S.I. e explorando a Teoria dos Campos Conceituais (TCC). No terceiro capítulo evidenciamos as considerações finais sobre a validade das hipóteses e as conclusões finais. Apresentamos, ainda, o prolongamento das discussões, indicando os limites desta tese e o que pretendemos fazer para dar continuidade à pesquisa.

PARTE 1:

FUNDAMENTOS TEÓRICOS E ELEMENTOS DE PESQUISA

Dividido em quatro capítulos, os fundamentos teóricos e elementos da pesquisa trazem discussões pertinentes para o embasamento do nosso trabalho. O primeiro capítulo traz um levantamento sobre algumas pesquisas realizadas e que possuem concepções semelhantes ao nosso objeto de pesquisa. O segundo capítulo trata da Estatística enquanto ciência, e aborda uma visão educacional dessa disciplina no Ensino Superior. Além disso, discorre sobre as competências estatísticas (letramento, pensamento e raciocínio) e descreve a ideia da concepção de um dispositivo pedagógico. O terceiro capítulo mostra tópicos históricos da estatística e a apreciação de conceitos e procedimentos estatísticos que conduzem para a Estimação Estatística. O quarto e último capítulo aborda a Teoria dos Campos Conceituais que orientou na construção do conceito de estimação.

1. Utilizando os passos do estado da arte: um breve panorama bibliográfico das pesquisas realizadas no campo da Educação Estatística

Com o intuito de colaborar de forma relevante com a Educação Estatística, apresentamos aqui uma investigação em relação as contribuições acadêmicas já existentes que norteiam sobre a temática da nossa pesquisa. Dessa forma, estaremos considerando as discussões já propostas e inserindo um trabalho original no processo de produção do conhecimento na área da Educação Estatística. Os avanços das pesquisas, nas diferentes áreas do conhecimento, provocam a produção contínua de novos e numerosos estudos acadêmicos. Esse acúmulo de produções, muitas vezes, pode gerar uma desordem de informações sobre os temas que já foram abordados e as descobertas já realizadas no âmbito acadêmico. Soares e Maciel (2000) entendem que a evolução da ciência promove a necessidade de estudos exploratórios que estejam preocupados com a ordenação das informações e dos resultados já obtidos, definindo, assim, ao que chamam de estado da arte. Laranjeira (2003) salienta que o termo Estado da Arte vem da tradução literal do inglês *tate of the art*, que significa o diagnóstico de algo. Ferreira (2002) define estado da arte como um estudo bibliográfico de caráter inventariante que, além de mapear produções teóricas de determinada área do conhecimento, tem o objetivo de discutir em que aspectos e dimensões as produções acadêmicas estão sendo conduzidas. A autora enfatiza, ainda, a importância de realizar esse mapeamento utilizando diferentes bancos de dados, observando teses de doutorado, dissertações e artigos publicados

em periódicos. Romanowski e Ens (2006) entendem que o Estado da Arte é uma sistematização de dados que abrangem toda uma área do conhecimento nos diferentes aspectos que geram produções. Neste sentido, entendemos que elaborar um estado da arte, para realizar uma pesquisa, pode auxiliar na verificação de tendências e na descoberta de lacunas que podem ser exploradas. Observando os possíveis passos que articulam o processo de elaboração do estado da arte, apuramos um panorama bibliográfico de trabalhos que circundam com as ideias da nossa pesquisa. A Figura 1, a seguir, expõe um modelo esquemático com as sete etapas que Romanowski e Ens (2006) consideram necessárias.

Figura 1 - Etapas para a construção de um estado da arte



Fonte: Adaptado de Romanowski e Ens (2006).

O primeiro passo é a determinação dos descritores que direcionam a busca das informações pretendidas, sendo assim, fizemos o uso dos seguintes descritores: Educação Estatística no Ensino Superior e *Éducation Statistique dans l'Enseignement Supérieur*. O descritor na língua francesa deve-se ao fato de realizarmos busca em um banco de dados francês, tendo em vista que o estudo do nosso doutorado está inserido em um contexto de cotutela entre o Brasil e a França. A busca do corpus, para nossa investigação, ocorreu inteiramente de forma digital por produções brasileiras e francesas referentes a teses, dissertações e artigos. Dessa forma, para a realização do segundo passo, foram escolhidos os seguintes bancos de dados: Catálogo de Teses e Dissertações (CAPES)³, Biblioteca Digital Brasileira de Teses e Dissertações⁴, Boletim de Educação Matemática (BOLEMA)⁵ e Agence Bibliographique de l'Enseignement Supérieur (ABES)⁶. As buscas nos remeteram a inúmeros trabalhos já existentes sobre o assunto, dos quais foram selecionados alguns, observando o terceiro passo da construção do Estado da Arte, que são as escolhas dos critérios de seleção das publicações. Adotamos como critério, três eixos que sustentam nossa pesquisa, são eles: letramento, pensamento e raciocínio estatísticos; a estatística inferencial; e a Teoria dos Campos Conceituais envolvida com a temática da Estatística. Nosso foco de escolha centrou-se nos trabalhos que corroboram com a Educação Estatística no Ensino Superior realizados entre os

³ Disponível em: <https://catalogodeteses.capes.gov.br/catalogo-teses/#/>.

⁴ Disponível em: <http://bdtd.ibict.br/vufind/>.

⁵ Disponível em: <http://www.periodicos.rc.biblioteca.unesp.br/index.php/bolema>.

⁶ Disponível em: <http://www.theses.fr/>.

anos de 2010 a 2017, sendo que, alguns trabalhos foram selecionados por conterem aspectos de relevância análoga à nossa pesquisa. Dessa forma, seguindo o quarto passo do Estado da Arte, coletamos três artigos que foram nominados de A1, A2 e A3 e quatro dissertações que chamamos de D1, D2, D3 e D4 e, ainda, oito teses denominadas de T1, T2, T3, T4, T5, T6, T7 e T8. A Tabela 1, a seguir, apresenta o ano, o título e os autores das produções selecionadas.

Tabela 1 - Produções científicas selecionadas para análise

 http://www.periodicos.rc.biblioteca.unesp.br/index.php/bolema		
ANO	AUTORES(S)	TÍTULO
A1	2011 Verônica Kataoka, Claudette Vendramini, Claudia da Silva e Maria de Oliveira	Evidências de Validade de uma Prova de Letramento Estatístico: um estudo com estudantes universitários de cursos tecnológicos.
A2	2011 Arlindo José de Souza Junior e Sandra Gonçalves Vilas Bôas Campos.	Trabalho de Projetos no Processo de Ensinar e Aprender Estatística na Universidade.
A3	2013 Celi Espasandin Lopes	Educação Estatística no Curso de Licenciatura em Matemática.
 https://catalogodeteses.capes.gov.br/catalogo-teses/#/		
ANO	AUTORES(S)	TÍTULO
D1	2012 Gustavo Salomão Viana.	Atitude e Motivação em relação ao desempenho acadêmico de alunos do curso de graduação em Administração em disciplinas de Estatística.
D2	2012 Marcia Elisa Berlikowski.	Análise das atitudes e imagem em relação à Estatística: um estudo comparativo com alunos da graduação.
D3	2013 Geisiane Rodrigues dos Santos.	A Educação Estatística no Ensino Superior, o trabalho com Projetos e o uso de Tecnologias.
T1	2014 Geraldo Bull da Silva Junior.	O Ensino de Estatística na Formação Inicial do Engenheiro de Produção.
T2	2015 Fernando Carvalho Dalbão.	As contribuições do Raciocínio Estocástico para a Formação do Economista.
 http://bdtd.ibict.br/vufind/		
ANO	AUTORES(S)	TÍTULO
T3	2012 Gislaine Donizeti Fagnani da Costa.	A metodologia de projetos como uma alternativa para ensinar estatística no Ensino Superior.
T4	2013 Claudinei Aparecido da Costa.	Educação matemática nos cursos superiores de tecnologia: revelações sobre a formação estatística.
D4	2015 Leonardo Bertholdo de Assis.	A formação do usuário de estatística pelo desenvolvimento da literacia estatística, do raciocínio estatístico e do pensamento estatístico através de atividades exploratórias.
 http://www.theses.fr/		
ANO	AUTORES(S)	TÍTULO
T5	2010 Bernard Ernest Coutanson.	La question de l'éducation statistique et de la formation de l'esprit statistique à l'école primaire en France: étude exploratoire de quelques caractéristiques de situations inductrices d'un enseignement de la statistique au cycle III.

(Continua)

	ANO	AUTORES(S)	TÍTULO
T6	2013	Vladimir Lira Veras Xavier de Andrade.	Os conceitos de medidas de tendência central e de dispersão na formação estatística no ensino médio no Brasil e na França. Abordagem exploratória no quadro da teoria antropológica do didático e da Teoria dos Campos Conceituais.
T7	2015	Marthe-Aline Jutand.	Études des phénomènes de transposition didactique de la statistique dans le champ universitaire et ses environnements: une contribution à la pédagogie universitaire.
T8	2016	Diane Diaz.	Les facteurs influençant la réussite des activités collaboratives médiées par les TICE dans une situation de formation universitaire à la statistique.

Fonte: Do autor (2018).

O quinto e o sexto passo, indicados para reger o estado da arte, são o da leitura das produções selecionadas para uma síntese preliminar, e a organização das tendências, que serão descritas a seguir.

1.1. Apreciação das tendências das produções selecionadas

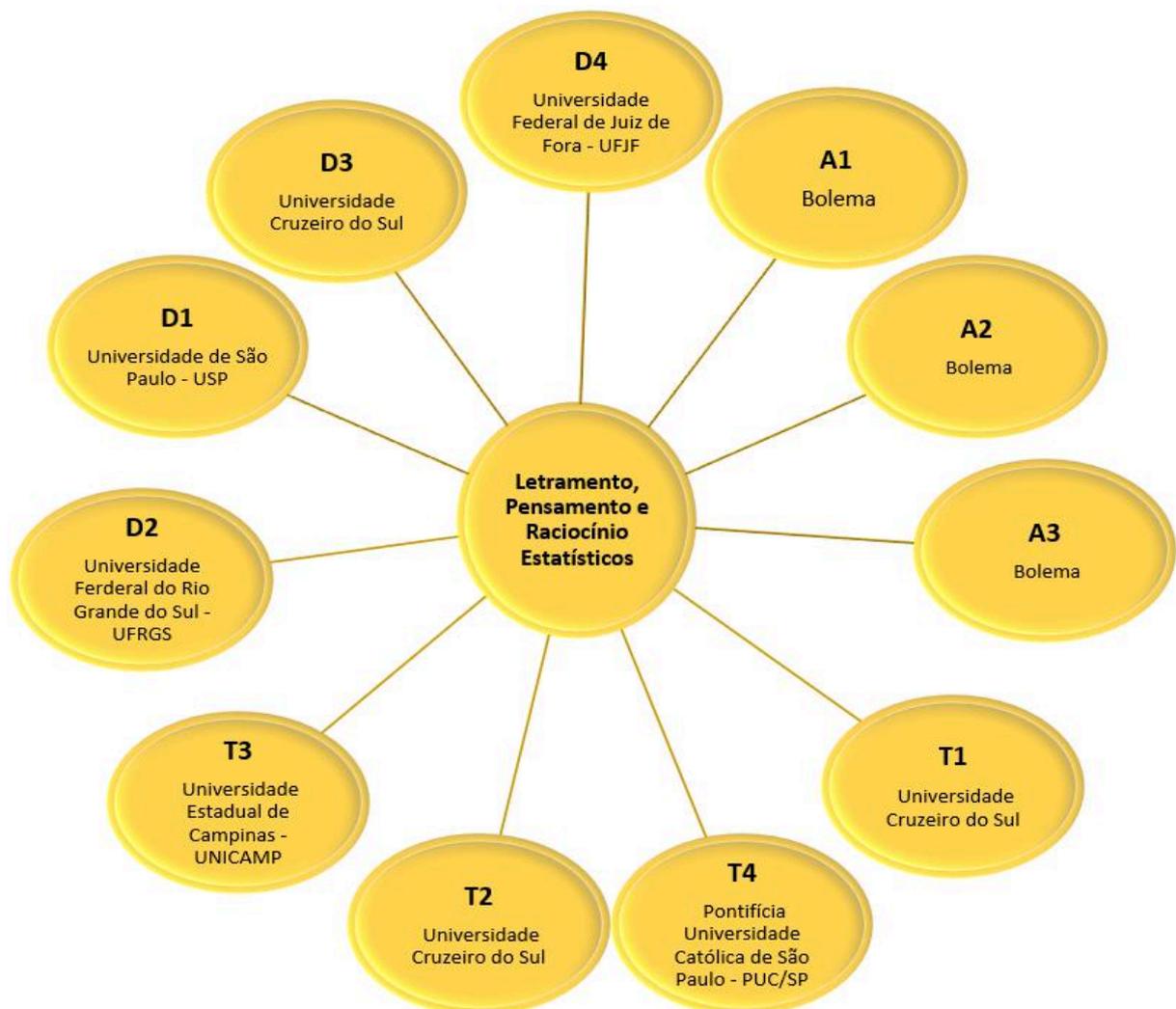
A análise e a compreensão dos estudos já realizados sobre um determinado tema são necessárias para o processo de evolução da ciência, afim de promover novas discussões na busca de possíveis elementos que possam ser abordados. Para uma melhor observação e análise quanto aos trabalhos investigados, construímos algumas figuras que identificam os eixos que foram apurados como critérios de seleção e as produções envolvidas.

O primeiro eixo é relativo ao tema que versa sobre letramento, pensamento e raciocínio estatísticos como mostra a Figura 2.

A primeira produção analisada, aqui denominada A1, trata de um artigo intitulado “Evidências de Validade de uma Prova de Letramento Estatístico: um estudo com estudantes universitários de cursos tecnológicos”, Kataoka *et al.* (2011) buscaram evidências de validade numa prova de letramento estatístico que foi aplicada a 236 universitários oriundos de cursos tecnológicos. Para tanto, as autoras utilizaram um questionário que contou com questões relativas a medidas de tendência central, probabilidade, variabilidade e interpretação de tabela de dupla entrada. O objetivo principal era identificar esse questionário como instrumento capaz de medir o nível de letramento estatístico dos estudantes concluintes da disciplina de Estatística em cursos tecnológicos. As análises foram qualitativas e apoiadas em resultados obtidos com o uso da Teoria de Resposta ao Item - TRI, mais especificamente pelo modelo de Rasch. Durante

as análises foram constatados indícios de relação entre o nível de letramento estatístico e as habilidades matemáticas dos estudantes. A escolha desse artigo, para a nossa análise, justificase pelo fato de estarmos preocupados com uma melhor compreensão referente ao processo de desenvolvimento do raciocínio estatístico. No entanto, apesar das boas ideias contidas no artigo, as autoras concluem ser necessário um maior estudo em relação as questões que foram aplicadas para que o questionário possa auxiliar, de fato, na identificação do nível de letramento estatístico.

Figura 2 - Eixo: letramento, pensamento e raciocínio estatísticos



Fonte: Do autor (2018).

O artigo A2, intitulado “Trabalho de Projetos no Processo de Ensinar e Aprender Estatística na Universidade”, Junior e Campos (2011) relataram os resultados de uma pesquisa de mestrado que se preocupou em contribuir com o crescimento científico dos alunos em relação a Estatística. Os autores enfatizam a utilização de novas tecnologias como recurso para

estabelecer conexões do conteúdo de Estatística com situações do cotidiano. Para tanto, foram aplicados situações-problema em estudantes de um curso de Licenciatura e Bacharelado em Matemática onde ficou constatado um melhoramento da compreensão Estatística quando essa é trabalhada com algum tipo de software. O artigo expõe a necessidade encontrada nos alunos de compreenderem o porquê e para quê servem os conteúdos vistos em sala de aula, enfatizando a importância de contextualizar a teoria. Dessa forma, Junior e Campos (2011), ressaltam a importância de um letramento estatístico que priorize vincular o conhecimento estatístico com a formação de um cidadão crítico, detentor de conhecimentos que irão refletir em sua vida pessoal e profissional. O trabalho A2 nos chamou atenção pelo fato de estar preocupado com a contextualização da estatística em situações não convencionais aos que geralmente são exercidas em sala de aula, considerando, ainda, a Estatística como instrumento de formação de um cidadão crítico, capaz de reconhecer e compreender informações estatísticas nos mais variados meios de comunicação. Além disso, defende o uso da tecnologia como ferramenta para melhor compreensão da Estatística.

O artigo A3, intitulado “Educação Estatística no Curso de Licenciatura em Matemática”, Lopes (2013) incentivou um maior desenvolvimento do pensamento estatístico dos futuros professores de matemática para auxiliá-los na percepção sobre o trabalho a ser desenvolvido com seus futuros alunos. A publicação defende uma maior apropriação do letramento estatístico e promove uma discussão sobre como a concepção da Estatística deve ser postulada. Para tanto, sugere que a programação de disciplinas de Estatística seja revista e que a sua apropriação vá além da resolução de problemas, ou seja, deve também promover a realização de atividades de investigação e problematização de situações diversas. A autora declara a necessidade de entender a Estatística como uma ciência de análise de dados e ter clareza sobre a interface dela com a matemática.

Em D1, temos uma dissertação intitulada “Atitude e Motivação em relação ao desempenho acadêmico de alunos do curso de graduação em Administração em disciplinas de Estatística”, Viana (2012) defendeu o fato de que a Estatística deve ser ensinada por meio de aplicações que fazem sentido na formação do estudante. O autor revela que o curso de Administração forma profissionais que devem possuir um bom letramento estatístico, uma vez que vai ao encontro das competências e das habilidades focadas no processo decisório do administrador. Dessa forma, 278 estudantes de Administração foram submetidos a uma investigação do modo como se dá a interação da atitude perante a Estatística, e da motivação acadêmica com o desempenho acadêmico do estudante nas disciplinas de Estatística. Foi realizado um estudo quantitativo por meio de uma Escala de Atitudes dos alunos frente a

Estatística. Com a pesquisa o autor concluiu que, de modo geral, os estudantes que conseguiram observar as aplicações da Estatística em disciplinas mais específicas do curso de Administração, obtiveram maiores pontuações, tanto em relação à atitude perante a Estatística como em relação à motivação acadêmica. Além de tratar da questão do letramento estatístico, a dissertação D1 nos chamou atenção pelo fato de trabalhar com uma Escala de Atitude, tendo em vista que nossa pesquisa também fará referências sobre as atitudes dos estudantes pesquisados frente a disciplina de Estatística.

A dissertação D2, intitulada “Análise das atitudes e imagem em relação à Estatística: um estudo comparativo com alunos da graduação”, Berlikowski (2012) tratou de uma pesquisa que objetivou analisar a imagem e a atitude que os alunos de cursos superiores possuem em relação a Estatística. Para tanto, foi realizado, por meio de um questionário, um estudo comparativo antes e depois de cursarem a disciplina de Estatística. A investigação contou com 101 estudantes matriculados na disciplina de Estatística nos cursos de graduação de Administração, Ciências Contábeis e Sistemas de Informação de uma faculdade particular do Rio Grande do Sul. Para análise dos dados, foram utilizadas ferramentas da Estatística Descritiva e confrontados com pesquisas que já trataram do mesmo tema. O autor revela que, quando a Estatística é aplicada em situações cotidianas vivenciadas pelos estudantes, eles tendem a ter um crescimento em relação ao letramento estatístico. Percebemos, mais uma vez, a inquietude de trabalhos que evidenciam a necessidade que os estudantes possuem em reconhecerem a aplicabilidade dos conceitos estatísticos em situações cotidianas e não somente de exercícios teóricos de fixação do conteúdo.

Na dissertação D3, intitulada “A Educação Estatística no Ensino Superior, o trabalho com Projetos e o uso das Tecnologias”, Santos (2013) investigou como os estudantes do Ensino Superior mobilizam conhecimentos estatísticos no curso de Administração com auxílio de recursos tecnológicos. A pesquisa foi norteada pela seguinte proposição: Quais as contribuições que o trabalho com Projetos em Modelagem Matemática, mediado pelo uso de tecnologias, pode trazer à aprendizagem estatística dos estudantes de um curso de Administração? Para responder a essa pergunta o autor utilizou uma abordagem qualitativa, com o intuito de se considerar mais o processo do que o resultado. Construiu-se, então, um estudo de caso, a partir dos cenários de investigação constituídos pelos Projetos de Investigação em Administração desenvolvidos por alunos atores no seu processo de ensino e aprendizagem. Os resultados indicaram que a prática e a teoria proporcionam uma aprendizagem permeada por significados decorrentes de uma proposta pedagógica em que o professor possibilita ao estudante ser coautor do processo de ensino e aprendizagem. Santos (2013) afirma que o trabalho com Projetos, cujos

temas sejam de interesse do estudante e relacionados ao cotidiano, permite mobilizar os conhecimentos estatísticos e matemáticos de forma a ressignificá-los. A inserção das tecnologias favorece na resolução dos cálculos matemáticos e na análise sobre os dados e, dessa forma, o autor conclui que existe um avanço no letramento estatístico.

Em D4, temos uma dissertação intitulada de “A formação do usuário de estatística pelo desenvolvimento da literacia estatística, do raciocínio estatístico e do pensamento estatístico através de atividades exploratórias”, Assis (2015) objetivou observar e analisar como atividades exploratórias podem contribuir para a formação do usuário de Estatística, tendo em vista uma posição crítica e ativa diante de informações estatísticas. A pesquisa foi realizada com estudantes de um curso de Administração que já haviam estudado a disciplina de Estatística básica anteriormente. Para tanto, foi utilizado um questionário com três atividades de um livro didático reconhecido pelo autor como instrumento de percepção do pensamento estatístico. Por meio da análise dos registros das resoluções, obteve como resultados indícios positivos no desenvolvimento do pensamento estatístico que foi julgado como essencial para a formação do usuário de Estatística.

A tese T1, intitulada de “O Ensino de Estatística na Formação Inicial do Engenheiro de Produção”, Junior (2014) investigou as necessidades do futuro engenheiro referentes ao conhecimento estatístico. Para tanto, o autor construiu dados para análise por meio de questionários e entrevistas semiestruturadas. Na análise dos questionários aplicados, o autor pontua as deficiências de raciocínio estatístico e literacia, mencionando as dificuldades em relação a interpretações gráficas e dos conceitos de variância e desvio padrão. Nas entrevistas realizadas, ficou evidenciada a necessidade de se articular a Estatística a outros componentes curriculares desde o início do curso. A pesquisa destacou três diferentes percepções em relação a Estatística por parte dos estudantes, quais sejam: foi considerada como uma Ciência de análise de dados, foi compreendida como apenas um item na formação do Engenheiro e, terceira e última, foi vista como uma disciplina de caráter matemático. Junior (2014) destaca que os estudantes relataram que começaram a compreender melhor os conceitos estatísticos a partir do momento em que passaram a utilizá-los como instrumentos de resolução de problemas e descrição de processos de outras disciplinas. A pesquisa nos chamou atenção, para análise, pelo fato de mencionar que os estudantes consideraram a Estatística mais relevante após o conhecimento da Estatística Inferencial, onde foi possível analisar dados e tomar decisões.

Na tese T2, intitulada “As contribuições do Raciocínio Estocástico para a Formação do Economista”, Dalbão (2015) apresentou uma pesquisa qualitativa e interpretativa cujo objetivo foi investigar as contribuições do raciocínio estocástico para a formação de futuros economistas

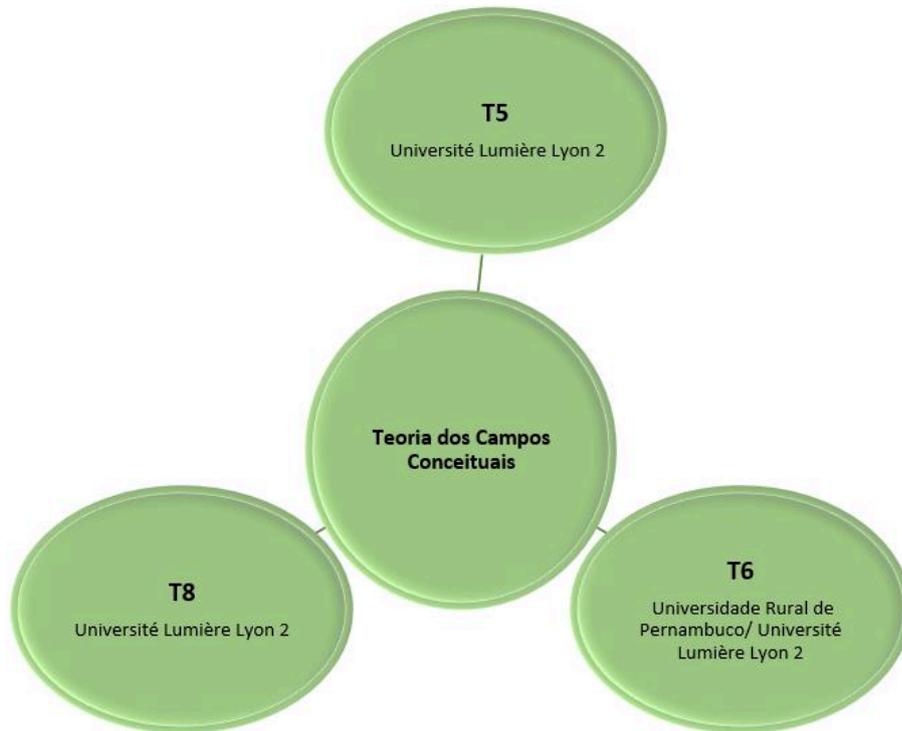
do curso superior em Ciências Econômicas. Para tanto, a pesquisa analisou as entrevistas que foram realizadas com oito estudantes do quinto semestre que participaram da construção de um projeto envolvendo os temas das áreas de microeconomia e macroeconomia que interferem no seu cotidiano. Os resultados da pesquisa evidenciaram que os estudantes mobilizaram conhecimentos matemáticos e estatísticos para realizarem a análise das problemáticas e, nesse movimento, adquiriram novos conhecimentos estocásticos com o desenvolvimento do raciocínio necessário para a tomada de atitudes nas situações propostas.

Em T3, temos uma tese intitulada “A metodologia de projetos como uma alternativa para ensinar estatística no Ensino Superior”, Costa (2012) investigou uma alternativa metodológica para ensinar Estatística no Ensino Superior focando aspectos afetivos presentes na relação professor-aluno, acreditando na mudança dos paradigmas científicos que procuram ressignificar o papel das emoções no pensamento humano e na interação entre cognição e afetividade. A pesquisa aconteceu num curso superior de Nutrição, onde foi realizado um projeto interdisciplinar envolvendo disciplinas específicas do curso de Nutrição e Bioestatística. O material gerado para a análise foi constituído de observações dos depoimentos dos estudantes que trouxe indícios de melhores contribuições para o desenvolvimento do raciocínio estatístico em situações semelhantes às do exercício da profissão de Nutricionista. Dessa forma, com a aplicação de atividades mais práticas para o curso de Nutrição, a pesquisa assinala que houve uma modificação e ressignificação de fatores sociais e afetivos, crenças e atitudes negativas com relação à matemática, adquiridas ao longo da escolaridade.

A última pesquisa explorada para esse primeiro eixo analisado, foi a tese T4, intitulada “Educação matemática nos cursos superiores de tecnologia: revelações sobre a formação estatística”, Costa (2013) objetivou compreender a configuração e as potencialidades da Estatística na formação dos estudantes das áreas tecnológicas. Para tanto, o autor investigou o surgimento da Educação Estatística e suas competências que compõem o núcleo: o letramento, o raciocínio e o pensamento estatísticos. A pesquisa foi de natureza qualitativa, analisando as entrevistas realizadas com os estudantes. Os resultados apontaram a inexistência da utilização de softwares estatísticos que poderiam contribuir, de forma positiva, para melhorar a compreensão e aplicação dos conceitos e ferramentas que a Estatística oferece para a formação universitária.

O segundo eixo observado, trata da Teoria dos Campos Conceituais envolvendo a temática da Estatística. A Figura 3 foi construída com as produções selecionadas.

Figura 3 - Eixo: Teoria dos Campos Conceituais



Fonte: Do autor (2018).

Na tese T5, intitulada “La question de l’éducation statistique et de la formation de l’esprit statistique à l’école primaire en France: étude exploratoire de quelques caractéristiques de situations inductrices d’un enseignement de la statistique au cycle III”⁷, Coutanson (2010) abordou a Teoria dos Campos Conceituais envolvendo uma análise realizada em livros didáticos de matemática utilizados pelos estudantes do Ensino Fundamental da França. A pesquisa procurou explorar como a Estatística é ensinada nos anos iniciais de escolaridade, promovendo uma discussão sobre as conceitualizações básicas para o entendimento inicial da Estatística.

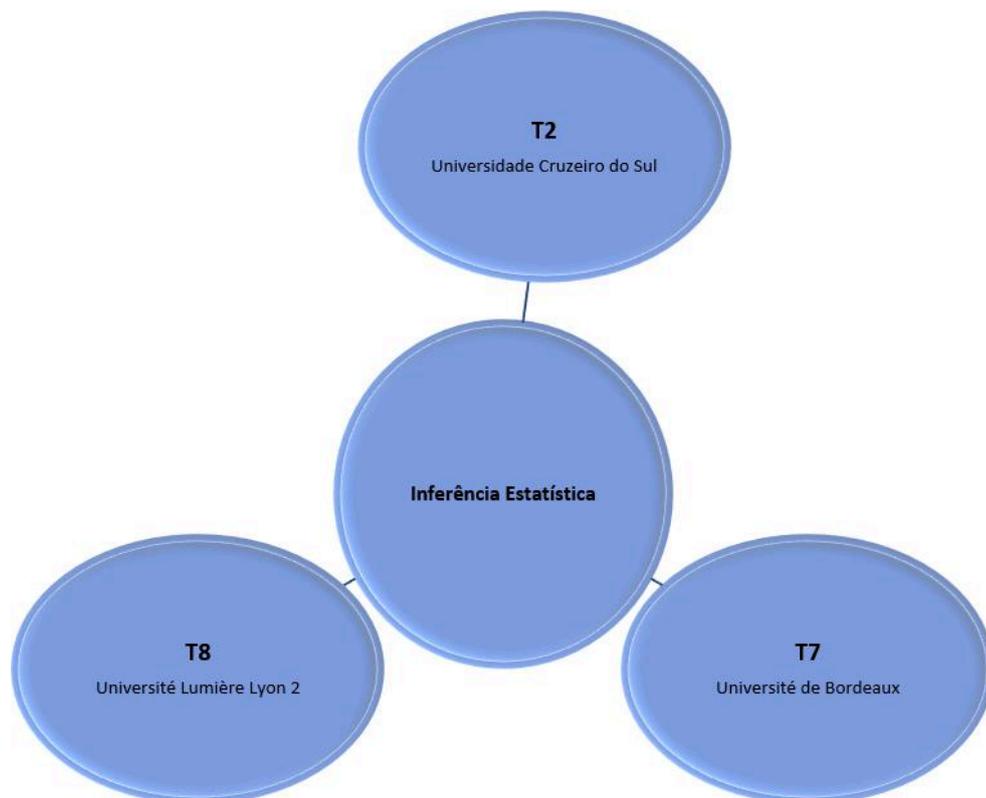
Em T6 temos a tese intitulada “Os conceitos de medidas de tendência central e de dispersão na formação estatística no ensino médio no Brasil e na França. Abordagem exploratória no quadro da teoria antropológica do didático e da Teoria dos Campos Conceituais”, Andrade (2013) propôs uma investigação nos livros didáticos utilizados para o Ensino de Estatística no Ensino Médio do Brasil e da França que tratam da Estatística Descritiva. A pesquisa aborda a Teoria dos Campos Conceituais no sentido de averiguar a necessidade de explorar os conceitos estatísticos com uma diversidade de situações para que o estudante possa tomar posse de um novo conceito.

⁷ Tradução nossa: A questão da educação estatística e a formação do espírito estatístico na escola primária na França: estudo exploratório de algumas características das situações que induzem um ensino de estatística no ciclo III.

A produção T8 é uma tese intitulada “Les facteurs influençant la réussite des activités collaboratives médiées par les TICE dans une situation de formation universitaire à la statistique”⁸, Diaz (2016) investigou o ensino da Estatística em cursos universitários de ciências sociais e humanas promovendo uma discussão sobre as principais dificuldades e obstáculos encontrados pelos estudantes. A pesquisa salienta que a disciplina de Estatística é aplicada para um público heterogêneo que, muitas vezes, não percebe o seu grau de importância e utilidade, fomentando a necessidade de empregar diversas situações onde o estudante possa melhor construir os conceitos estatísticos. Dessa forma, a autora fundamenta sua pesquisa utilizando a Teoria dos Campos Conceituais que proporciona uma concepção para a construção de novos conceitos.

O terceiro e último eixo observado foi o que apresentou, no corpo das produções analisadas, a Estatística Inferencial, como retrata a Figura 4.

Figura 4 - Eixo: Inferência Estatística



Fonte: Do autor (2018).

⁸ Tradução nossa: Fatores que influenciam o sucesso de atividades colaborativas relacionadas a TIC em uma situação de formação estatística com base universitária.

A tese T2, já observada no eixo que versou sobre letramento, pensamento e raciocínio estatísticos, apresenta, em sua abordagem, a preocupação com o tratamento da Estatística Inferencial, fazendo uso aplicado desse conteúdo em situações relevantes para a formação dos estudantes de Ciências Econômicas. Versando sobre o raciocínio estocástico, a pesquisa enfatiza a importância do conteúdo de probabilidade para a tomada de decisões dos futuros profissionais.

Em T7, temos a tese intitulada “Études des phenomenes de transposition didactique de la statistique dans le champ universitaire et ses environnements: une contribution à la pédagogie universitaire”⁹, Jutand (2015) salientou a Estatística como a ciência de análise de dados e tomada de decisão, onde a universidade deve contribuir para a formação dos futuros profissionais. A pesquisa enfatiza a importância de treinar estudantes na disseminação e tradução de resultados de estudos estatísticos.

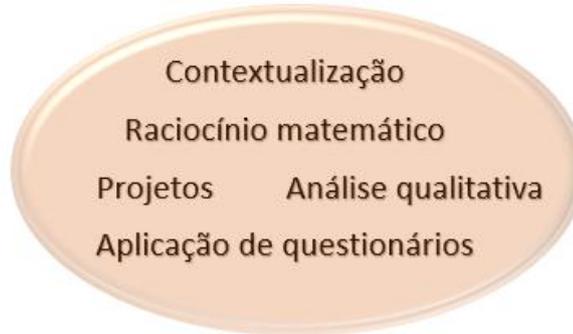
A última produção analisada, T8, já examinada no eixo que versou sobre a Teoria dos Campos Conceituais, trata em seu texto sobre a Estatística Inferencial fazendo uso desse conteúdo estatístico em questões que foram aplicadas para promover uma análise das resoluções e das maiores dificuldades encontradas pelos estudantes. Destaca, ainda, a importância de trabalhar com as ferramentas da Estatística Inferencial para a formação de profissionais críticos e aptos a realizarem análises e interpretações estatísticas nas mais diversas situações a serem encontradas.

1.2. Conclusões sobre as tendências observadas nas produções selecionadas

O último passo para a realização do estado da arte proposto por Romanowski e Ens (2006) sugere a elaboração de conclusões preliminares em relação às principais tendências encontradas nos trabalhos que foram selecionados para análise. Desse modo, após a breve descrição dos trabalhos, construímos a Figura 5 que é composta pelas ações que observamos e percebemos serem as maiores tendências encontradas em relação aos três eixos que foram examinados.

⁹ Tradução nossa: Estudos dos fenômenos de transposição didática da estatística no campo universitário e seus ambientes: uma contribuição para a pedagogia universitária.

Figura 5 - Ações observadas nos trabalhos analisados



Fonte: Do autor (2018).

Contemplamos, no eixo que tratou do letramento, pensamento e raciocínio estatísticos, que os autores dos trabalhos A1, A2, A3, D2 e T1, atentam para a importância da contextualização do ensino da Estatística por meio da aplicação de situações-problema que retratem cenários do cotidiano vivenciados pelos estudantes para obterem informações que possam direcionar na busca de um possível entendimento sobre a construção do letramento estatístico. Segundo Junior e Campos (2011) o ensino da Estatística deve priorizar a formação de um cidadão crítico com conhecimentos que refletirão em sua vida pessoal e profissional, estabelecendo conexões do conteúdo de Estatística com situações do cotidiano.

Na mesma direção, porém com uma proposta de utilização de situações-problema voltadas para a área específica da formação universitária, os trabalhos D3, T2 e T3, carregam em seus textos a ideia do ensino da Estatística por meio de projetos para um melhor entendimento sobre a construção do letramento estatístico. Esses projetos foram desenvolvidos com o engajamento interdisciplinar onde a Estatística aparece como ferramenta de apoio para auxiliar o futuro profissional em suas ações técnicas. De acordo com Costa (2012) a aplicação de conceitos estatísticos em situações semelhantes às do exercício de sua profissão pode proporcionar, aos estudantes, uma maior compreensão desses conceitos.

Apuramos, ainda no primeiro eixo, que os trabalhos D1, D2 e T3, entendem que a afetividade e a interpretação de uma escala de atitude em relação a disciplina de Estatística, podem colaborar para uma melhor compreensão na formação do letramento estatístico. Viana (2012) e Costa (2012) refletem que a identificação do relacionamento da atitude, assim como a afetividade e a motivação acadêmica em relação a Estatística, possibilitam levantar fortes fatores de aprendizagem que deverão ser considerados pelos docentes.

Observamos, também, que os resultados obtidos das análises dos trabalhos investigados, revelam a importância do letramento estatístico, fazendo uma comparação e diferenciação com o raciocínio matemático. Lopes (2013) salienta que a Estatística fornece meios para lidar com

dados que levem em conta a onipresença da variabilidade, o que a diferencia da matemática e de outras ciências. Entendemos e pontuamos, desde já, que muitos problemas estatísticos não possuem uma única solução matemática, pois começam com um questionamento e terminam com uma ponderação fundamentada em resultados teórico-práticos. A ciência estatística requer um tipo diferente de pensar, pois os dados não são somente números, são números inseridos em um contexto. Diante dessas ponderações realizadas no primeiro eixo analisado, refletimos que nossa pesquisa está ancorada em ações contundentes, tendo em vista que aplicamos, para os nossos estudantes, dois questionários com situações-problema envolvendo situações que discorrem do nosso objeto de estudo, que é a Estimação Estatística. Além disso, um dos questionários é composto por questões que observam a relação dos estudantes com a estatística. Dessa forma, entendemos estar contribuindo e reforçando a importância da contextualização e a influência da afetividade para aprendizagem da Estatística.

No segundo eixo explorado, a Teoria dos Campos Conceituais envolvendo a temática da Estatística, os trabalhos T5 e T6 examinaram livros didáticos para refletirem sobre a concepção introdutória dos conceitos estatísticos no ensino básico. Os autores utilizam a Teoria dos Campos Conceituais para fundamentar a pesquisa onde, segundo Andrade (2013), a observação dessa teoria proporciona a compreensão da importância de trabalharmos com uma diversidade de situações para que os estudantes possam tomar posse de um novo conceito. O trabalho T8, utilizou a teoria para promover uma discussão em relação aos principais obstáculos enfrentados pelos estudantes na resolução de problemas estatísticos, onde, segundo Diaz (2016), estão relacionados com a história dos estudantes, a formação cultural anterior ao Ensino Superior e a afetividade. Nesse segundo eixo analisado, podemos perceber uma carência de trabalhos que conectam a Teoria dos Campos Conceituais com a Estatística. Além disso, das três produções analisadas, apenas a T8 faz referência ao Ensino Superior. Diante disso, nossa proposta foi utilizar a Teoria dos Campos Conceituais para analisar e interpretar as respostas dos estudantes para as questões que propusemos. O último eixo analisado, que foi a Estatística Inferencial, nos mostrou a preocupação dos trabalhos T2, T7 e T8 com a formação crítica dos futuros profissionais em relação a tomada de decisões. Jutand (2015) defende que é necessário desenvolver capacidades para explicar as etapas de um processo estatístico e comunicar, de forma compreensível, para fins de difusão e apropriação. O pensar estatístico requer, inúmeras vezes, ideias de variabilidade, distribuição, estimação e testes de hipóteses. Apesar das ramificações existentes dentro da Estatística Inferencial, os trabalhos analisados fazem um aparato geral da indução estatística, mas não especificam tópicos dos conteúdos analisados. Diante disso, nossa pesquisa diferencia-se por tratar, especificamente, da Estimação Estatística,

onde extraímos informações pertinentes para contribuir com a Educação Estatística. Outra tendência recorrente é a análise que a maioria dos trabalhos selecionados empregaram. Observamos que A1, A2, A3, D3, D4, T1, T2 e T4 foram produzidos utilizando apenas uma análise qualitativa sobre seus dados. Os demais trabalhos, como D2, T5, T6 e T8, utilizaram análises qualitativas e quantitativas e, apenas D1, fez somente uso de uma análise quantitativa. Em nossa pesquisa, abarcamos uma análise qualitativa e quantitativa para nossos dados, envolvendo, como já mencionado, a Teoria dos Campos Conceituais aliada com a utilização do software estatístico SPAD, para análise descritiva, e do software CHIC, para análise estatística implicativa.

Tendo em vista a qualidade inquestionável dos trabalhos examinados, compreendemos que nossa pesquisa, que explorou a construção do raciocínio estatístico por meio da Estimação Estatística, apresenta-se como colaboradora para o crescimento da Educação Estatística diante dos dados obtidos.

2. A Estatística: visões conceituais e educacionais

É notório que aplicações exigentes de conhecimentos estatísticos estão sendo cada vez mais presentes em nosso cotidiano por meio de diversas fontes de informações que nos são transmitidas para comunicar, orientar, educar e melhor compreender o mundo atual. A utilização da Estatística está difundida nos diferentes ramos do conhecimento humano, tendo em vista sua aplicabilidade em construir modelos de interpretações para as mais variadas informações. Esses modelos são gerados a partir de concepções matemáticas, utilizando técnicas para o levantamento de dados, organização, análise e interpretação dos resultados de alguma informação a ser estudada e revelada. Além disso, uma participação crítica e reflexiva diante dessas informações recebidas, combina com a engrenagem que movimenta os saberes estatísticos. Mas afinal, o que é a Estatística? Sabemos que o ensino da Estatística é o propulsor dos conhecimentos necessários para se conseguir observar e atuar, de forma proativa, nas diversas situações em que a Estatística se revela. Mas o que é uma educação estatística? Diante disso, propomos uma breve discussão em relação a essência do significado de Estatística enquanto ciência e uma outra discussão sobre a principal fonte de aquisição para uma boa formação estatística que é a educação. Encerramos com a proposta de uma investigação em relação a um dispositivo pedagógico que será definido e enquadrado nos moldes que a pesquisa se propôs: a formação do raciocínio estatístico por meio da estimação estatística.

2.1. A Estatística enquanto ciência

Principiamos nossas reflexões reproduzindo Régnier (1997) que menciona um texto epistemológico de Antoine Augustin Cournot (1801-1877)¹⁰ sobre Estatística:

A estatística é uma ciência moderna: as ideias dos mais antigos não estavam inclinadas para o trabalho de precisão; os meios de pesquisa e de comunicação faltavam; finalmente (o que é mais surpreendente), apesar da variedade de especulações filosóficas, eles não pareciam suspeitar da existência de um princípio de compensação que sempre acabava manifestando influência nas causas regulares e permanentes, atenuando cada vez mais as causas irregulares e fortuitas. Hoje em dia, pelo contrário, a estatística se desenvolveu de maneira exuberante; e é preciso cuidar algumas aplicações prematuras e abusivas que podem vir a desacreditá-la por algum tempo e atrasar a era tão desejável onde os dados de uma experiência servirão como base para teorias cujo objeto são as várias partes da organização social. De fato, entende-se principalmente por estatística (como a etimologia indica) a coleção de dados que descrevem a aglomeração de homens em sociedades políticas: mas para nós a palavra assumiu um significado mais amplo. Entendemos por estatística, a ciência cujo objetivo é coletar e coordenar numerosos fatos que são sensivelmente independentes das anomalias do acaso e que denotam a existência de causas regulares cuja ação foi combinada com a de causas fortuitas (COURNOT, 1843, tradução nossa¹¹).

Cournot (1843) complementa dizendo que considera a Estatística uma ciência que observa os dados, utilizando os números como ferramentas que serão interpretados com o uso das técnicas de teorias probabilísticas. O objetivo da Estatística é observar e se aprofundar, o máximo possível, no conhecimento de uma determinada situação analisada, investigando, por meio de uma discussão racional, os dados e os acontecimentos que podem afetá-los. Régnier (1996) e Droesbeke e Tassi (1997) oferecem uma definição para Estatística trazida por Ronald Aylmer Fisher (1890-1962) considerado um dos maiores estatísticos da primeira metade do século XX, expondo que o objetivo da Estatística é o de reduzir dados. Uma massa de dados deve ser substituída por uma pequena quantidade que represente corretamente essa massa e, na

¹⁰ Matemático, filósofo e economista francês, propulsor das teorias marginalistas, conhecido por seus estudos sobre a oferta e a demanda nos termos da competição monopolística. Foi professor de Análise Matemática da Universidade de Lyon em 1834, e Reitor da Academia de Dijon de 1854 a 1862.

¹¹ “La statistique est une science toute moderne : le génie des anciens ne se portait pas volontiers vers des travaux de précision ; les moyens de recherche et de communication leur manquaient ; enfin (ce qui surprend davantage), malgré la variété de leurs spéculations philosophiques, ils ne paraissent pas avoir soupçonné l’existence d’un principe de compensation qui finit toujours par manifester l’influence des causes régulières et permanentes, en atténuant de plus en plus celle des causes irrégulières et fortuites. De nos jours, au contraire, la statistique a pris un développement en quelque sorte exubérant; et l’on n’a plus qu’à se mettre en garde contre les applications prématurées et abusives qui pourraient la décréditer pour un temps, et retarder l’époque si désirable où les données de l’expérience serviront de bases certaines à toutes les théories qui ont pour objet les diverses parties de l’organisation sociale. Em effet, l’on entend principalement par statistique (comme l’indique l’étymologie) le recueil des faits aux quels donne lieu l’agglomération des hommes em sociétés politiques: mais pour nous le mot prenda une acception plus étendue. Nous entendrons par statistique, la science qui a pour objet de recueillir et de coordonner des faits nombreux sensiblement indépendants des anomalies du hasard, et qui dénotent l’existence des causes régulières dont l’action s’est combinée avec celle des causes fortuites” (COURNOT, 1843). Disponível em: <https://gallica.bnf.fr/ark:/12148/bpt6k285042.texteImage>.

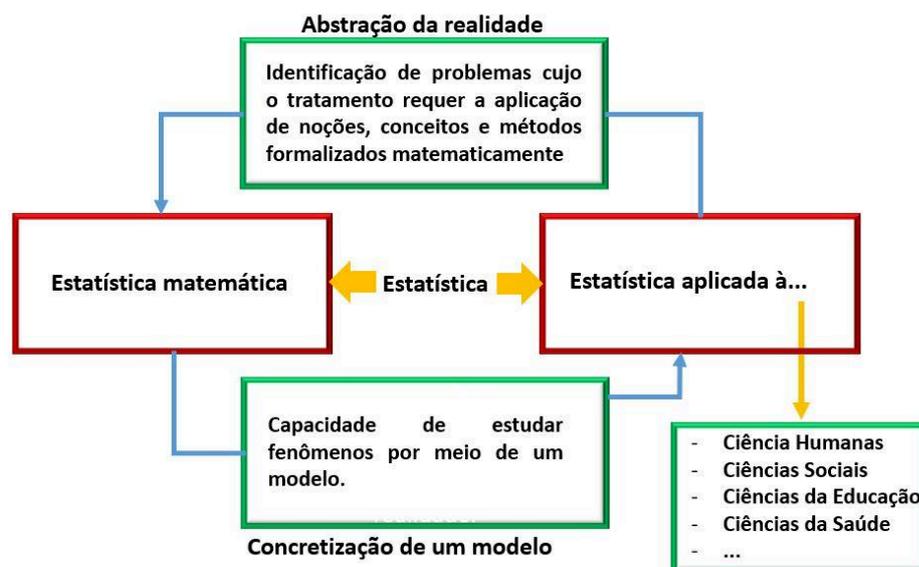
medida do possível, carregar todas as informações relevantes contidas nos dados originais para a realização de análises. Yule e Kendall (1948) apresentam a Estatística como um método científico que trata os dados para medir as propriedades das populações dos fenômenos naturais. Besson (1995) descreve a Estatística com uma visão mais ampla do que simplesmente a de uma ciência que analisa dados, mas também a coloca como ciência preocupada com questões interdisciplinares que a torna cada vez mais expansiva.

Diante das contribuições em relação ao entendimento da Estatística que aos poucos foram se incorporando para a formação de sua concepção enquanto ciência, buscamos uma definição que esteja de acordo com as tendências que essa ciência engloba nos dias atuais, encontrando em Régnier (1996):

A Estatística é a ciência que opera o estudo de métodos baseados em modelagem matemática, dos modos de utilização e processamento de dados, ou seja, informações, com o objetivo de conduzir e apoiar uma reflexão ou tomar uma decisão concreta sujeita aos caprichos do incerto. (tradução nossa)¹²

O autor supracitado apresenta um sucinto esquema para retratar a definição da Estatística enquanto ciência, observando os elementos que interagem para a sua aplicação, como mostra a Figura 6 a seguir:

Figura 6 - Esquema sobre a definição da Estatística



Fonte: Adaptado e traduzido de Régnier (1997).

¹² La statistique est la science qui procède à l'étude méthodique à partir de modélisations mathématiques, des modes d'utilisation et de traitement de données, c'est à dire de l'information, dans le but de conduire et d'étayer une réflexion ou de prendre une décision en situation concrète soumise aux aléas de l'incertain.

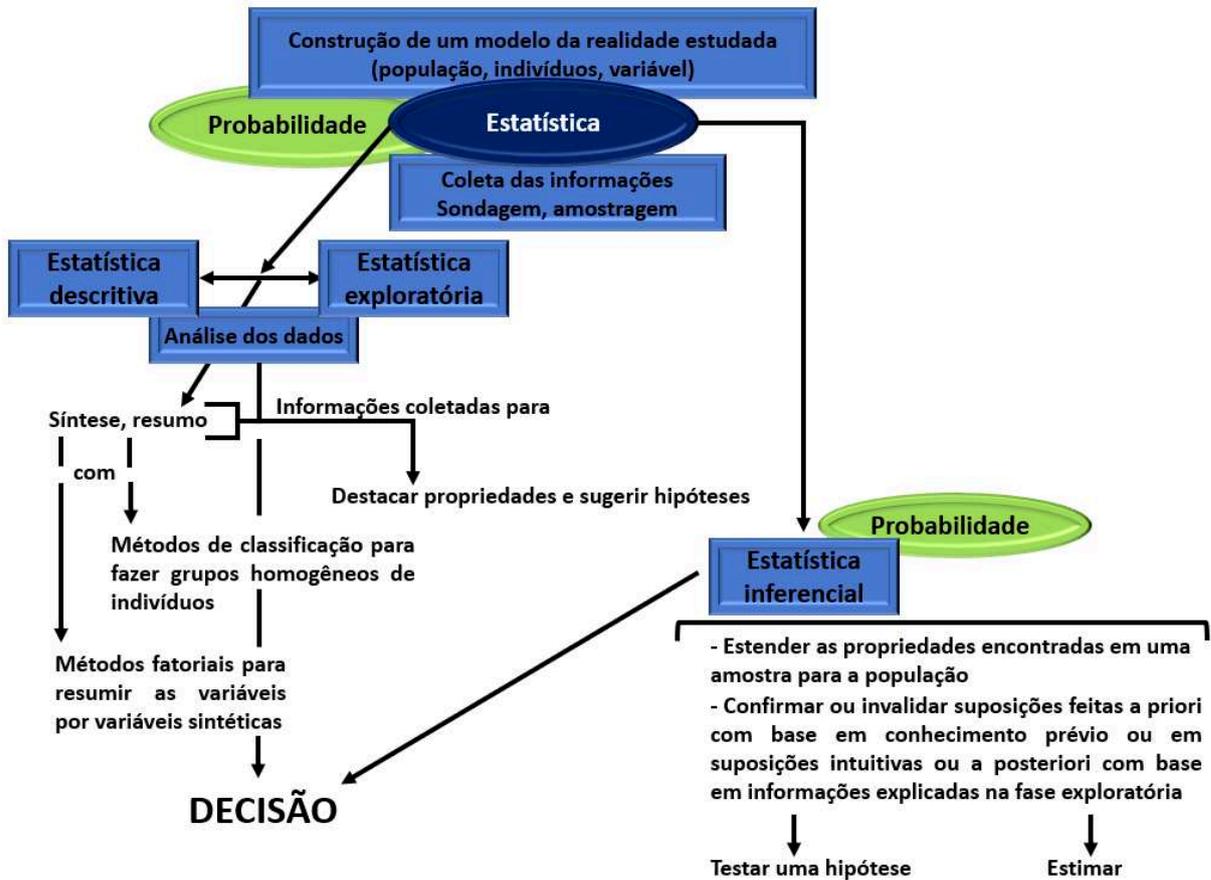
Ao propor o esquema, o autor mostra a Estatística como uma implicação dialética entre as aplicações da Estatística nas diversas áreas do conhecimento com a estatística matemática. Hand (1998) define a estatística matemática como as ideias estatísticas que são formalizadas pela matemática. O uso das ferramentas da matemática é essencial para formalizar modelos estatísticos que serão analisados para a compreensão de diversos fenômenos existentes nas mais variadas áreas do conhecimento. Campos (2007) reconhece a importância da matemática para a Estatística, observando que elas se diferenciam por apresentarem aspectos e valores que são pertinentes na Estatística, como a organização dos dados, a interpretação, a reflexão e a tomada de decisões. Régnier (1996) completa dizendo que se caracterizarmos a matemática como uma ciência do certo, então a Estatística poderia ser distinguida dela como uma ciência da incerteza que busca estabelecer os limites da incerteza.

Dessa forma, a aplicação da Estatística para o estudo de fenômenos das mais variadas áreas, necessita de uma abstração da realidade para a concretização de um modelo real que é formalizado por ferramentas matemáticas e por princípios estatísticos tais como os eventos aleatórios, variabilidade, chance maior ou menor de ocorrência, as quais são importantes para que um indivíduo seja capaz de argumentar com confiança e realizar as conclusões de suas investigações.

Droesbeke e Tassi (1997) ressaltam que a concepção da Estatística moderna pode ser compreendida observando-a em duas grandes áreas: a estatística descritiva e a estatística inferencial. Régnier (1996) compõe um esquema que apresenta os objetivos e as operações atribuídas para essa divisão, como mostra a Figura 7.

Régnier (1996) atribui para a Estatística, como meta central, o propósito de uma ferramenta que dará suporte para a tomada de decisão, seja ela baseada na exploração ou descrição de dados coletados sob circunstâncias conhecidas ou baseada em hipóteses testadas que carregam um certo controle de risco. Diante disso, a associação entre a probabilidade e a Estatística é indispensável enquanto áreas de conhecimento, uma vez que a Estatística utiliza teorias probabilísticas para explicar a frequência da ocorrência de fenômenos, tanto em estudos observacionais quanto em experimentos. No primeiro nível temos a estatística descritiva que estuda os modos de utilização e tratamento dos dados, no sentido da produção e descrição das informações. O segundo nível aponta a estatística inferencial como sendo responsável pela expansão das informações descritas para um domínio de validade não explorada diretamente, considerando um controle de riscos estabelecido por algum critério empregando o raciocínio indutivo.

Figura 7 - Esquema das operações e dos objetivos atribuídos para a Estatística observando sua divisão



Fonte: Adaptado e traduzido de Régnier (1997).

Das constatações apuradas, consideramos a estatística descritiva como o ramo da Estatística que aplica técnicas para descrever, organizar e abreviar um conjunto de dados que poderão ser interpretados pela descrição sem usar induções que estabeleçam correspondências para a massa não explorada diretamente. A estatística inferencial expande os dados descritos para a massa não explorada diretamente fazendo uso de ferramentas probabilísticas que avaliam os riscos dessa correspondência por meio de critérios. Em linhas gerais, podemos dizer que a Estatística, com base na probabilidade, fornece técnicas e métodos de análise de dados que auxiliam o processo de tomada de decisão nos mais variados problemas onde existe incerteza.

Notoriamente, a Estatística, enquanto ciência, vem se propagando dentro do contexto escolar e universitário, tornando-se, progressivamente, uma exigência do mundo moderno diante da necessidade de compreender códigos e linguagens que são expostos nos mais diversos meios para a sociedade. Régnier (2009) completa que “A estatística considerada como uma ciência construída pelos seres humanos ao longo dos séculos, é um objeto de aprendizagem,

particularmente no contexto do ensino escolar, universitário e de formação continuada” (tradução nossa¹³).

A necessidade de uma evolução em relação a compreensão dos estudos relacionados a aprendizagem da Estatística, promoveu um movimento que investiga os elementos que contornam essa ciência, percorrendo discussões sobre a importância do domínio da matemática para a compreensão estatística e a forma como a educação pode colaborar para melhor formar cidadãos aptos a observar, questionar e interpretar situações tanto da vida profissional como no cotidiano. Desse modo, manifesta-se a Educação Estatística aliando princípios matemáticos, educacionais e estatísticos.

Figura 8 - Tríade da Educação Estatística



Fonte: Do autor (2018).

Lopes (2003) considera que a tríade Estatística, Matemática e Educação possuem uma dependência na origem da Educação Estatística que centraliza seus objetivos no desenvolvimento do pensamento probabilístico e estatístico e se preocupa com os assuntos relacionados com o ensino e aprendizagem de conceitos nessa área. A seguir, promovemos reflexões sobre a Educação Estatística voltada para o Ensino Superior, buscando obter os subsídios necessários que orientaram nossas discussões sobre o ensino e aprendizagem da estimação estatística.

2.2. Educação Estatística

A linguagem estatística e seus conceitos são utilizados para auxiliar na condução de inúmeras informações de diversas naturezas que a sociedade consome. Anúncios publicitários, pesquisas de opinião, comunicações científicas, utilizam a Estatística para fundamentar seus

¹³ La statistique considérée comme une science construite par les êtres humains tout au long des siècles et même des millénaires, est aussi un objet d'apprentissage, en particulier dans le cadre d'un enseignement scolaire ou universitaire mais aussi dans celui de la formation continue.

propósitos e transmiti-los de forma organizada por meio de tabelas, gráficos ou redações que empregam atributos estatísticos. Por consequência, surge a preocupação de como os indivíduos interagem diante dessas informações e, por conseguinte, como a educação pode contribuir para uma formação capaz de atender essas demandas. Carzola (2002) reconhece que a Estatística, antes restrita a algumas áreas, nunca esteve tão presente na vida do cidadão comum. Com o desenvolvimento e a abrangência dos estudos em Estatística, levanta-se a preocupação com a Educação Estatística.

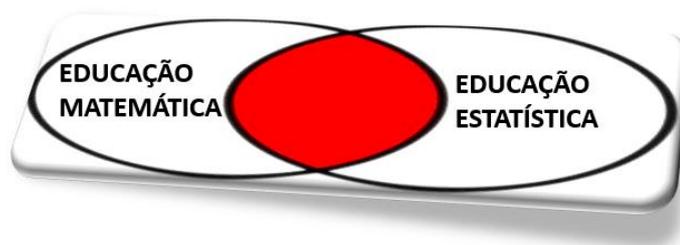
Em 1948, o interesse sobre o ensino de Estatística começou a ser discutido de forma mais extensiva com a constituição do Statistical Education Committee pelo ISI (International Statistical Institute), onde, em 1991, passou a ser chamado de International Association for Statistical Education (IASE)¹⁴, promovendo uma cooperação internacional que estimula a melhoria da educação estatística mundial em todos os níveis de instrução. Régnier (1996) aponta que um artigo publicado pela UNESCO nos anos 50, expôs assuntos educacionais sobre a formação do espírito estatístico e, em 1959, um relatório proferido por uma comissão dos Estados Unidos que tratava do ensino da matemática, declarou a necessidade da inserção de conhecimentos probabilísticos e estatísticos para a formação dos cidadãos. Na França, também no ano de 1959, em um seminário sobre educação matemática organizado pela Organisation Européenne de Coopération Economique, foi enfatizado que temáticas estatísticas vinham ganhando cada vez mais espaço em assuntos públicos, havendo a necessidade de inserir conteúdos estatísticos e probabilísticos no ensino da matemática. No Brasil, apesar do ensino de Estatística estar presente desde o final do século XVIII, ligado ao cálculo de probabilidades e destinado à formação de engenheiros militares, a discussão passou a ser mais explorada após a implementação das diretrizes nos Parâmetros Curriculares Nacionais (PCN) do ano de 1997, que oficializaram o ensino de probabilidade e estatística na educação básica. Carzola (2002) aponta que a educação estatística no Brasil tem seu marco histórico em uma Conferência Internacional realizada na Universidade Federal de Santa Catarina no ano de 1999 que tratou sobre as experiências e expectativas do ensino de Estatística. No ano de 2010, foi criado um Grupo de Trabalho (GT - 12), composto por pesquisadores que atuam na área de educação estatística, da Sociedade Brasileira de Educação Matemática (SBEM), objetivando explorar o ensino e aprendizagem da Estatística, observando aspectos cognitivos, afetivos, epistemológicos e o desenvolvimento de materiais que dão suporte ao ensino.

¹⁴ Disponível em: <https://iase-web.org/>.

Os princípios sobre o ensino da Estatística, como podemos perceber, foram vinculados com a educação matemática, provocando discussões em relação as principais distinções dessas duas ciências e o lugar que cada uma delas ocupa na educação. No contexto escolar, a Estatística está inserida como uma parte da matemática, levando a um possível entendimento de que elas possuem um desenvolvimento didático semelhante. Campos (2007) observa a existência de algumas peculiaridades comuns no âmbito educacional entre a Estatística e a Matemática, mas atenta ao fato de que é necessário esclarecer aspectos discordantes relacionados ao estudo da didática dessas duas ciências. Batanero (2001) discorre que a Educação Estatística não pode ser considerada como uma área da Educação Matemática, justificando que a Estatística não é uma área da Matemática pelo fato de existir provocações estatísticas de ordem filosófica, social e ética sobre procedimentos que não são fechados como os da matemática. A autora ainda reflete que é necessário experimentar e avaliar métodos de ensino que estejam adaptados para os aspectos específicos da Estatística, pois os princípios gerais do ensino da Matemática nem sempre poderão ser empregados. Carzola (2002) salienta que a educação estatística não deve priorizar a transmissão de técnicas para trabalhar com fórmulas e cálculos, mas a promoção da flexibilidade do pensamento durante a resolução de problemas e a habilidade para analisar dados.

A matemática possui uma natureza mais fechada em relação aos resultados de seus procedimentos, voltada para o determinismo, para a lógica e para a exatidão, enquanto a Estatística lida com problemas mais abertos, cujo resultados demandam análises e interpretações por meio de métodos que tentam lidar com a incerteza. Entendemos que a educação matemática e a educação estatística possuem, entre elas, muitos compartilhamentos metodológicos, mas ocupam propósitos teóricos diferentes. Apoiamos Coutinho (2013) ao revelar um diagrama que apresenta uma configuração entre a educação matemática e a educação estatística como mostra a Figura 9 a seguir:

Figura 9 - Relação entre a Educação Matemática e a Educação Estatística



Fonte: Adaptado de Coutinho (2013).

Existe uma independência entre as duas áreas, ocupando-se, cada uma, de suas questões particulares, mas havendo uma intersecção que mostra um certo grau de importância dos conhecimentos matemáticos para a aquisição de conhecimentos estatísticos. Em determinadas situações, os elementos matemáticos acompanham as técnicas estatísticas, mas desempenham um papel secundário diante das inúmeras ações de um processo que irá valorizar os aspectos específicos da Estatística como a variabilidade, a aleatoriedade e a incerteza que não combinam com as concepções determinísticas e lógicas da matemática.

Wodewotzki e Jacobini (2004) descrevem a Educação Estatística como um campo de investigação que tem como objetivo o ensino e a aprendizagem de conceitos estatísticos, assim como, o desenvolvimento das interpretações sobre os cálculos realizados. Lopes (2003) entende que a Educação Estatística deve auxiliar na leitura e na interpretação de dados, fornecendo habilidades para que o indivíduo possa analisar criticamente os dados apresentados. Dessa forma, podemos ver que a Educação Estatística é um campo da educação que visa investigar e auxiliar no ensino e aprendizagem da linguagem estatística e de seus conceitos, podendo utilizar a Educação Matemática como coadjuvante em um processo que prioriza a capacitação de habilidades subjetivas para que um indivíduo possa ser capaz de analisar e interpretar resultados obtidos diante das mais diversas situações em que a Estatística se faz presente.

2.3. Educação Estatística no Ensino Superior: a Estatística como uma disciplina de serviço

Continuamos nossas reflexões sobre a Educação Estatística observando agora alguns pontos significativos no processo de ensino e aprendizagem de uma disciplina de Estatística no Ensino Superior. Antes de iniciarmos, vamos localizar a formação estatística que foi observada em nossas análises, considerando dois tipos diferentes de público. O primeiro público, do qual não tratamos em nossa pesquisa, diz respeito ao que procura uma graduação em Estatística, onde o indivíduo receberá instruções que o formará como estatístico, possuindo conhecimentos específicos para todos os aspectos abrangentes da Estatística. O segundo, nosso público visado, é constituído de indivíduos oriundos de diferentes cursos de graduação que possuem uma ou algumas disciplinas obrigatórias de Estatística em sua grade curricular. Wada (1996) nomeia essas disciplinas de cursos de serviço ou disciplinas de serviço, que objetivam a instrumentalização de futuros profissionais para o uso adequado das ferramentas estatísticas. Essa formação procura alcançar o que Carzola (2002) chama de usuários de Estatística, que são os consumidores das informações estatísticas, mas entendemos que também poderão produzir, por conta da instrução recebida. A Educação Estatística possui um olhar mais atento para esse

público, tendo em vista que o processo de ensino nem sempre é satisfatório como é apontado por Barceló, Capilla e Zúnica (2000), Carzola (2002), Batanero (2001), Grácio e Oliveira (2005), Mantovani e Viana (2004), entre outros.

O baixo aproveitamento dessas disciplinas de serviço levanta discussões para a Educação Estatística, apontando a Matemática como um dos principais fatores que contribuem para as atitudes negativas dos estudantes em relação a Estatística. No Ensino Básico, a Estatística está inserida no contexto do ensino da Matemática e, portanto, transmitida por professores licenciados em Matemática que, segundo Carzola (2002), nem sempre estão capacitados para o ensino da Estatística, afetando a aprendizagem em todos os níveis de ensino. Barbeta (2006) expõe que muitos estudantes chegam ao Ensino Superior com pouco ou nenhum conhecimento estatístico. Essa percepção da Estatística inserida na Matemática no Ensino Básico, leva a compreensão dos estudantes que chegam ao Ensino Superior de que a Estatística é uma continuidade da Matemática, o que a torna rejeitada por muitos que não obtiveram um bom desempenho em Matemática durante o Ensino Básico. A desconstrução dessa percepção não é tarefa fácil, levando ao debate de como a Matemática deve ser empregada no Ensino Superior diante de uma disciplina de Estatística. De modo geral, a Matemática exige técnicas de resoluções que acarretam uma única solução para um determinado problema, distanciando o estudante da natureza Estatística que exige uma postura mais crítica diante do enfrentamento da incerteza e de problemas com possibilidades variadas para uma tomada de decisão. Silva (2007) salienta que o reforço dos fundamentos matemáticos é importante quando o ensino é voltado para capacitar estatísticos, mas discutível quando os estudantes forem usuários da Estatística. Hand (1998) reconhece a importância da Matemática para o ensino da Estatística, revelando que quanto maior o conhecimento matemático, menor a propensão de erros por parte dos estudantes na resolução dos problemas. Contudo, o autor defende que o uso de softwares estatísticos para o desenvolvimento dos cálculos, prioriza o desenvolvimento de habilidades em análise estatística, diminuindo os esforços em relação ao fundamento matemático da análise. A Educação Estatística deve priorizar o desenvolvimento de uma formação estatística que fomente o espírito crítico capaz de reconhecer o uso das fórmulas matemáticas inseridas nos conceitos estatísticos afim de promover reflexões diante dos resultados encontrados. Carzola (2002) reconhece a Matemática como umas das dificuldades encontradas na formação do usuário da Estatística, revelando outros possíveis pontos como mostra a Figura 10 a seguir:

Figura 10 - Dificuldades na formação de usuários da Estatística



Fonte: Adaptado de Carzola (2002).

A fundamentação teórica da Estatística está na Matemática, dessa forma, pelo fato de ser um campo considerado problemático por muitos estudantes que ingressam no Ensino Superior apresentando dificuldades em operações básicas da Matemática, o que os levam a transferir esse sentimento para a Estatística. A natureza da Estatística, como já mencionamos, requer o trabalho com experimentos aleatórios e conceitos de natureza probabilística, submetendo o estudante a realizar procedimentos para tomada de decisões em meio à incerteza. Outro ponto relevante citado pelo autor é a carga horária das disciplinas de Estatística que, por vezes, possui um número reduzido de horas. Apesar da sua inquestionável importância na formação de qualquer profissional, a Estatística, por ser uma disciplina de serviço, insere-se como coadjuvante nas diversas áreas do conhecimento, possuindo uma pequena carga horária para o seu ensino na grade curricular dos cursos de graduação. Sendo assim, os professores acabam sanando gargalos advindos da Educação Básica, como os procedimentos matemáticos, deixando de lado o desenvolvimento do pensamento estatístico.

Diante das reflexões levantadas, entendemos que é necessário refletir sobre a principal função das disciplinas de Estatística no Ensino Superior, observando na Educação Estatística competências necessárias para uma formação Estatística que atenda as demandas da sociedade. Rumsey (2002), Gal e Garfield (1997), Campos (2007), delMas (2002) e outros, defendem em suas pesquisas o desenvolvimento de três importantes competências: o letramento, o pensamento e o raciocínio estatístico. Dessa forma, iniciamos o próximo item da nossa pesquisa para reconhecer e compreender essas competências.

2.4. Competências para a Educação Estatística

O ensino da Estatística deve oportunizar e contribuir para que os estudantes tenham um melhor desenvolvimento cognitivo de competências fundamentais para o consumo de diversas informações que são propagadas em nosso cotidiano. A Estatística possui ferramentas que colaboram na formação de um cidadão crítico, capaz de realizar interpretações e análises diante de inúmeros resultados que permeiam nossas vidas.

Muitos estudiosos, tais como Gal e Garfield (1997), Gal (2002), Rumsey (2002), delMas (2001), Selz & Maillochon (2009), Régnier (2009) e outros, revelam em suas pesquisas, indicativos que sustentam a necessidade de desenvolver nos estudantes o letramento, o pensamento e o raciocínio estatístico, vistos como competências fundamentais para o aprendizado de conceitos elementares da Estatística e na compreensão de suas diferentes aplicações. Diante disso, os pesquisadores indicam desenvolvê-las e observá-las de modo unificado, pois entendem que, para haver uma compreensão global do que é e para que serve a Estatística, o incremento harmônico dessas três competências faz-se necessário.

Apresentamos, a seguir, considerações sobre cada uma dessas competências, onde procuramos explorá-las de acordo com percepções voltadas para a Estimação Estatística.

2.4.1. Letramento estatístico

Começamos nossas considerações justificando o uso do termo letramento. Derivada da palavra inglesa literacy, a expressão surge como forma de diferenciação do significado da palavra alfabetização. A alfabetização é definida como o processo de aprendizagem onde se desenvolve a habilidade de ler e escrever, enquanto o letramento significa o desenvolvimento e uso competente da leitura e da escrita nas práticas sociais. Soares (2004) menciona que em meados de 1980 se dá, simultaneamente, a invenção do letramento no Brasil, do illetrisme na França e da literacia em Portugal, para nomear fenômenos distintos da alfabetização. A autora prossegue informando que, apesar da palavra literacy já estar dicionarizada desde o século XIX nos Estados Unidos e na Inglaterra, foi em 1980 que ela passou a fazer parte de discussões também nas áreas da Educação. No Brasil, muitos pesquisadores da área da Educação Estatística, utilizam o termo literacia e, ainda que Soares (2004) explique que as palavras letramento e literacia sejam sinônimos, empregamos o termo letramento por entendermos ser a tradução mais fidedigna para a nossa língua materna.

Sabemos que o letramento, por si só, representa um fenômeno que define a capacidade de compreensão dos indivíduos em relação a importância de possuir o domínio da leitura e da escrita para a tomada de decisões diante de situações cotidianas, tanto no âmbito pessoal como no profissional. Kataoka *et al.* (2011) argumentam que uma pessoa letrada transforma seu estado cognitivo, social, cultural e econômico, modificando o modo de avaliar informações que chegam ao seu conhecimento.

Uma das primeiras definições colocadas para o termo letramento estatístico vem de Wallmann (1993) que o descreveu como sendo uma competência que ajuda compreender e avaliar criticamente resultados estatísticos que diariamente surgem na vida de todos nós. Watson (1997) define o letramento estatístico como sendo a capacidade de entendimento do texto e do significado das implicações contidas nas informações estatísticas inseridas em seu contexto formal, identificando três estágios para que ocorra o seu desenvolvimento, são eles: o entendimento básico da terminologia estatística; o entendimento da linguagem estatística e os conceitos inseridos num contexto de discussão social e o desenvolvimento de atitudes de questionamento nas quais se aplicam conceitos mais sofisticados para contradizer alegações que são feitas sem fundamentação estatística apropriada.

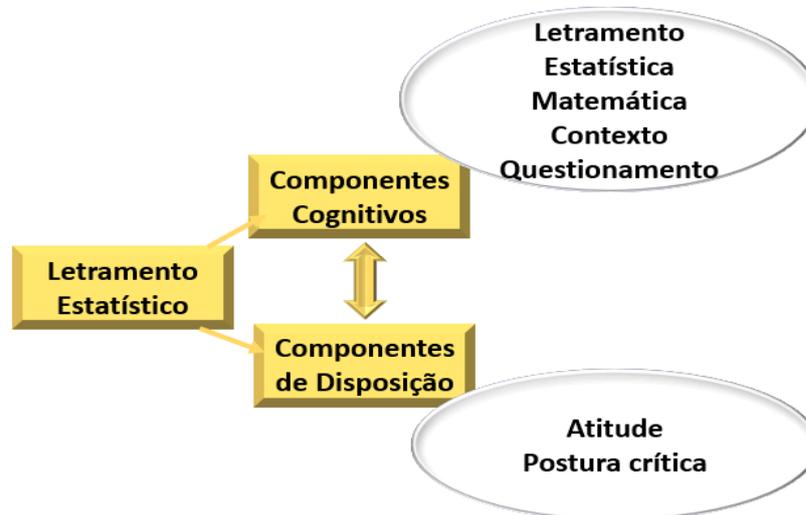
Assim sendo, podemos compreender que o termo letramento estatístico evoca o fenômeno definido pelo estudo de argumentos que utilizam a Estatística como referência, empregando habilidades na organização e representação de dados e nas interpretações necessárias para tomadas de decisões diante das mais variadas situações.

Gal (2002) pensa no letramento estatístico refletindo sobre o interesse de indivíduos consumidores de informações estatísticas que aparecem nos mais diferentes meios de comunicação. A proposta do autor, com a qual concordamos, sugere que o letramento estatístico seja composto por um componente cognitivo e por um componente de disposição, que deverão ser correlacionados para que os indivíduos tenham um melhor desempenho na extração das informações que apresentam argumentos estatísticos (Figura 11).

Os componentes cognitivos são os encarregados pelas competências que interpretam e avaliam criticamente a informação estatística, sendo compostos por cinco elementos: o letramento, o conhecimento estatístico, o conhecimento matemático, o conhecimento de contexto e a habilidade de elaborar questões críticas. Por sua vez, os componentes de disposição designam as competências para discutir e exprimir os resultados gerados pela informação estatística, sendo compostos por dois elementos: as atitudes que moldam a visão de mundo do indivíduo e a postura crítica, ou seja, a capacidade de gerar questionamentos diante das

informações quantitativas que podem ser, de acordo com Silva (2007), unilaterais, viesadas ou incompletas, seja de maneira intencional, ou não.

Figura 11 - Composição do letramento estatístico



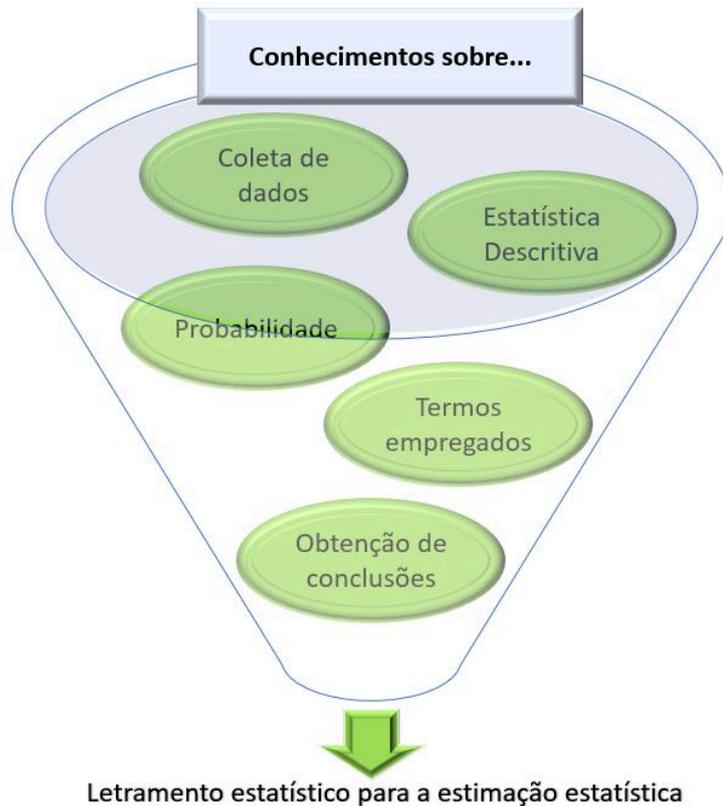
Fonte: Adaptado de Gal (2002).

Observa-se que o letramento estatístico, especificamente em relação a estimação estatística, pode ser explorado quando Gal (2002) disserta sobre os conhecimentos de conceitos e procedimentos básicos de Estatística e Probabilidade, que formam os componentes cognitivos. Ao propor cinco tópicos para abranger os quesitos necessários para a compreensão e interpretação de informações estatística, vislumbramos as relações existentes com a estimação estatística, como mostra a Figura 12.

O primeiro tópico refere-se ao conhecimento dos motivos e das maneiras pelas quais a coleta de dados aconteceu. A estimação estatística necessita de uma coleta de dados para a constituição de uma amostra e, assim, realizar inferências. Dessa forma, conhecer e entender o processo da coleta de dados é um ponto fundamental para que a estimação estatística aconteça. O segundo tópico aborda a necessidade da familiaridade com as ideias básicas relacionadas com a Estatística Descritiva. Para realizar inferências a estimação estatística necessita de informações como a média, variância e desvio padrão que são ferramentas advindas da Estatística Descritiva. No terceiro tópico é exposta a necessidade da familiaridade com os termos empregados. Reconhecer e compreender os conceitos de parâmetros e estimadores, por exemplo, é essencial para a aplicação da estimação. O quarto tópico ocupa-se da compressão das noções básicas de probabilidade. A estimação estatística faz uso da probabilidade por meio dos modelos de distribuição de probabilidades para conjecturar informações sobre o comportamento de uma variável em estudo. O quinto e último tópico revela sobre o

conhecimento de como as conclusões e inferências estatísticas são obtidas. Diante de uma situação proposta, o indivíduo deve reconhecer a necessidade do uso de estimativas pontuais ou intervalares para conseguir inferir sobre o objeto em estudo.

Figura 12 - Tópicos do letramento estatístico para a estimação estatística



Fonte: Do autor (2018).

De modo geral, um indivíduo letrado estatisticamente e, em especial, apto para efetuar e interpretar estimações estatísticas, deve possuir: conhecimentos da coleta de dados que produziu uma amostra; instrução quanto as ferramentas da Estatística Descritiva que servem de base para a elaboração dos cálculos da estimação; competências probabilísticas para compreender as conjecturas de um contexto analisado; identificação dos agentes empregados e suas funções (estimadores, parâmetros,...) e, por fim, reconhecer as ações mais acuradas para a apuração e interpretação das conclusões.

Identificados alguns dos possíveis tópicos para o letramento estatístico referente a estimação estatística, entendemos que o indivíduo deve receber uma escolarização que supra essas necessidades e dialogue, de forma natural, utilizando informações que se recebe a todo momento. Watson *et al.* (2003) argumenta que o ensino da Estatística deve oportunizar aos

estudantes situações que exijam o uso de conceitos estatísticos e probabilísticos para a formação de cidadãos críticos e letrados estatisticamente.

Concluimos evocando Rumsey (2002) que interpreta o letramento estatístico introduzindo as ideias de competência estatística e cidadania estatística. De acordo com a autora, além dos conhecimentos básicos das ferramentas estatísticas, a cidadania estatística exige um olhar crítico para as informações, demandando o desenvolvimento do pensamento e do raciocínio estatístico. Dessa forma, para que tenhamos indivíduos capazes de enfrentar situações da vida moderna que exigem o letramento estatístico, é preciso fomentar e investir no ensino para o aperfeiçoamento do raciocínio e do pensamento estatístico.

2.4.2. O pensamento estatístico

Iniciamos nossas reflexões apoiadas nas considerações encontradas em algumas pesquisas. O trabalho de Carzola (2002) considera o pensamento estatístico tão necessário quanto a capacidade de ler e escrever, para que tenhamos uma cidadania plena. A tese de doutorado de Lopes (2003) mostra que pensar estatisticamente consiste na capacidade de uma pessoa compreender mensagens simples e diretas presentes no cotidiano, bem como as que envolvem processos complexos de inferência. A autora complementa, ainda, que o domínio dessa forma de pensar oferece ao indivíduo maiores possibilidades de exercer sua cidadania. Mallows (1998) considera o pensamento estatístico como a capacidade de relacionar dados quantitativos com situações concretas, admitindo a existência da variabilidade e da incerteza, explicitando o que os dados podem revelar em relação a um determinado problema. Diante disso, refletimos que pensar estatisticamente possibilita ao indivíduo assimilar informações quantitativas expressas em textos, tabelas ou gráficos, gerando uma visão crítica que o ajudará em possíveis tomadas de decisões para a sua vida pessoal e profissional. O pensamento estatístico é uma ação do pensamento que distingue e percebe informações contextualizadas em diversas realidades, mas com elementos comuns que são próprios da Estatística.

Assim sendo, percebemos que o pensamento estatístico demanda para o ensino da Estatística não apenas a preocupação com as resoluções práticas dos procedimentos que envolvem algoritmos, mas deve estar atento para o processo completo, para a reflexão e para a interpretação dos resultados. Ferreira *et al.* (2010) pontuam que uma característica particular do pensamento estatístico é favorecer a habilidade de perceber o processo de maneira global, compreendendo os significados de suas diferentes relações para especular questões.

Algumas inquietações provenientes dessas considerações sobre o pensamento estatístico são expostas por Chance (2002) que faz a seguinte pergunta: Como podemos ensinar o pensamento estatístico e como podemos determinar se os estudantes estão pensando estatisticamente? Podemos refletir, diante do questionamento, que o pensamento é uma disposição natural do indivíduo que irá se expandir, com o passar dos anos, adquirindo novas habilidades cognitivas por meio, por exemplo, da educação. O ensino da estatística deve contribuir para promover a evolução do pensamento, capacitando o indivíduo a resolver e interpretar problemas quantitativos cada vez mais complexos. Campos (2007) defende que as estratégias de pensamento usadas pelos professores e suas assimilações para a resolução de problemas, contribuem para o desenvolvimento do pensamento estatístico nos estudantes. Concordamos com Campos (2007) quanto a sua reflexão que expõe a difícil tarefa de desenvolver o pensamento estatístico nos estudantes. O autor aponta que não é possível ensinar diretamente a pensar estaticamente, mas argumenta sobre a viabilidade do uso de estratégias que estimulam a valorização de hábitos mentais que promovam essa ação.

Chance (2002) sugere algumas práticas mentais que podem auxiliar e fomentar o desenvolvimento do pensamento estatístico. Relacionamos suas ideias, considerando o pensar na resolução e interpretação de um problema de estimação estatística, como mostra a Figura 13.

Julgamos que essas ideias podem conduzir para um melhor exercício da prática do pensamento estatístico avistando situações que abarquem a estimação estatística. Com todas as ideias interligadas, identificamos que a visão geral do processo se refere ao planejamento do ciclo investigativo, onde o indivíduo deverá levantar questões dos primeiros passos decisórios, inteirando-se do assunto e levantando possíveis hipóteses para a resolução. A obtenção de dados surge com a proposta de considerar, sempre, a melhor forma de conseguir dados que sejam relevantes para a constituição de uma amostra que representará uma determinada situação. A estimação estatística ocupa uma ou mais amostras para conseguir inferências, tendo a escolha dos dados, papel fundamental para atingir resultados satisfatórios. A análise constante das variáveis envolvidas, como a escolha de um bom estimador, também é indispensável para a estimação estatística, tendo em vista o julgamento do parâmetro que estará sendo analisado.

Figura 13 - Hábitos mentais de Chance (2002) adaptados para a Estimação Estatística



Fonte: Adaptado de Chance (2002).

Chance (2002) expõe o ceticismo onipresente em relação aos dados, nos remetendo a compreensão de que o indivíduo nunca perca o senso crítico, observando as escolhas e adaptando-as como julgar necessário. A incredulidade nos dados para a estimação, pode gerar reflexões que estimulem o pensar estatístico, pois o indivíduo estará realizando inferências com base em pequenos grupos que serão remetidas a grupos maiores. A relação dos dados com o contexto de uma determinada situação infere nas interpretações das conclusões que poderão ser expostas em termos não estatísticos, fazendo com que o indivíduo divulgue os resultados de forma criativa para ser compreendida facilmente pelos consumidores da informação.

Todas essas ações descritas poderão ser percebidas e empregadas pelos educadores em avaliações ou em intervenções com práticas estatísticas, para que se observe indícios do desenvolvimento do pensamento estatístico. Wodewotzki *et al.* (2010) argumenta que o pensamento estatístico pode ser bem desenvolvido por meio de relações que os estudantes fazem com situações concretas e aplicadas, bem como possam perceber que o resultado de uma estimação, por exemplo, evoca uma tendência e não uma certeza, interpretando os resultados e explorando os dados sob diversos ângulos.

Concluimos nossas considerações sobre o pensamento estatístico, apreciando Wild e Pfannkuch (1999) que comparam o pensamento estatístico com o ato de respirar. Os autores

entendem que diversas informações são interpretadas utilizando o pensamento estatístico sem ser percebido e complementam dizendo que uma disciplina de Estatística deve ser ensinada para que as pessoas respirem mais efetivamente. Dessa forma, o ensino da Estatística é peça chave para o aperfeiçoamento do pensamento estatístico, tornando o conteúdo mais significativo se for implementado de modo paralelo ao contexto que circunda a realidade do indivíduo.

2.4.3. O raciocínio estatístico

Nosso primeiro apontamento introduz a definição de raciocínio como um exercício da razão, a dedução de algo frente a uma ou várias premissas que irão estimular uma tomada de decisão. É um processo interno, um conjunto de atividades mentais, cuja percepção das conexões de ideias atreladas a certas regras, dão suporte para o desenvolvimento de soluções na presença de uma determinada informação. Em outras palavras, o raciocínio é uma competência humana que possibilita solucionar problemas.

Sobre a especificidade do raciocínio estatístico, por se tratar da compreensão de uma ciência, o indivíduo deve possuir ferramentas básicas da Estatística advindas de sua escolarização para poder absorver e inferenciar frente a situações com essa demanda. Para Garfield (2002) a definição do raciocínio estatístico está vinculada com a maneira pela qual um indivíduo raciocina com ideias estatísticas e a forma de como dá sentido às informações estatísticas. Isso compreende fazer interpretações sobre dados, construções de tabelas, gráficos e, inúmeras vezes, o raciocínio estatístico envolve ideias de estimações e testes de hipóteses para que haja interpretações e inferências acerca dos resultados. Diante disso, podemos considerar que o raciocínio estatístico compreende um conjunto de processos mentais que articulam ferramentas estatísticas para efetuar cálculos e interpretar os resultados.

Ao prosseguirmos com os nossos apontamentos, considerando-se que a definição de raciocínio estatístico está relacionada, também, a execução de cálculos, compreendemos que é necessário distinguir, no que concerne ao raciocínio, a Estatística, da Matemática. Apesar de admitirmos a presença de competências comuns, o raciocínio matemático visa alcançar propriedades operacionais e deduções lógicas, enquanto o raciocínio estatístico ocupa-se de interpretações subjetivas ligadas, muitas vezes, por algum conceito matemático que é extrapolado por análises complexas e relações abstratas que são descritas, em concordância com Campos (2007), por meio de uma linguagem própria (letramento estatístico) diante da compreensão global de um fenômeno (pensamento estatístico).

Para melhor situar as disposições do raciocínio estatístico, Gal e Garfield (1997) manifestam as seguintes colocações:

- a) Os dados, para a Estatística, representam números inseridos num contexto que irão nortear as interpretações dos resultados;
- b) Conceitos e procedimentos matemáticos são utilizados como uma das partes da solução de problemas estatísticos. É necessário, entretanto, procurar resultados mais expressivos, induzindo o crescimento pela utilização de softwares que executam a parte operacional;
- c) A natureza fundamental de muitos problemas estatísticos é a de que eles, frequentemente, não possuem uma única solução matemática. Os problemas estatísticos comumente apontam um questionamento inicial e findam com uma opinião que deve ser fundamentada em certos resultados teórico-práticos.

Diante das colocações realizadas, podemos dissertar sobre o primeiro item que, por vezes, a Matemática manuseia números para a obtenção de resultados puramente numéricos, não objetivando algum tipo de contextualização. Entretanto, para a Estatística, os dados coletados para a efetivação de alguma análise, sempre devem ser considerados números em consonância com o tema proposto, ou seja, existe uma representatividade por parte dos números em relação ao contexto analisado. O segundo item considera o uso de procedimentos matemáticos como uma das parcelas responsáveis pela obtenção de resultados para a Estatística, observando que o processo prático pode ser melhor atingido fazendo uso de recursos computacionais. Não discordamos dessa consideração, porém entendemos que as técnicas matemáticas manuais sempre devem estar presentes pois, em concordância com Batanero (2001), se os alunos adquirirem só o conhecimento do tipo computacional, é provável que cometam erros em suas operações. O último item discorre que geralmente uma solução estatística encontrada para um problema não deve ser compreendida como única. Nesse sentido, os estudantes, ao resolverem um problema de Estatística, expressam suas conjecturas que não devem ser encaradas como certas ou erradas, pois necessitam ser analisadas quanto à qualidade de seu raciocínio, efetuando adequações de métodos, quando necessário, para fundamentar as comprovações descobertas.

Garfield e Gal (1999) elencam alguns tipos específicos de raciocínio que são esperados dos estudantes do Ensino Superior para um melhor aprendizado da Estatística. Combinamos a relação proposta pelos autores com as possíveis especificidades dos raciocínios que devem ser desenvolvidos pelos estudantes durante o aprendizado da estimação estatística, como mostra a Figura 14 a seguir.

Figura 14 - Raciocínios para a Estatística indicados por Garfield e Gal (1999), adaptados para a Estimação Estatística



Fonte: Adaptado de Garfield e Gal (1999).

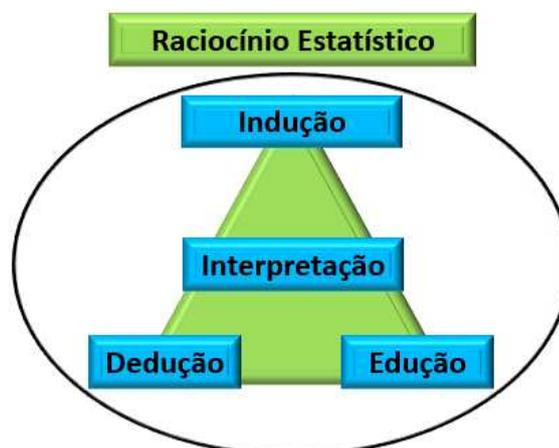
A estimação estatística, enquadrada no ramo da estatística inferencial, utiliza a coleta de dados para constituir uma ou mais amostras de um determinado grupo comum, para a submissão de análises e interpretações de possíveis tendências que serão expandidas para o grupo comum não coletado. Dessa forma, os raciocínios sobre os dados, e sobre a representação dos dados, são fundamentais para o reconhecimento de suas categorias (qualitativos, quantitativos discretos ou contínuos) e a exposição que os retratam mais adequadamente (tabular ou gráfica). Além disso, o raciocínio sobre amostras, incide no entendimento da relação que uma amostra possui com a sua população. Amostras grandes e bem selecionadas representarão melhor a população, e precauções deverão ser tomadas no exame da população quando as amostras forem pequenas. No que discorre o raciocínio sobre medidas estatísticas, apuramos que os estimadores e os parâmetros utilizados na estimação estatística podem ser caracterizados por medidas de tendência central e medidas de dispersão, como a média e a variância, por exemplo. O entendimento do uso das medidas mais apropriadas, em cada caso, é indispensável para apurar estimações, comparando diferentes distribuições e ajustando possíveis previsões. O raciocínio sobre incerteza, reporta-nos ao discernimento sobre probabilidades que propicia, por exemplo,

a compreensão dos intervalos de confiança diante de uma estimação intervalar. Além disso, esse raciocínio abrange desde a compreensão do ceticismo sobre os dados até as interpretações dos resultados das possíveis tendências averiguadas diante de uma informação estimada. O raciocínio sobre associações requer saber julgar e interpretar as ligações entre todas as variáveis envolvidas no processo da estimação, ajustando um planejamento e estabelecendo as relações entre os dados e a amostra, entre a amostra e a população e entre os estimadores e os parâmetros. É o raciocínio que facilita a visão geral do processo para combinar as ferramentas estatísticas necessárias para a obtenção de resultados.

Entendemos que todos esses raciocínios são necessários para a compreensão da estimação estatística e para todas as outras especificidades da Estatística. Contudo, especialmente no que se refere ao ramo da estatística inferencial, é necessária, ainda, a abstração dos resultados numéricos obtidos pelos processos estatísticos empregados para a composição de uma interpretação que esteja adequada ao contexto analisado. Essa interpretação estatística consiste em ler e comentar, com competência, os resultados desses cálculos. Dessa forma, observamos que as características do raciocínio para a estimação estatística, permeiam a subjetividade e, como no pensamento estatístico, não é tarefa fácil de ser desenvolvida entre os estudantes. Sedlmeier (1999) aponta que a instrumentalização estatística, muitas vezes, não é o suficiente para perceber progressos no desenvolvimento do raciocínio, principalmente quando se trata da estatística inferencial.

Diante disso, evocamos Régnier (2009) que defende a integração da interpretação estatística com o raciocínio estatístico, estabelecendo três modos de raciocínio para compor essa relação: a dedução, a indução e a educação. A seguir, a Figura 15 mostra o esquema proposto pelo autor.

Figura 15 - Raciocínio estatístico integrador



Fonte: Adaptado e traduzido de Régnier (2009).

Da indução, espera-se a competência do entendimento dos dados para descrevê-los apropriadamente no contexto do problema analisado. A dedução requer o entendimento de um certo resultado para a compreensão dos elementos que foram envolvidos no problema analisado. A educação¹⁵ surge para compor essa relação no sentido de fazer emergir informações diante de um conjunto de dados sem estrutura aparentemente saliente, ou seja, é a capacidade do indivíduo em detectar relações entre duas ou mais ideias percebidas em um problema ou evocadas da memória. Ainda sobre educação, Spearman (1927) explica que é a capacidade de formular novas ideias ou relações a partir de ideias ou relações já estabelecidas.

O estímulo desses raciocínios pode gerar uma melhora no desenvolvimento do raciocínio estatístico, tendo em vista que raciocinar e interpretar estatisticamente, seja informações numéricas implícitas ou explícitas, requer um entendimento biunívoco, ou seja, diante de um resultado o indivíduo pode ser capaz de compreender os elementos que foram analisados para a chegada de uma decisão. Da mesma forma, a partir da análise dos dados, conseguir inferenciar. Ao nos referirmos sobre informações implícitas, buscamos o sentido da educação, onde uma informação, sem uma estrutura aparentemente estatística, pode evidenciar interpretações estatísticas. Régnier (2009) argumenta que o desenvolvimento do raciocínio estatístico faz com que o indivíduo tenha novas percepções diante de informações regulares ou fortuitas e novas ideias ao ler o mundo com um novo espírito, denominado espírito estatístico.

2.4.4. Compreendendo o espírito estatístico

A formação do espírito estatístico requer conhecimentos científicos advindos da educação estatística, que podem promover o estímulo de um raciocínio crítico para decifrar fenômenos, apropriando-se de uma cultura de manipulação e processamento de dados. Régnier (2003) argumenta que um cidadão responsável e crítico, deve estar atento a conhecimentos básicos da estatística para entender os diversos fenômenos que o circundam. No decorrer dos anos de nossa história, muitos pesquisadores observaram e analisaram os caminhos que levavam o ser humano a produzir e apropriar-se dos conhecimentos científicos, entre eles, destacamos Gaston Bachelard¹⁶ (1884-1962).

Bachelard (1996) destaca que uma das grandes dificuldades na formação de um espírito científico decorre em função de obstáculos epistemológicos, ou seja, crenças e conhecimentos

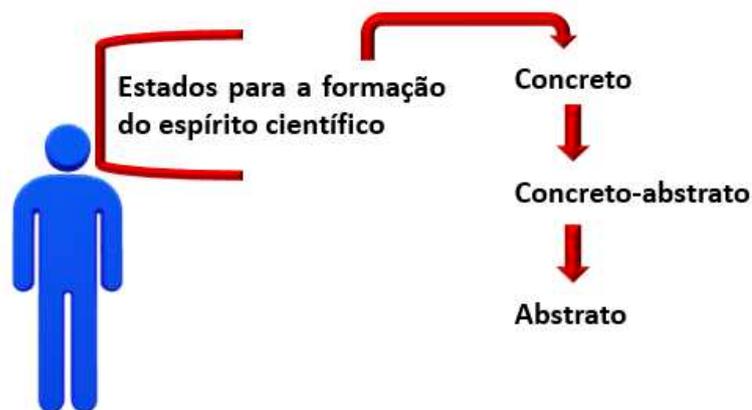
¹⁵ Essa palavra foi usada por Charles Edward Spearman (1863-1945), psicólogo inglês conhecido pelo seu trabalho na área da Estatística e modelos da inteligência humana.

¹⁶ Filósofo francês das ciências, da atitude poética e do tempo, autor de várias obras, entre elas, podemos citar a obra epistemológica: *A Formação do Espírito Científico* (1938).

já adquiridos pelos indivíduos, podendo gerar uma inércia, uma estagnação para novos conhecimentos. De acordo com o pesquisador, a ciência consegue progredir por meio da superação desses obstáculos e, nesse sentido, ela deve ir contra um conhecimento anterior, aniquilando saberes erroneamente obtidos ou extrapolando aquilo que barra o progresso. O termo ruptura, conforme Bachelard (1996), é usado para indicar uma descontinuidade entre o conhecimento comum e o conhecimento científico, descrevendo, assim, a necessidade de uma ruptura entre um conhecimento não questionado e, portanto, não científico, em favor de um conhecimento problematizado e, portanto, científico.

Na prática educacional, Bachelard (1989) fomenta a condução dos estudantes a criarem uma consciência científica que ultrapasse o mero aprendizado de técnicas tendo em vista as diversas situações a que serão submetidos pelo mundo que os cerca. O espírito científico, para uma formação individual, de acordo com Bachelard (1996), percorre três estados que estão apresentados na Figura 16 a seguir:

Figura 16 - Estados para o desenvolvimento do espírito científico segundo Bachelard



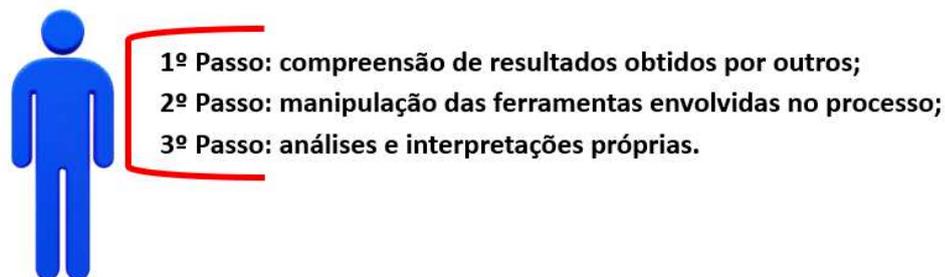
Fonte: Do autor (2018).

O estado concreto é fase na qual o indivíduo forma suas opiniões a partir das primeiras imagens que adquire do fenômeno. O concreto-abstrato é fase na qual o indivíduo, mesmo estando ligado em experiências anteriores, é capaz de descobrir algumas generalizações e de desvendar algumas descobertas. Por último, o estado abstrato, é a fase na qual o indivíduo é capaz de analisar as complexidades do problema e gerar novos conhecimentos por meio de seus questionamentos. Para Bachelard (1996) o espírito científico é adquirido por meio de relações que o indivíduo estabelece com seu contexto, bem como das experiências vivenciadas, enquanto ser que existe porque é ação, ou seja, em função da troca ocorrida entre sujeito e objeto.

Com base na abordagem iniciada por Bachelard (1996), Régnier (2003) propõe uma ideia para o treinamento da formação do espírito estatístico caracterizado pela natureza da relação de incerteza e erro, onde o espírito estatístico exige uma renúncia do uso sistemático de verdades para procurar controlar a verossimilhança e a plausibilidade. Dessa forma, o autor buscou o entendimento da ruptura de obstáculos epistemológicos para a inserção de um novo conhecimento, uma nova conquista do modo de raciocinar e de se posicionar diante de possíveis informações estatísticas. Para tanto, um ponto fundamental a ser compreendido, e aqui fazemos alusão ao processo de estimação estatística, seria a consciência da presença de amostragem flutuante, isto é, os resultados obtidos de uma estimação variam de acordo com as amostras. A investigação por meio do levantamento de uma série de amostras não apresentaria necessariamente resultados idênticos, mas estariam sujeitos a variação.

De maneira análoga aos estados que Bachelard (1996) propôs para o desenvolvimento do espírito científico, construímos um esquema interpretando o que Régnier (2005) indica como sendo as possíveis etapas para a formação do espírito estatístico, como mostra a Figura 17 a seguir:

Figura 17 - Etapas para a formação do espírito estatístico



Fonte: Adaptado de Régnier (2002).

O primeiro passo seria a observação e a compreensão de resultados já obtidos em situações estatísticas semelhantes. Analogamente, estaríamos no estado Concreto de Bachelard, onde o indivíduo forma suas primeiras ideias e adquire os primeiros conhecimentos científicos do fenômeno analisado. O segundo passo (Concreto-abstrato), o indivíduo manipula as ferramentas estatísticas e, mesmo possuindo algumas referências anteriores, pode gerar novas informações percebendo suas diferentes aplicações. O estágio final (Abstrato) almeja que o indivíduo consiga alcançar o raciocínio estatístico realizando interpretações próprias fazendo uso das ferramentas estatísticas. Entendemos que um conhecimento científico pode ser construído por meio de experiências que impugnam o conhecimento comum. O verdadeiro conhecimento só é adquirido quando um indivíduo experimenta densamente o contexto que o

cerca, admitindo dúvidas e permitindo explorá-las em função da busca de fatos que possam consolidar e fundamentar a ciência em questão. Assim, é necessário que durante uma formação estatística, o indivíduo possa vivenciar diversas situações para expandir sua visão de mundo. A seguir, a Figura 18 representa um sintético esquema onde Régnier (2005) indica os principais elementos que contribuem para o treinamento da formação do espírito estatístico.

Figura 18 - Elementos para a formação do espírito estatístico



Fonte: Adaptado e traduzido de Régnier (2003).

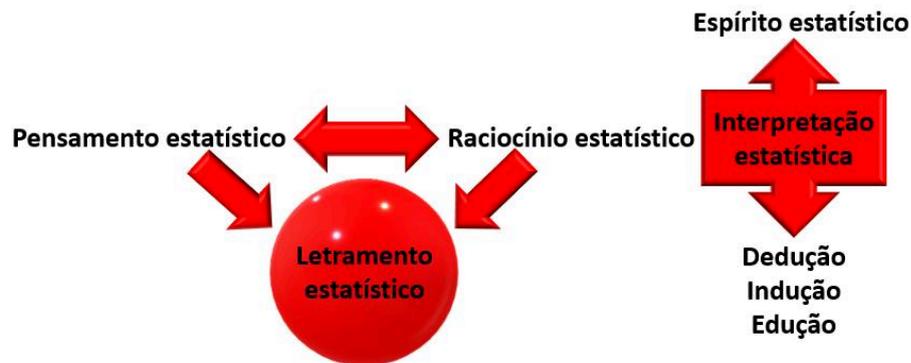
O esquema proposto revela o desenvolvimento do pensamento estatístico enquanto uma competência de perceber quando e como aplicar os conhecimentos e os procedimentos estatísticos, observando o contexto e a influência das informações que podem gerar condições de incerteza ou, como citado anteriormente, a presença de amostragem flutuante. O raciocínio estatístico necessita da integração da interpretação estatística que só será alcançada quando o indivíduo justificar suas conclusões realizando indução estatística, ou seja, competência para argumentar cientificamente sua tomada de decisão.

Uma formação estatística não pode estar restrita apenas na aprendizagem do uso das ferramentas, dos métodos e dos algoritmos processuais para obtenção de soluções, mas deve, também, mobilizar ações que facilitem uma interpretação mais crítica e reflexiva para os resultados obtidos nas mais variadas situações. Diante disso, entendemos que o espírito estatístico promove uma condição essencial para o raciocínio científico das concepções estatísticas (descritivas e indutivas) visando gerar argumentos questionadores e reflexivos para a tomada de decisões em situações cotidianas e profissionais.

2.4.5. Reflexões finais sobre o letramento, o pensamento e o raciocínio estatístico

Para consolidar uma boa formação estatística todas as competências citadas devem atuar em conjunto para atender as demandas operacionais, interpretativas e de tomadas de decisão. delMas (2002) percebe o letramento estatístico como o objetivo que engloba a meta geral do ensino da Estatística, considerando os aspectos do pensamento e do raciocínio estatístico. Campos (2007) declara entender que não é produtivo pensar em uma formação estatística considerando as competências estatísticas de forma independente, pois elas se complementam e somente juntas irão abranger a compreensão global da Estatística. Rumsey (2002) argumenta que para o desenvolvimento do letramento estatístico faz-se necessário pensar sempre em outras capacidades correlatas que irão se somar para promover uma melhor compreensão dos conceitos estatísticos. Régnier (2003) indicou um raciocínio estatístico integrador, com observação nas interpretações estatísticas, além da compreensão do que o autor chama de espírito estatístico. Diante disso, para tentarmos complementar com as análises realizadas, elaboramos um esquema para visualizar e definir o papel de cada competência como mostra a Figura 19 a seguir:

Figura 19 - Competências estatísticas



Fonte: Do autor (2018).

Concordamos com Chance (2002) que percebe o letramento estatístico como a competência que exerce o entendimento e a interpretação da informação Estatística apresentada e apoiamos Silva (2007) que percebe essa competência como dependente do raciocínio e do pensamento estatístico. O letramento estatístico é a competência dos indivíduos de ler, escrever, descrever e interpretar informações estatísticas, cujo nível de percepção será mais apurado, conforme a expansão dos níveis de pensamento e raciocínio. Assentimos com Silva (2007) que observa o avanço bilateral entre os níveis do pensamento e do raciocínio estatístico, levando ao

entendimento que essas competências estão mutuamente conectadas. Ao pensamento estatístico, compete o planejamento, a vigia e a generalização durante toda a aplicação de um processo estatístico. O raciocínio estatístico, além da habilidade de trabalhar com as ferramentas e com os conceitos da Estatística, compete interpretar os resultados obtidos por meio dos raciocínios de dedução, indução e educação, constituindo um espírito crítico e científico chamado de espírito estatístico. Elaboramos uma tabela para melhor visualizar as ações e as demandas do letramento, pensamento e raciocínio estatístico.

Tabela 2 - Ações e demandas das competências estatísticas

	LETRAMENTO	PENSAMENTO	RACIOCÍNIO
Ações relacionadas	Ler Escrever Descrver Interpretar	Planejar Generalizar Vigiar	Questionar Explicar Interpretar resultados
Demanda	Conhecimentos de conceitos e métodos estatísticos	Acompanhamento do processo estatístico, reconhecimento da variação	Indução, dedução, educação e espírito estatístico

Fonte: Do autor (2018).

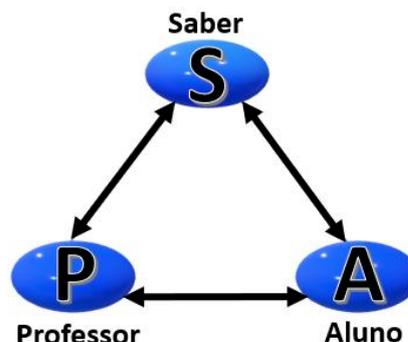
No decorrer das discussões até aqui realizadas, buscamos alguns dos subsídios que nos apoiaram na execução de nossas análises e colaboraram com a construção de um dispositivo pedagógico para a Estatística, definido no próximo item, que foi explorado afim de contribuir de forma agregadora para a Educação Estatística e para um melhor entendimento do processo de construção do raciocínio estatístico durante a aprendizagem da estimação estatística.

2.5. Compreensão de um dispositivo pedagógico para a Estatística

Considerando as reflexões iniciadas pelos relatos das experiências vivenciadas nas pesquisas efetivadas e observadas durante as análises para a composição do nosso Estado da Arte e, juntamente com as ponderações sobre a ciência Estatística e o entendimento das atribuições da Educação Estatística e de suas competências, propomos um modelo de observação que foi denominado de dispositivo pedagógico para melhor compreendermos a construção do raciocínio estatístico por meio dos conceitos que agregam a estimação estatística. Esse modelo de observação foi concebido pelas abstrações que realizamos dos conceitos de dispositivo pedagógico elaborados por Foucault (1995), Bernstein (1996), Larrosa (1995) e

outros, possuindo uma identidade própria como é da natureza de um dispositivo que, segundo Foucault (2000), não é estático mas sim, permanentemente repensado com base no contexto inserido e nas aplicações a que se propõe. Foucault (1995) anuncia o dispositivo pedagógico como uma rede heterogênea envolvendo discursos¹⁷, instituições, instalações, regulamentos e metodologias, que pode assumir uma função estratégica, pensando no sentido de sala de aula, entre quem ensina e quem é ensinado. Bernstein (1996) interpreta um dispositivo pedagógico como um modelo para analisar o processo pelo qual uma disciplina ou um campo científico de conhecimento é transformado ou “pedagogizado” para constituir o conhecimento escolar, o currículo e relações a serem transmitidas. Larrosa (1995) entende o dispositivo pedagógico como um lugar de transformação, onde um indivíduo pode experimentar novos conhecimentos. Diante da complexidade dos aspectos que envolvem a definição de um dispositivo pedagógico, abstraímos a ideia de que é um mecanismo organizado que serve para cumprir uma determinada ação, neste caso dentro de um processo de ensino e aprendizagem, para que o estudante possa assimilar novos conhecimentos e adquirir novas habilidades. Nossa intenção com o dispositivo pedagógico foi de promover um ambiente em sala de aula capaz de gerar informações para a Educação Estatística a partir do ensino da estimação estatística, observando o processo de apropriação desse novo conhecimento e a construção do raciocínio estatístico necessário. Nosso dispositivo encaminhou-se no sentido de atender demandas da didática estatística, buscando os fatores que auxiliam e interferem no processo do ensino e da aprendizagem da Estatística, abrangendo questões de natureza filosófica, epistemológica e pedagógica. Tendo em vista as relações de avizinhamiento entre o ensino da Matemática e o ensino da Estatística, evocamos Brousseau (1986) que vislumbrou, na didática da matemática, uma interação triangular de ensino composta por professor, aluno e o saber.

Figura 20 - Triângulo das situações didáticas



Fonte: Adaptado de Brousseau (1986).

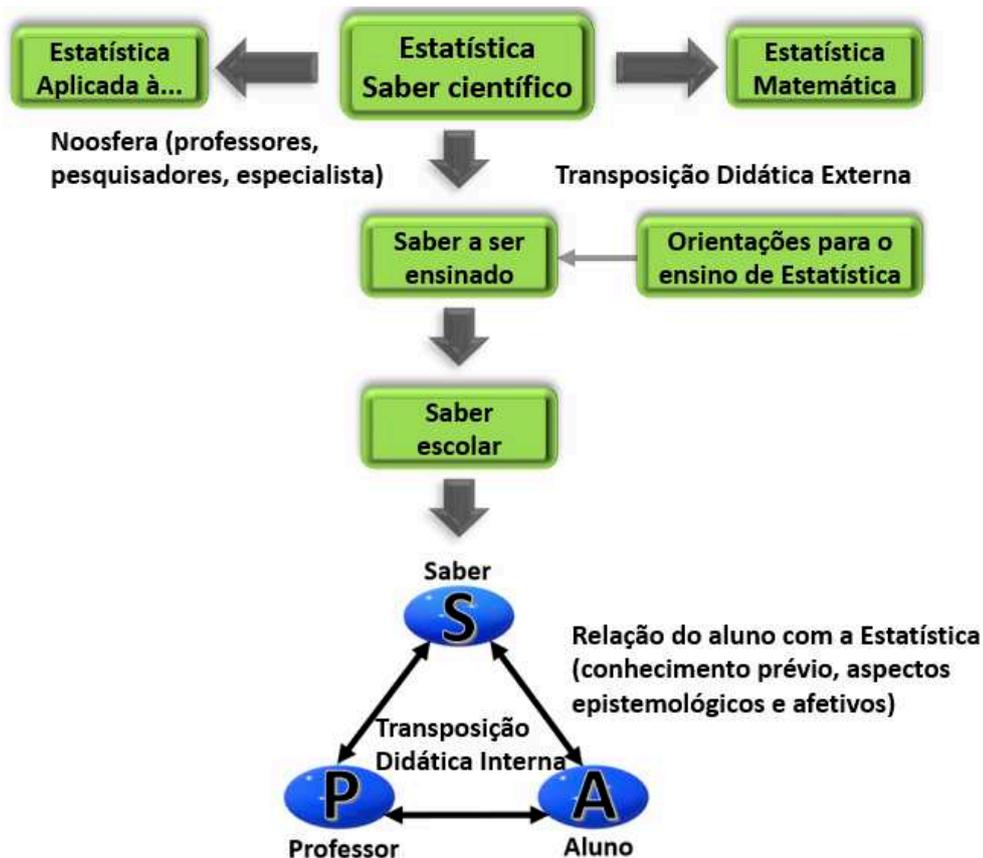
¹⁷ Conjunto de enunciados que constrói conhecimentos (FOUCAULT, 1996).

Nesse triângulo um contrato de ensino é instalado e definido como um conjunto de regras, na maioria das vezes implícitas, que regem uma interação tripartite onde, segundo Brousseau (2000),

Esse contrato define os papéis, os lugares e funções de cada parte. Ele define as atividades esperadas do professor com os alunos, os respectivos locais de preocupação com relação ao conhecimento tratado e até as condições gerais em que essas relações com o conhecimento evoluirão durante o ensino. (tradução nossa)¹⁸

A análise realizada por Brousseau (1986) para a didática matemática, se amplia para outras áreas do conhecimento, conduzindo na Estatística o desenvolvimento da didática da Estatística, como aponta Andrade e Régnier (2009) na elaboração de um esquema representado na Figura 21 a seguir:

Figura 21 - Triângulo das situações didáticas para a Estatística

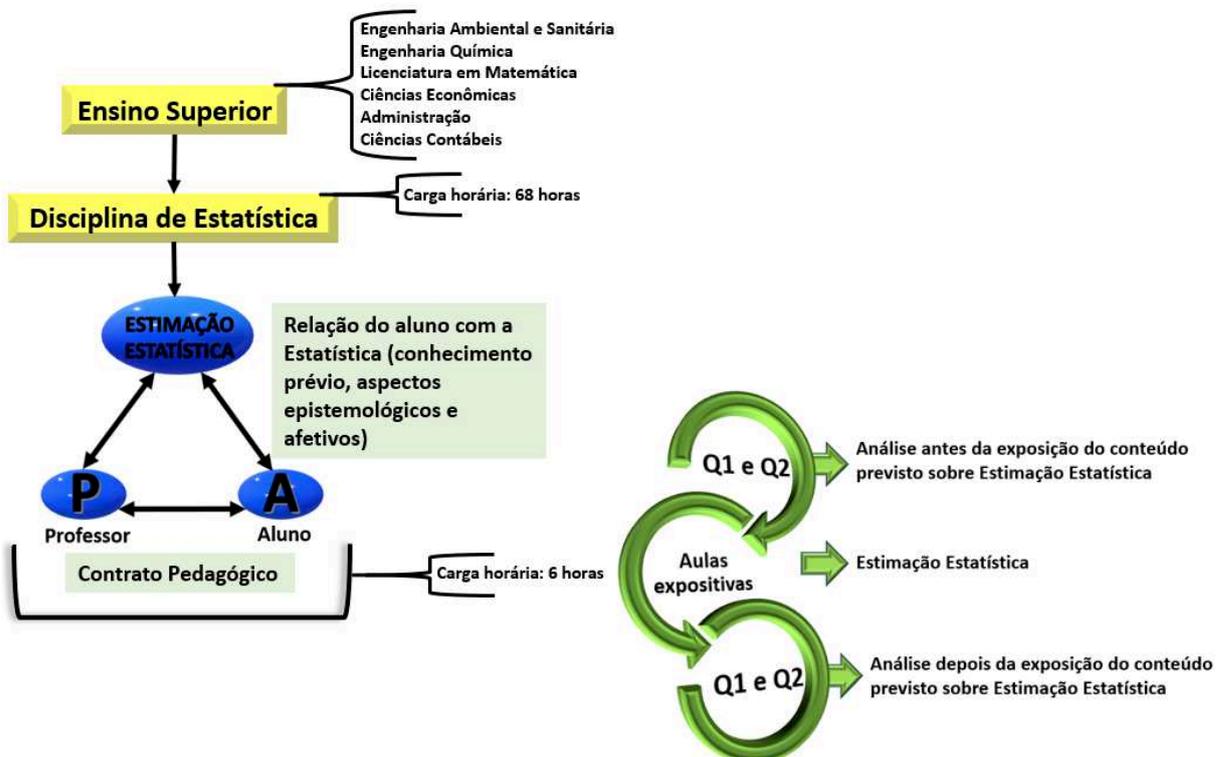


Fonte: Adaptado de Andrade e Régnier (2009).

¹⁸ Ce contrat fixe les rôles, places et fonctions de chaque partie. Il fixe les activités attendues du professeur comme des élèves, les places respectives de chacun au regard du savoir traité, et même les conditions générales dans lesquelles ces rapports au savoir évolueront au cours d'un enseignement.

Os autores supracitados relacionam o Triângulo das Situações Didáticas de Brousseau (1986) com a didática da Estatística, observando a trajetória do saber que é denominada de Transposição Didática. O termo foi introduzido em 1975 pelo sociólogo Michel Verret e explorado por Yves Chevallard em 1985 em seu livro *La Transposition Didactique*, indicando os processos de transposições ocorridas de um determinado saber quando passa do campo científico para o campo escolar. Na fonte desse processo está a Noosfera, que segundo Chevallard (1991), é uma Instituição imaginária, uma “esfera pensante”, composta por pesquisadores, professores, especialistas ligados a Universidades e Ministérios de Educação, que irão definir que saberes (Saber Científico) devem ser ensinados e de que forma devem chegar à sala de aula (Orientações para o ensino). A última etapa do processo, chamada por Chevallard (1991) de Transposição Didática Interna, coloca o professor no papel de transformador desse saber para os alunos, intermediando com eles a sua gestão em sala de aula e definindo as funções que cada um deverá assumir para promover o conhecimento.

Figura 22 - Contexto do Dispositivo Pedagógico



Fonte: Do autor (2018).

Chevallard (1991) expõe, ainda, que em cada esfera de conhecimento o saber é organizado de maneira diferente e vai adquirindo uma característica própria. Diante disso, estabelecemos uma relação para o dispositivo pedagógico, compreendendo que ele pode ser

interpretado como uma esfera que transpõe o saber de acordo com o contexto pretendido. Nosso contexto está relacionado a uma instituição de Ensino Superior que oferece a Estatística como uma disciplina de serviço para vários cursos de graduação. Dessa forma, definimos o dispositivo pedagógico da nossa pesquisa como um conjunto de ferramentas que deverão atender as demandas que a instituição de ensino exige da disciplina de Estatística oferecida em seus diferentes cursos de graduação, organizadas no espaço e no tempo de acordo com uma meta de aprendizagem.

O dispositivo pedagógico foi observado no Ensino Superior inserido numa disciplina de Estatística que possuía a atribuição de desenvolver conhecimentos da estatística descritiva e inferencial, com uma carga horária de 68 horas, comum aos seis cursos de graduação envolvidos na pesquisa. A meta de aprendizagem considerada no dispositivo foi a estimação estatística, conteúdo subsequente ao da estatística descritiva e, portanto, todos os estudantes já possuíam conhecimentos prévios adquiridos no decorrer do desenvolvimento da disciplina. Respeitando o plano de ensino da disciplina de Estatística observado na pesquisa, foram ocupadas 6 horas-aula destinadas para o tópico de estimação, onde ocorreram as seguintes ações:

- a) Aplicação de dois questionários;
- b) Exposição do conteúdo de estimação estatística;
- c) Reaplicação dos questionários.

Anteriormente à exposição do tópico da disciplina de Estatística que abordou a estimação estatística, foram aplicados dois questionários com a intenção de verificar os conhecimentos prévios adquiridos pelos estudantes e outras informações pertinentes para a pesquisa. O primeiro questionário aplicado, que denominamos de Q1 (Anexo A), foi elaborado pelo professor Doutor Jean-Claude Régnier, que o emprega em suas inúmeras pesquisas realizadas na França e em várias partes do mundo na área da Educação Estatística, sendo traduzido e adaptado para o nosso contexto.

Figura 23 - Contexto do questionário Q1



Chevallard (1991) considera que a construção de um novo saber é motivada pela utilização de situações-problema que tratam de conhecimentos antigos e novos. Os antigos, e os prévios conhecimentos, nem sempre conseguem atender a demanda da problemática, motivando a expansão do novo conhecimento e justificando a importância da relação didática entre o saber, o estudante e o professor. Nesse sentido, Q1 foi elaborado com tópicos que investigam o sentimento do estudante em relação a Estatística, observando a afetividade e o grau de importância que é dado para a disciplina. Gal & Ginsburg (1994) destacam que a identificação dos aspectos afetivos e das atitudes positivas e negativas que os estudantes possuem em relação a Estatística, podem promover uma melhor análise durante a investigação do processo de ensino e aprendizagem. O segundo questionário, elaborado pelo autor da tese e denominado de Q2 (Anexo B), foi composto por questões que envolvem elementos da estimação estatística em possíveis questões de enfrentamento da vida cotidiana.

Figura 24 - Contexto do questionário Q2



Fonte: Do autor (2018).

Henry (2005) destaca que uma das principais características da estatística é sua importante ancoragem nas situações do cotidiano. Essas questões não possuem associação com nenhum conhecimento específico dos cursos de graduação pesquisados, objetivando promover um resgate dos conhecimentos já adquiridos pelos estudantes e o entendimento de suas possíveis aplicações em situações comuns. Após a aplicação dos dois questionários, foi realizada a exposição do conteúdo de estimação estatística, compondo, assim, a segunda ação do dispositivo pedagógico. A exposição do conteúdo seguiu o plano de ensino que rege a disciplina de Estatística determinado pela instituição, onde foi abordado a estimação de parâmetros com ênfase na estimação da média e da proporção por meio de aulas expositivas com a utilização de material teórico fornecido pelo professor pesquisador e com recursos computacionais em um laboratório de informática. A última ação realizada foi a reaplicação dos questionários Q1 e Q2, para que pudessemos observar as possíveis mudanças nas respostas

dos estudantes e o desenvolvimento do novo conhecimento. O cruzamento das informações dos dois questionários aplicados, antes e depois da exposição do conteúdo de estimação estatística, nos forneceu elementos para promover discussões que contemplaram questões pertinentes para a Educação Estatística.

3. Inferência Estatística: um olhar sobre a estimação estatística

No sentido comum, a inferência é o raciocínio que deduz uma nova proposta diante de um conjunto de orientações consideradas verdadeiras. Já no quadro da Estatística, é o raciocínio que torna possível descrever uma determinada população a partir de observações feitas em uma amostra. Em particular, é investigar não somente em que medida os valores dos parâmetros, mas, também, as relações, as correlações e as regularidades encontradas em uma amostra são generalizáveis para uma determinada população a ser pesquisada. A estimação estatística é um dos processos estatísticos que atua nessa investigação, avaliando parâmetros por meio de estimadores obtidos em uma amostra, com base no cálculo de probabilidades. Para construir uma estimação, são necessários vários fatores que integram sua concepção, entre eles, conceitos advindos da Estatística Descritiva e da Probabilidade. Nossa pesquisa está voltada para uma análise da estimação pontual e intervalar, sendo essa última dividida entre intervalos de confiança para a média e para a proporção. Para tanto, buscamos reunir alguns elementos pertinentes para a compreensão da estimação estatística, começando pela realização de um breve estudo histórico para contextualizar a sua essência e a exposição de alguns conceitos que promovem o seu uso.

3.1. Alguns elementos sobre a história da Estatística

Conhecer a legítima origem de qualquer ramo do conhecimento não é tarefa simples, pois muitos são os fatores que podem influenciar, ao longo dos anos, na construção de uma definição para um determinado saber que a humanidade utiliza em prol de seu desenvolvimento. Entretanto, organizamos alguns apontamentos que nos dão indícios da essência dessa ciência hoje conhecida como Estatística. Na visão popular a Estatística é abreviadamente compreendida, ainda nos dias de hoje, pela manipulação de dados numéricos que são apresentados e interpretados em quadros, tabelas e gráficos. Embora saibamos que essa é uma visão limitada, os primórdios registros são baseados nessa perspectiva.

A motivação inicial está diretamente centrada no estudo do Estado, com interesse original calcado na coleção de informações referentes a população e a economia. Estas informações eram fundamentais para que os governantes pudessem construir e reger seus programas de governo. Nesse sentido, os primeiros registros do cerne da Estatística estão apontados na Suméria¹⁹, datados por volta de 5000 à 2000 a.C., decorrentes de levantamentos censitários para fins tributários e militares. Echeveste *et al.* (2005) ressaltam as descobertas no Antigo Egito²⁰, 2000 a.C., onde os faraós faziam uso sistemático de informações de caráter estatístico sobre a falta de mão-de-obra relacionada com a construção das pirâmides. Na era pré-colombiana²¹, nas civilizações maias, astecas e incas, também foram encontrados registros tratando do assunto, o que nos fez refletir sobre a extensão mundial da importância dada, ao longo dos anos, para a manipulação de informações quantitativas no desenvolvimento da humanidade e como a sua utilização não foi exclusiva de um determinado período ou região. Em regra geral, uma ciência não nasce do absoluto, mas emerge de um contexto particular descrito, entre tantos outros elementos, pela necessidade humana de evoluir e aperfeiçoar novos conhecimentos. Embora a Estatística tenha sido reconhecida como uma ciência coerente entre os anos de 1885 e 1925, sua participação desde os primórdios da humanidade é notória. Para Droysbeke e Tassi (1997) o nascimento de uma determinada ciência está atrelada ao nível de conhecimento matemático decorrente do momento e da aptidão de seus estudiosos. A Estatística não fugiu dessa lógica, uma vez que seus primeiros grandes pensadores foram, também, vanguardistas na construção da história da Matemática. O termo Estatística é recente, especialmente em comparação com a atividade correspondente de coleta de dados que foram observadas desde a Antiguidade. Droysbeke e Tassi (1997) fazem duas associações em relação ao surgimento da palavra Estatística. A primeira vem da expressão em latim *collegium statisticum*, como eram conhecidas as exposições sobre os assuntos do Estado e, a segunda, com a universidade alemã Göttingen, mais precisamente na pessoa do professor Gottfried Achenwall (1719-1772). A associação com o professor é justificada pelos autores devido ao fato de que ele provavelmente tenha sido um dos primeiros a propor um ensino tratando de Estatística. Até o século XVIII a Estatística foi explorada apenas no âmbito de sua área hoje conhecida como descritiva, onde os estudos eram voltados para revelar as regularidades em acontecimentos de cunho econômico e sociológico. Esse período, como descrevem Magnello e Loon (2016), corresponde ao tipo de Estatística chamada de Estatísticas Vitais que tinha a

¹⁹ Região do sul da Mesopotâmia, atual Iraque e Kuwait, considerada como um dos berços da civilização ocidental.

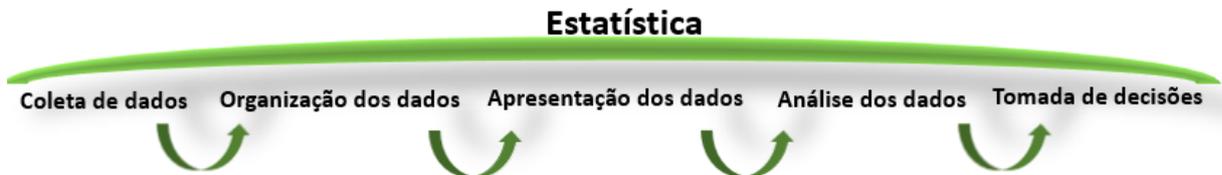
²⁰ O Antigo Egito foi uma civilização da Antiguidade oriental do Norte da África, hoje o país moderno do Egito.

²¹ Considerada a Era marcada antes do aparecimento dos europeus no continente americano.

preocupação de descrever censos populacionais e construir tabelas oficiais que informavam a quantidade de casamentos, divórcios, nascimentos, crimes, entre outras informações sociais. Evidentemente que o sentido da palavra Estatística, enquanto ciência, foi modificando-se ao longo dos anos por meio das concepções científicas necessárias e correspondentes de cada período. Salsburg (2009) expõe que os estudiosos do século XVIII, ficaram muito inquietos ao manipularem e perceberem certas regularidades em dados quantitativos, fazendo com que eles explorassem ferramentas matemáticas apropriadas para enunciar leis e realizar previsões. Inicia-se, assim, no século XIX, o desenvolvimento da Estatística Matemática e suas aplicações. Magnello e Loon (2016), descrevem a origem da Estatística Matemática com os estudos da Teoria das Probabilidades no final do século XVIII e dos trabalhos do suíço Jacob Bernoulli (1654-1705), dos franceses Abraham de Moivre (1667-1754) e Pierre-Simon de Laplace (1749-1827) e do alemão Johann Carl Friedrich Gauss (1777-1855). O século XIX, chegou trazendo uma ciência com a visão crédula e filosófica de que o Universo era regido, de forma regular, como o mecanismo de um enorme relógio. Como informa Salsburg (2009), a ciência acreditava na existência de um número pequeno de fórmulas matemáticas que pudessem descrever a realidade e fazer previsões de eventos futuros. Kendall, Kendall e Stuart (1997), mencionam que a humanidade precisou de centenas de anos para se acostumar com um mundo onde vários eventos não tinham uma razão definida ou eram determinados por fatores de possibilidades tão remotas que somente podiam ser razoavelmente representados por modelos probabilísticos. No final do século XIX, muitos erros haviam sido revelados em função da discrepância entre o observado e o previsto. Nas antigas ciências, como física e química, as leis de Isaac Newton (1643-1727), como relata Salsburg (2009), mostravam-se como meras e grosseiras aproximações. O período entre 1900 e 1915, como narram Driesbeke e Tassi (1997), foi considerado de transição entre a visão original e a nova visão de Estatística que necessitava de técnicas matemáticas, probabilidades e sofisticados métodos de estudos para a interpretação de certas informações. Dessa maneira, gradualmente, a ciência passou a trabalhar com um novo paradigma para modelos estatísticos. Na virada do século XIX a evolução da Estatística tomou uma nova direção, pois o conceito dessa ciência proporcionou o uso mais amplo da matemática e suas aplicações para diversas áreas do conhecimento. Dessa forma, foi edificada o que hoje chamamos de Estatística Inferencial, com uso sistemático da probabilidade nos papéis definidos de coleta e resumo de dados empíricos. Salsburg (2009) complementa que, ainda no final do século XX, quase toda a ciência passou a usar modelos estatísticos. Surgiram grandes pesquisadores no assunto, entre eles, Ronald Aylmer Fisher (1890-1962) reconhecido por muitos como o maior estatístico do século. Reconhecemos a Estatística, nos dias atuais, como

uma ciência que possui a arte de coletar, analisar, interpretar e tirar inferências de dados quantitativos. Segundo Freund (2006) tudo que se tratar, por pouco que seja, de coleta, processamento, interpretação e apresentação de dados, pertence ao domínio da Estatística.

Figura 25 - Domínios da Estatística



Fonte: Do autor (2018).

Hodiernamente a Estatística está centrada em função de duas grandes áreas, sendo uma conhecida como Estatística Descritiva e a outra como Estatística Inferencial.

Figura 26 - Grandes áreas da Estatística



Fonte: Do autor (2018).

A primeira trabalha com números que descrevem fatos, tornando questões complexas mais fáceis de serem observadas para um melhor entendimento e, a segunda, atua no campo da tomada de decisões, fazendo, por exemplo, estimativas que são, justamente, nosso objeto de estudo nesta pesquisa.

3.2. Alguns elementos sobre a história da estimação estatística

Dentro das duas principais grandes áreas da estatística, tratamos, nesta pesquisa, daquela que está preocupada nas tomadas de decisões, a Estatística Inferencial, especificamente o que versa a estimação estatística. Tentarmos situar o surgimento da estimação na história da humanidade seria muito pretencioso da nossa parte, tendo em vista que o ato de estimar, no sentido de ponderar algum acontecimento, é algo inerente ao ser humano. A todo momento realizamos algum tipo de previsão, seja olhando para o céu prevendo se choverá, ou não, ou

avaliando um time de futebol e prognosticando o seu sucesso em uma partida ou, simplesmente, avaliando se o tempo que possuímos será suficiente para a realização de alguma tarefa. Entretanto, podemos fazer algumas alusões na história da humanidade em que houve a prática de estimações. De acordo com Ponce (1986), “Havia no Egito antigo um dispositivo admirável para a época que permitia conhecer com boa exatidão o crescimento das águas do rio Nilo e prognosticar o volume da futura colheita”. Os egípcios estimavam a extensão da área útil de suas propriedades após cada inundação anual no vale do rio Nilo, para o pagamento de impostos e a plantação de seus cultivos. Na Antiguidade, os cultivos eram realizados estimando o suficiente a ser plantado para a subsistência da população e o pagamento de impostos. Podemos citar, ainda, as previsões de Thomas Robert Malthus (1766-1834) que criou uma teoria demográfica fundamentada na relação entre o crescimento da população e o crescimento da oferta de alimentos para explicar uma possível falta de comida no futuro. Malthus afirmava que a oferta de alimentos não acompanharia a tendência do crescimento populacional. Contudo, Dreesbeke e Tassi (1997) acrescentam fatos pontuais sobre a trajetória da história da estimação estatística levando-se em consideração a cientifização das técnicas formais para a sua prática. Ronald Aylmer Fisher (1890-1962) é reconhecido pela escrita de artigos que popularizaram o uso da estimação estatística no meio científico. Entretanto, Dreesbeke e Tassi (1997) declaram que não seria correto destinar o surgimento formal da estimação estatística para Fisher, uma vez que as primeiras ideias surgiram desde o início do cálculo de probabilidades. Os mesmos autores trazem uma cronologia de obras que já apontavam o uso da estimação, como o caso da obra de Jacques Bernoulli (1654-1705), intitulada de *Ars Conjectandi* (1713)²². Alguns anos mais tarde, uma obra de Thomas Bayes (1701-1761) publicada dois anos após a sua morte em 1763²³, registra o uso da estimação com uma linguagem de notações mais modernas. Essas notações são normalmente utilizadas até os dias de hoje, como, por exemplo, a letra grega Teta θ representando os parâmetros e, a variável X , como representação de uma variável aleatória. Ainda, cronologicamente, Dreesbeke e Tassi (1997) assinalam os trabalhos de Pierre Simon de Laplace (1749-1827), *Mémoire sur la probabilité des causes* (1774) e *Théorie analytique des probabilités* (1812). Os trabalhos de Laplace estão a mais de cem anos antes dos artigos que Fisher escreveu tratando da estimação, e serviram de base para fundamentar todos os seus trabalhos. Dreesbeke e Tassi (1997) revelam que muitos estatísticos consideraram justo dar o

²² Obra publicada oito anos após a sua morte pelo seu sobrinho Nicolas Bernoulli e descreve a Lei dos Grandes Números e o Teorema Dourado, considerado um marco na história das probabilidades.

²³ Obra intitulada de *An Essay towards solving a Problem in the Doctrine of Chances*, que tratou de ensaios para a resolução de problemas de probabilidade.

título de primeiro teórico da estatística moderna para Laplace. No final do século XIX, em 1894, surge os trabalhos de Karl Pearson (1857-1936), utilizando a distribuição normal²⁴ para efetuar os cálculos dos parâmetros populacionais. A etapa seguinte da trajetória da estimação estatística está marcada em 1922 e 1925 com as publicações de dois artigos de Fisher, *On the mathematical foundations of theoretical statistic* e *Theory of statistical estimation*, respectivamente, onde houve uma popularização científica do uso da estimação. No primeiro artigo, Fisher expõe que um dos propósitos da estatística é a prática da redução de dados. Uma massa de dados deve ser substituída por uma pequena quantidade que represente corretamente essa massa, com o máximo de informações relevantes contidas nos dados originais. Expõe, ainda, que o objetivo é fazer o uso dessa pequena quantidade representativa da população, chamada de amostra, para a retirada de inferências, por meio da construção de estimativas para parâmetros populacionais. Na década de trinta surge a técnica da estimação intervalar, descrita em um artigo intitulado “*The dependence of the income in small farms upon their area, the outlay and the capital invested in cows*”, do agrônomo polonês Waclaw Pytkowski (1904-1989). Entretanto, Dreesbeke e Tassi (1997) alertam que a teoria dos intervalos de confiança é de autoria do matemático polonês Jerzy Neyman (1894-1981), o qual expôs suas primeiras ideias no apêndice de um artigo que foi publicado em 1934, definitivamente desenvolvido e apresentado nos artigos “*Outline of a theory of statistical estimation based on the classical theory of probability*” em 1937 e “*L’estimation statistique traitée comme un problème classique de probabilité*” em 1938. Dessa forma relatamos de maneira concisa uma ideia da trajetória histórica da estimação estatística, onde podemos observar que seu caminho está diretamente ligado a história das teorias probabilísticas.

3.3. Alguns elementos sobre a história da probabilidade

Como observamos no texto anterior, a história da estimação estatística está atrelada na história da probabilidade e, sendo a estimação um campo da Estatística Inferencial, “falar de inferência estatística sem evocar a evolução do conceito de probabilidade é um absurdo” (DROESBEKE; TASSI, 1997, p. 22, tradução nossa²⁵). A estimação estatística utiliza as teorias probabilísticas para analisar e compreender a frequência da ocorrência de eventos, tanto em estudos observacionais quanto em experimentos que modelam a aleatoriedade e a incerteza, de

²⁴ É uma distribuição de probabilidades, também conhecida como Lei Normal, Lei Laplace-Gauss, utilizada para modelar fenômenos naturais e será visitada no decorrer desse trabalho.

²⁵ Parler d’inférence statistique sans évoquer l’évolution du concept de probabilité est un non-sens.

forma a realizar previsões de fenômenos futuros. Originalmente, a probabilidade está relacionada aos jogos de azar²⁶. Entre os mais antigos, citamos o Tali, um jogo praticado com o astrágalo²⁷ e, de acordo com David (1962), é o antepassado do nosso atual jogo de dados, onde a transição foi ocorrendo ao longo de dois mil anos. Os indícios desse jogo aparecem na história da civilização egípcia, primeira dinastia, datada de 3500 a.C. Bellhouse (2000) indica que um dos primeiros documentos que apresenta um raciocínio probabilístico está em um poema chamado “De Vetula”, escrito no ano de 1250 por um erudito eclesiástico francês, Richard de Fournival²⁸ (1201-1260). Esse poema descreve um cálculo de combinações referentes ao lançamento de três dados. Entretanto, Viali (2008) ressalta que as primeiras ideias que formaram a base para o desenvolvimento científico da probabilidade ocorreram bem mais tarde, como no trabalho de Jérôme Cardano (1501-1576), intitulado de “Liber de ludo aleae” escrito por volta do ano de 1526 e publicado em 1663, na cidade de Lyon, na França. Outro trabalho a ser citado é o de Galileu Galilei (1564-1642), “Sulla scoperta dei dadi”, que foi publicado apenas no ano de 1656, e expôs cálculos com raciocínios probabilísticos utilizando permutações. Contudo, Driesbeke e Tassi (1997) destacam o trabalho considerado como o precursor no estudo das teorias probabilísticas, intitulado de “Summa de arithmetica, geometria, proportioni et proportionalita”, publicado no ano de 1494, cujo autor é Luca Bartolomeo Pacioli²⁹ (1445-1517). O trabalho de Pacioli apresentou um problema conhecido como o problema dos pontos³⁰ na divisão de uma aposta e, como descreve Pichard (1997), contribuiu, de forma fundamental, na concepção do cálculo de probabilidades. Montucla (1968) lembra que os nomes de Blaise Pascal (1623-1662) e Pierre Fermat (1601-1665) são reconhecidos como sendo os pioneiros no desenvolvimento das teorias probabilísticas, justamente em função das tratativas da solução do problema exposto no trabalho de Pacioli. Antoine Gombaud³¹ (1607-1684) foi quem apresentou e solicitou uma solução do problema dos pontos para Pascal, além de outros problemas sobre jogos de dados praticados na época. Pascal, por sua vez, pediu auxílio para Fermat o que motivou uma célebre troca de correspondências, entre eles, envolvendo sete cartas contendo informações para a solução de problemas. Lightner (1991) descreve que, na terceira carta, datada em 29 de julho de 1654, Pascal responde para Fermat

²⁶ Segundo Houaiss : ár. *az-zahr* 'flor', p.ext. vulg. 'dado' porque se pintava uma flor numa das faces do dado.

²⁷ Um pequeno osso do calcanhar de um animal específico que possui a forma semelhante de um tetraedro irregular.

²⁸ Médico e poeta francês.

²⁹ Também conhecido como Frei Luca dal Borgo, monge franciscano e célebre matemático italiano.

³⁰ O problema apresenta a seguinte situação: um jogo equitativo termina quando um dos jogadores vence seis partidas. Suponha-se que por algum motivo o jogo tenha que ser interrompido, e neste momento, um dos jogadores venceu em cinco partidas e um outro jogador em apenas três. Como as apostas devem ser repartidas?

³¹ Conhecido como Chevalier de Méré, foi um nobre jogador francês. Seu nome é relacionado ao cálculo matemático de jogos de azar.

concordando e complementando com a sua solução encontrada, cuja ideia principal foi o princípio do que hoje conhecemos como a definição clássica da probabilidade, tratando da razão entre a enumeração dos casos favoráveis sobre os casos possíveis de um jogo de azar. No mesmo ano, como aponta Mesnard (1991), Pascal endereça uma carta para a Academia Parisiense destacando a importância da descoberta que foi denominada como a Geometria do Acaso, mas não propôs uma definição explícita para a probabilidade. Dessa forma, Droesbeke e Tassi (1997) ressaltam que vincular o nome de alguém como o autor da criação de um conceito é uma tarefa difícil. No que diz respeito a probabilidade, já haviam registros das primeiras tentativas anteriormente ao ano de 1654. Outras obras haviam sido publicadas, na mesma época, como a de Christiaan Huygens (1629-1695)³² escrita entre 1650 e 1660, intitulada “*Ratiociniis in ludo aleae*”. Huygens introduziu e utilizou, em sua obra, a noção de esperança matemática e a amostragem com, e sem reposição de elementos dentro de uma amostra e, além disso, “Sua reflexão sobre o conceito de probabilidade parece ter tido mais influência no próximo século do que a de Pascal e Fermat” (DROESBEKE; TASSI, 1997, p. 24, tradução nossa³³). Com o passar dos anos, a teoria das probabilidades foi superando o marco original dos jogos de azar passando a ser um ramo da matemática pura, com aplicações nas mais diversas áreas das ciências.

É notável que uma ciência que tem sua origem nos estudos sobre jogos de azar pudesse ter se elevado ao nível dos mais importantes assuntos do conhecimento humano. As questões mais importantes da vida são, em sua maioria, essencialmente problemas de probabilidade (LAPLACE, 1840, p. 1 e 2, tradução nossa³⁴)

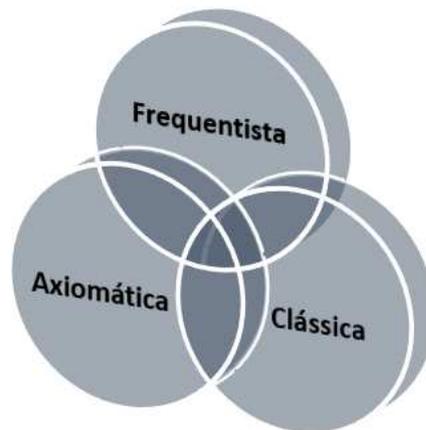
Atualmente, como apontam Godino, Batanero e Cañizares (1996), encontramos diversas formas de conceber o conceito de probabilidade, entre as quais destacamos a clássica, a frequentista e a axiomática.

³² Foi um físico, matemático, astrônomo neerlandês. Inventou o relógio de pêndulo dito de Huygens.

³³ Sa réflexion sur le concept de probabilité semble avoir eu plus d'influence dans le siècle qui suivra que celle de Pascal et de Fermat.

³⁴ Il est remarquable que cette science, qui a pris son origine dans l'étude des jeux de chance, soit devenue l'objet le plus important de la connaissance humaine. Les questions les plus importantes de la vie ne sont en réalité, pour l'essentiel, que des problèmes de probabilités.

Figura 27 - Conceitos de probabilidade



Fonte: Do autor (2018).

Para realizarmos uma breve apresentação dos conceitos de probabilidade que foram mencionados, utilizamos algumas definições na linguagem matemática atual, com o intuito de facilitar a compreensão das teorias probabilísticas que objetivam o estudo dos fenômenos aleatórios. Diante disso, definimos fenômeno aleatório como sendo um acontecimento que, quando observado repetidamente, sob as mesmas condições, produz resultados distintos. Para exemplificar, podemos citar o lançamento de um dado honesto³⁵. De acordo com Ardilly (2004) quando a possibilidade de repetir o fenômeno está nas mãos do experimentador, o fenômeno aleatório é denominado de experimento aleatório. Outra definição importante é a de espaço dos resultados, ou espaço amostral, que corresponde ao conjunto de todos os resultados possíveis de um fenômeno aleatório, normalmente denotado pela letra maiúscula grega ômega Ω . Encerramos definindo evento, ou acontecimento, que é todo subconjunto de um espaço amostral, simbolicamente representado por uma letra maiúscula latina

3.3.1. A definição Clássica de probabilidade

A noção clássica de probabilidade, principiada por volta da metade do século XVII por Fermat e Pascal, foi definida, pela primeira vez, na obra “Teoria Analítica das Probabilidades”, de Pierre-Simon Laplace (1749-1827), em 1812, baseada no modelo de jogos de azar com resultados igualmente prováveis. Formalmente, a probabilidade era reduzida ao elemento quociente entre dois números inteiros e positivos, onde o numerador indica o número de casos favoráveis à ocorrência de um evento, e o denominador indica o número de casos possíveis.

³⁵ Ao nos referirmos sobre dado honesto, dado equilibrado, ou, ainda, moeda honesta, moeda equilibrada, estamos admitindo que todas as suas faces possuem a mesma probabilidade de resultado em um lançamento.

Diante disso, podemos apresentar a definição clássica de probabilidade do seguinte modo: para a descoberta da probabilidade de um evento A , que é um subconjunto finito de um espaço amostral Ω , de resultados igualmente prováveis, usamos

$$P(A) = \frac{\text{número de elementos de } A}{\text{número de elementos de } \Omega} = \frac{n(A)}{n(\Omega)}, \text{ com } 0 \leq P(A) \leq 1$$

Sendo assim, a função do conjunto $P(A)$ satisfaz as seguintes condições:

- a) $P(A) \geq 0$, para todo $A \subset \Omega$;
- b) Se A e B são eventos mutuamente exclusivos, então $P(A \cup B) = P(A) + P(B)$;
- c) $P(\Omega) = 1$.

Demonstração:

- a) Como $n(\Omega) > 0$ e $n(A) \geq 0$, segue que $P(A) \geq 0$.
- b) Se A tem n eventos simples e B possui m eventos simples, como A e B são eventos mutuamente exclusivos, então significa que não têm eventos simples em comum, logo o número de eventos de $A \cup B = n + m$.
- c) Como o número de eventos simples de Ω é $n(\Omega)$, então $P(\Omega) = 1$.

Para exemplificar, tomemos a realização de um sorteio com bilhetes numerados de 1 a 100, perguntando qual a probabilidade de sortearmos um múltiplo de três. Desse modo, temos um espaço amostral $n(\Omega) = 100$ e um evento $n(A) = 33$, logo $P(A) = \frac{33}{100} = 0,33$ ou 33%.

As aplicações para a definição clássica de probabilidade são limitadas, devendo-se isso ao fato de que todos os casos possíveis devem ter a mesma probabilidade ou, em termos atuais, a equiprobabilidade dos elementos do espaço dos resultados, gerando uma tautologia, pois está definindo essencialmente a probabilidade com seus próprios termos. Outra restrição encontrada está relacionada com a finitude do espaço de resultados, não sendo possível trabalhar quando o espaço amostral é infinito. A definição clássica nos levaria a uma indeterminação do tipo $\frac{\infty}{\infty}$, quando fôssemos trabalhar com alguma variável de tipo contínuo, por exemplo. Diante dessas limitações encontradas, fez-se necessária uma nova abordagem do fenômeno, afim de estabelecer um conceito mais abrangente.

3.3.2. A definição frequentista de probabilidade

De acordo com Coutinho (2007) o livro de Bernoulli, “Ars Conjectandi”, apontou o fato de que poucos eram os casos onde a definição clássica de probabilidade poderia ser utilizada. Na busca de ampliar o conceito de probabilidade, Bernoulli propõe uma interpretação frequentista ou experimental, que consiste na observação das várias repetições de um experimento realizado sob as mesmas condições. Com um grande número de repetições, podemos obter uma boa estimativa para a probabilidade da ocorrência de um determinado evento, é o que Bernoulli justificou por meio da Lei dos Grandes Números³⁶, conhecida pelo nome de Teorema de Bernoulli. Assim, como a probabilidade clássica, a probabilidade frequentista utiliza uma razão que consiste no número de vezes, dentre o total de repetições efetuadas, nas quais o evento aconteceu, com o número total de repetições do experimento aleatório. Em 1928, Richard Edler von Mises (1883-1953), em seu livro intitulado “Probability, Statistic and Truth”, apresentou uma formalização da probabilidade frequentista utilizando o limite de uma função. A definição para a probabilidade frequentista pode ser escrita do seguinte modo: seja ε um experimento aleatório, onde, sob as mesmas condições teóricas, são realizados n ensaios independentes do experimento ε , com n suficientemente grande. Em cada ensaio o evento A pode ou não ocorrer. Seja $n(A)$ o número de ocorrências do evento A , temos:

$$P(A) = \lim_{n \rightarrow \infty} \frac{n(A)}{n}, \text{ com } 0 \leq P(A) \leq 1, \text{ onde } n(A) \text{ é o número de vezes que } A \text{ ocorreu}$$

em n ensaios do experimento ε , n é o total de repetições do experimento ε e $\frac{n(A)}{n}$ é frequência relativa $f_r(A)$ do evento em ε . Ao fazermos um grande número de repetições experimento, nas mesmas condições, de modo que as repetições não dependam dos resultados anteriores, observamos a formação de um padrão no sentido de uma estabilização, ou seja, a frequência relativa de ocorrência do evento A tende para um valor constante P , o que consideramos, então, como a probabilidade de ocorrência do evento A . Importante observar que a frequência relativa do evento A é uma aproximação da probabilidade de A . As duas se igualam apenas no limite. Em geral, para um valor razoavelmente grande de n , torna a $f_r(A)$ uma boa aproximação de $P(A)$, o que pode ser entendido pela Lei dos Grandes Números. A definição de probabilidade frequentista admite um lema análogo ao da probabilidade clássica, desse modo, a frequência

³⁶ É um teorema fundamental da teoria da probabilidade que descreve o resultado da realização da mesma experiência repetidas vezes. A média aritmética dos resultados obtidos tende a ser uma aproximação do valor esperado.

relativa $f_r(A)$ definida na classe de eventos do espaço amostral Ω satisfaz as seguintes condições:

- a) Para todo evento A , $0 \leq f_r(A) \leq 1$;
- b) Se A e B são dois eventos mutuamente exclusivos, então $f_r(A \cup B) = f_r(A) + f_r(B)$;
- c) $f_r(\Omega) = 1$.

Essa definição, embora útil, assim como a definição clássica, apresenta algumas críticas. Dentre elas, a vaga ideia de repetirmos um experimento n vezes, ou $n \rightarrow \infty$. Podemos questionar quantas repetições serão necessárias para obtermos um bom resultado. Gnedenko (2008) menciona que durante as repetições de um experimento aleatório, a frequência relativa de um evento tende a se estabilizar. Esse valor de estabilidade é uma estimativa probabilística do evento e é independente de quem realiza o experimento. Destacamos que a definição clássica e a definição frequentista de probabilidade, antes de serem definições formais, são consideradas como formas de se calcular probabilidades.

3.3.3. A definição axiomática de probabilidade

Toda matemática passou por um processo de axiomatização³⁷ a partir da segunda metade do século XIX. A probabilidade, por sua vez, tornou-se um instrumento eficaz do conhecimento matemático, recebendo, também, axiomas. Em 1933, o matemático russo Andrei Nikolaevich Kolmogorov (1903-1987) apresentou os axiomas para a probabilidade, “[...] desenvolveu sua teoria axiomática em seu livro intitulado *Grundbegriffe der Wahrscheinlichkeitsrechnung*, inspirado no desenvolvimento da teoria da medida.” (DROESBEKE; TASSI, 1997, p. 37, tradução nossa³⁸). No prefácio do seu livro, cujo o título traduzido é *Conceitos Básicos da Teoria da Probabilidade*, Kolmogorov revela que objetivou explicitar e sistematizar o conjunto de axiomas que já estavam sendo utilizados, embora de forma implícita, pela maioria dos teóricos contemporâneos que tratavam sobre *Cálculo de Probabilidades*. Kolmogorov propôs uma abordagem axiomática em que a probabilidade é definida como uma função de conjuntos que satisfaz as propriedades da frequência relativa e estende o conceito de probabilidade para situações mais gerais. Ele incorporou a matemática da probabilidade baseando-se no contexto da teoria geral de conjuntos de pontos mensuráveis,

³⁷ Termo empregado para designar questões que não necessitam de nenhuma prova ou comprovação para serem aceitas como válidas, são verdades inquestionáveis.

³⁸ “[...] développer sa théorie axiomatique dans son ouvrage intitulé *Grundbegriffe der Wahrscheinlichkeitsrechnung*, en s’inspirant du développement de la théorie de la mesure.

conhecida atualmente como teoria da medida, definindo a probabilidade como uma medida em que a probabilidade do universo é unitária. Boyer (1974) observa que Kolmogorov não construiu conceitos de probabilidade e, sim, fundamentos matemáticos para o cálculo de probabilidade por meio de funções mensuráveis e integração, ou seja, uma estrutura matemática para a teoria de probabilidade, de um modo análogo ao da Geometria de Euclides. A axiomática de Kolmogorov contém ferramentas matemáticas necessárias para o estudo da probabilidade, mas não fundamenta uma ligação deste conceito com a realidade, ou seja, não realiza aplicações. A definição axiomática de Kolmogorov pode ser escrita do seguinte modo: seja Ω um conjunto não vazio. Uma probabilidade em Ω é uma função do conjunto \mathcal{P} que associa a subconjuntos A de Ω um número real $P(A)$ que satisfaz:

- a) Para todo $A \subseteq \Omega$ vale que $0 \leq P(A) \leq 1$;
- b) $P(\Omega) = 1$;
- c) Se $A \subseteq \Omega$, $B \subseteq \Omega$ e $A \cap B = \emptyset$ então $P(A \cup B) = P(A) + P(B)$.

Segue o teorema: seja Ω um conjunto não vazio e P uma probabilidade em Ω , então são verdadeiras as seguintes afirmações:

- a) $P(\emptyset) = 0$;
- b) $P(A^c) = 1 - P(A)$;
- c) Se $A \subseteq B$, então $P(A) \leq P(B)$ e $P(B - A) = P(B) - P(A)$;
- d) $P(B - A) = P(B) - P(B \cap A)$;
- e) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Demonstrações:

- a) $P(\emptyset) = P(\emptyset \cup \emptyset) = P(\emptyset) + P(\emptyset)$, ou seja, $P(\emptyset) = P(\emptyset) + P(\emptyset)$, logo $P(\emptyset) = 0$;
- b) $\Omega = A \cup A^c$ e $A \cap A^c = \emptyset$, então $P(A \cup A^c) = 1$, de onde extraímos $P(A^c) = 1 - P(A)$;
- c) Se $A \subseteq B$ então $B = A \cup (B - A)$ e, portanto, $P(B) = P(A) + P(B - A)$, desta igualdade decorrem que $P(A) \leq P(B)$ e $P(B - A) = P(B) - P(A)$;
- d) $B - A = B - B \cap A$, resultado decorrente de VI;
- e) $A \cup B = A \cup (B - A)$, sendo $A \cap (B - A) = \emptyset$, o resultado decorre do axioma III e da propriedade VII.

A definição axiomática de probabilidade é uma definição matemática, mas não estabelece o valor associado ao evento de Ω . A determinação da probabilidade de um evento depende da declaração de um modelo apropriado para o experimento aleatório considerado. Observamos, ainda, que a definição clássica de probabilidade pode ser alcançada considerando

um caso particular da definição axiomática de Kolmogorov, considerando um universo finito e supondo acontecimentos elementares equiprováveis. Apresentadas as definições de probabilidade, ressaltamos o estudo dessa importante base para a compreensão da estatística, principalmente no que diz respeito à inferência estatística. A combinação entre a probabilidade e a estatística por meio de métodos científicos torna-se, cada vez mais, uma importante ferramenta para o estudo e interpretações de fenômenos aleatórios que são impactantes na vida da sociedade moderna.

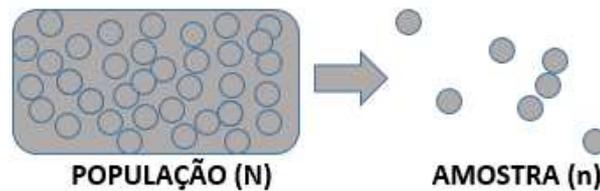
3.4. Breve estudo sobre alguns conceitos fundamentais da Estatística para situar a estimação

Como já mencionado, é necessário verificarmos alguns conceitos advindos da Estatística Descritiva e da Probabilidade para a construção de uma estimação. Dessa forma, procuramos evidenciar aqueles que consideramos ser os mais pertinentes para a compreensão da estimação de parâmetros populacionais que serão tratados por meio da estimação pontual, dos intervalos de confiança da média, e da proporção.

3.4.1. População e amostra

Para a realização de uma pesquisa é necessário levantar informações que possam responder ao questionamento que é proposto pela investigação. Nesse sentido, o início básico de uma pesquisa é a coleta dessas informações. Uma pesquisa censitária, por exemplo, como apresentam Lethielleux e Chevalier (2016), designa a extenuante coleta de informações de todo conjunto de uma determinada população a ser estudada. Alcançar um levantamento censitário é, geralmente, muito oneroso e demorado, optando-se por trabalhar com partes da população que são denominadas de amostras. Freund (2006) assinala que na Estatística o termo população, representado convencionalmente por uma letra maiúscula do alfabeto, é empregado para designar um conjunto de objetos, indivíduos ou resultados experimentais acerca do qual se pretende estudar alguma característica comum. Essa população, existente ou conceptual, pode ser finita quando possui um número limitado de elementos que é passível de contagem, ou infinita quando apresenta um número ilimitado de elementos que não são passíveis de contagem.

Figura 28 - População e amostra



Fonte: Do autor (2018).

O termo amostra, como descreve Mountassir (2016), simbolizado por uma letra minúscula do alfabeto, possui o sentido de subconjunto finito extraído de uma população. Para ser representativa, a amostra tem que apresentar a capacidade de reproduzir as mesmas características importantes da população de origem. A representatividade da amostra está intimamente ligada ao seu tamanho.

3.4.2. Variáveis aleatórias

Quando uma variável tem resultados ou valores que tendem a variar de uma observação para outra, em razão de fatores relacionados com a probabilidade, chama-se variável aleatória. Uma variável aleatória é uma função com valores numéricos, cujo valor é determinado por fatores de probabilidade. De modo geral, definimos uma variável aleatória associada a uma amostra. Nesse sentido, considerando X uma variável populacional que se deseja estudar, uma amostra aleatória de X é o conjunto de n variáveis aleatórias independentes (X_1, X_2, \dots, X_n) de forma que cada uma delas possui a mesma característica da variável populacional.

Figura 29 - Tipos de variáveis aleatórias



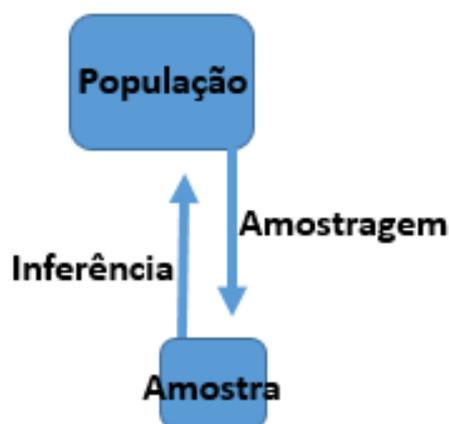
Fonte: Do autor (2018).

As variáveis aleatórias são discretas ou contínuas. Uma variável aleatória é considerada discreta se toma valores que podem ser contados e, contínua, quando pode tomar qualquer valor de determinado intervalo, e portanto necessita o conceito de densidade de probabilidade.

3.4.3. Uma síntese sobre Amostragem

Os processos de amostragem, como aponta Ardilly (2004), são praticados no mundo todo por pesquisadores das mais diferentes áreas do conhecimento na realização de pesquisas científicas sendo, na maior parte dos casos, a única forma de obter informações sobre uma determinada realidade a ser pesquisada. Stevenson (2001) descreve a amostragem como o processo de extração da amostra de uma população a ser pesquisada. Salienta, ainda, que pode haver alguma confusão entre sondagem e amostragem. No entanto, a amostragem é apenas uma das fases da sondagem, haja visto que seu objetivo é selecionar amostra de uma população para um estudo estatístico. Dessa forma, definimos a sondagem como o estudo estatístico de uma população realizado por meio de uma amostra. A amostragem, por sua vez, é uma fase da sondagem relativa ao processo da coleta de amostra de uma população. Refletimos, ainda, que a amostragem é um procedimento que visa selecionar parte da população a ser investigada, a qual chamamos de amostra, para que possamos estimar uma ou mais características pretendidas e, assim, inferir sobre a totalidade dessa população como simula a Figura 30 a seguir:

Figura 30 - Representação de amostragem



Fonte: Do autor (2018).

A amostragem possui dois métodos para a seleção da amostra, quais sejam, os probabilísticos e os não probabilísticos. Os probabilísticos garantem que todos os elementos de uma população possuem a mesma probabilidade de ser escolhidos, ao contrário dos não

probabilísticos. Os processos de amostragem podem apresentar riscos, pois uma amostra que não seja representativa de uma população pode gerar interpretações errôneas durante uma análise estatística. Stevenson (2001) adverte que esses riscos poderão ser conhecidos e interpretados pela Estatística quando selecionamos amostras por meio de critérios probabilísticos, isto é, ao acaso. Dessa forma, nossa pesquisa está sob a perspectiva dos processos de amostragem probabilísticos, mais especificamente dos processos de amostra aleatória simples com e sem reposição de elementos. Uma amostra aleatória simples é uma amostra de tamanho n planejada a partir de uma população de tamanho N de tal maneira que todas as possíveis amostras de tamanho n possuem a mesma probabilidade de seleção. Sendo N o tamanho de uma população e $n < N$ o tamanho de uma amostra, os números possíveis de amostras aleatórias simples, considerando os critérios com e sem reposição de elementos, está descrito a seguir:

Com reposição, temos $k = \text{número de amostras} = N^n$

Sem reposição, temos $k = \text{número de amostras} = \binom{N}{n} = \frac{N!}{n!(N-n)!}$

Para exemplificar, tomamos uma população $N = (A, B, C, D)$ de onde iremos obter amostras de tamanho $n = 2$ com e sem reposição de elementos. Dessa forma, para efetuar uma retirada sem reposição, temos $N = 4$ e $n = 2$, onde $\binom{N}{n} = \binom{4}{2} = \frac{4!}{2!(4-2)!} = \frac{24}{4} = 6$ amostras. As possíveis 6 amostras seriam: (A, B) , (A, C) , (A, D) , (B, C) , (B, D) e (C, D) .

Com reposição, temos $N = 4$ e $n = 2$ onde $N^n = 4^2 = 16$ amostras. As 16 amostras possíveis seriam: (A, A) , (A, B) , (A, C) , (A, D) , (B, B) , (B, A) , (B, C) , (B, D) , (C, C) , (C, A) , (C, B) , (C, D) , (D, D) , (D, A) , (D, B) e (D, C) . Como podemos observar, a ordem dos elementos, nesse caso, importa, pois na amostragem com reposição as amostras são ordenadas.

Contudo, a amostragem aleatória simples constitui um processo de seleção de uma amostra, realizado normalmente por sorteio, onde todos os elementos da população tenham a mesma probabilidade de serem escolhidos e que todo subconjunto de n elementos tenha a mesma chance de fazer parte da amostra. Constitui o método probabilístico de amostragem que permite a generalização, ou seja, garante cientificamente a aplicação de técnicas estatísticas de inferência.

3.4.4. Distribuição de frequências

Ao concluir a coleta de dados, o pesquisador tem à sua disposição um conjunto de observações que deverão ser organizadas e apresentadas de forma que facilite a análise

pretendida. As técnicas que nos permitem organizar, resumir e apresentar os dados estatísticos são do âmbito da Estatística Descritiva que não serão apresentadas nesta pesquisa. Entretanto, exibimos a distribuição de frequências que é a forma padrão de apresentação tabular de dados em Estatística.

O estudo das distribuições de frequências é dividido em dois casos distintos: dados tabulados e agrupados por frequências sem intervalos de classes, e dados tabulados e agrupados por frequências com intervalos de classes. A distribuição de frequências é o agrupamento de dados em classes associadas com suas respectivas medidas de frequências absolutas n_i e medidas de frequências relativas f_i de ocorrências. As medidas de frequências absolutas n_i representam o número de casos observados que foram enquadrados em cada classe e que são provenientes da contagem das variáveis da pesquisa para as quais a mensuração da característica em estudo resultou em valores que pertencem a classe.

As medidas de frequências relativas f_i são os valores que representam as frequências absolutas em termos de proporção em relação ao número total de variáveis. Dessa forma, temos $f_i = \frac{n_i}{N}$, onde N é o somatório de todas as frequências absolutas, enquanto o somatório de todas as frequências relativas sempre resultará no número um, uma vez que estamos tratando de proporções $\sum_{i=1}^{i=k} f_i = 1$.

Os dados agrupados, sem intervalos de classes, estão associados com a frequência absoluta e com a frequência relativa. Este tipo de distribuição de frequência é aplicado para as variáveis aleatórias discretas, quando a variedade de valores assumidos não for muito grande.

Tabela 3 - Distribuição de frequências sem intervalos de classes

Valores das variáveis	Frequência absoluta n_i	Frequência relativa f_i
x_1	n_1	f_1
x_2	n_2	f_2
...
x_k	n_k	f_k
Somatório	$N = \sum_{i=1}^{i=k} n_i$	$\sum_{i=1}^{i=k} f_i = 1$

Fonte: Do autor (2018).

Os dados agrupados com intervalos de classes, estão associados com a frequência absoluta e com a frequência relativa. Este tipo de distribuição de frequências é aplicado, principalmente, para as variáveis aleatórias contínuas.

Tabela 4 - Distribuição de frequências com intervalos de classes

Valores das variáveis	Centros dos intervalos	Frequência absoluta n_i	Frequência relativa f_i
$[x_1, x_2[$	$c_1 = \frac{x_1 + x_2}{2}$	n_1	f_1
$[x_2, x_3[$	$c_2 = \frac{x_2 + x_3}{2}$	n_2	f_2
...
$[x_{k-1}, x_k[$	$c_k = \frac{x_{k-1} + x_k}{2}$	n_k	f_k
Somatório		$N = \sum_{i=1}^{i=k} n_i$	$\sum_{i=1}^{i=k} f_i = 1$

Fonte: Do autor (2018).

A Tabela apresenta as classes por meio de intervalos da forma $[x_{k-1}, x_k[$, que denota o intervalo fechado à esquerda e aberto à direita. Nesse sentido, x_{k-1} é o limite inferior da classe e pertence a esse intervalo. A variável x_k é o limite superior da classe e não pertence a esse intervalo, pertencendo, então, ao intervalo seguinte. Quando os dados estiverem agrupados com intervalos de classe, Freund (2006), esclarece que se trata de uma variável contínua e, por convenção, aceitamos que as frequências estejam distribuídas uniformemente ao longo da classe. Dessa forma, o ponto médio da classe $c_k = \frac{x_{k-1} + x_k}{2}$ é o valor representativo da classe.

3.4.5. Média

A Estatística geralmente utiliza um único número para resumir um grupo de números. Intuitivamente, a melhor aposta seria encontrar um número que esteja centrado nas observações que estão sendo estudadas. Lethielleux e Chevalier (2016) destacam que esse número é proveniente de um grupo que a Estatística chama de medidas de tendência central. Andrade (2013) destaca em sua tese que entre as medidas de tendência central estão a média, a moda e a mediana, onde, para a nossa pesquisa, destacamos o uso da média aritmética. Entretanto, não devemos esquecer os conceitos de média geométrica e de média harmônica.

Matematicamente, a média é definida como a soma dos valores dividida pelo número de valores somados. Usando o conceito físico de ponto de equilíbrio, a média indica o centro de um conjunto de valores. A média de uma variável aleatória X , é também chamada de valor esperado de X , cuja representação é $E(X)$. Dessa forma, podemos escrever que a média aritmética dos valores de $X(X_1, X_2, \dots, X_n)$ é $E(X) = \frac{1}{N} \sum_{i=1}^N X_i$, ou seja, a soma de todos os valores dos elementos de X , dividido pela número total desses elementos.

A seguir, expomos as expressões comumente utilizadas para representar a média aritmética, por meio de somatório, apresentando uma diferenciação na notação da escrita entre uma média populacional e uma média amostral:

$$\text{Para identificar a média populacional: } \mu = \frac{1}{N} \sum_{i=1}^{i=N} X_i$$

$$\text{Para identificar a média amostral: } \bar{X} = \frac{1}{n} \sum_{i=1}^{i=n} X_i$$

Stevenson (2001) ressalta a importância de distinguirmos, em termos de notação, a média de uma amostra e a média de uma população. A letra grega μ (pronuncia-se mi) representa a média populacional e \bar{X} (x barra) a média amostral, onde X_i são os elementos da população ou da amostra, N é o número total de elementos de uma população e n o número total de elementos de uma amostra.

A definição de média envolve o conceito de somatório, dessa forma, aplicando-se as propriedades à formulação da média, podemos derivar algumas propriedades que serão úteis em algumas situações, permitindo-nos adiantar resultados sem a necessidade de efetuarmos extensos cálculos.

Propriedade 1: Seja k uma constante real, onde $E(k) = k$.

$$\text{Demonstração: } E(k) = \frac{1}{N} \sum_{i=1}^{i=N} k = \frac{k}{N} \sum_{i=1}^{i=N} 1 = \frac{k \cdot N}{N} = k$$

Para exemplificarmos essa primeira propriedade, vamos supor uma amostra (5,5,5,5). A média encontrada usando os elementos da amostra é $E(X) = \frac{1}{3} \sum_{i=1}^{i=4} X_i$, dessa forma, temos $\frac{5+5+5+5}{4} = 5$.

Propriedade 2: Seja k uma constante real, onde $E(kX) = kE(X)$.

$$\text{Demonstração: } E(kX) = \frac{1}{N} \sum_{i=1}^{i=N} (kX_i) = k \left(\frac{1}{N} \sum_{i=1}^{i=N} X_i \right) = kE(X)$$

De acordo com a segunda propriedade, para uma amostra (1, 2, 3, 4), a média é $E(X) = \frac{1+2+3+4}{4} = 2,5$. Considerando $k = 5$, obtemos um novo conjunto para a amostra (5, 10, 15, 20), onde teremos a média $\frac{5+10+15+20}{4} = 12,5$. Dessa forma, nota-se que $E(5X) = 5E(X)$.

Propriedade 3: Seja k uma constante real, onde $E(X + k) = E(X) + k$.

Demonstração: $E(X + k) = \frac{1}{N} \sum_{i=1}^{i=N} (X_i + k) = \frac{1}{N} (\sum_{i=1}^{i=N} X_i + \sum_{i=1}^{i=N} k)$, dessa forma, podemos escrever $\frac{1}{N} \sum_{i=1}^{i=N} X_i + \frac{1}{N} k \sum_{i=1}^{i=N} 1 = E(X) + \frac{kN}{N} = E(X) + k$.

Tomamos, por exemplo, uma amostra (1, 2, 3, 4), onde temos como média $E(X) = \frac{1+2+3+4}{4} = 2,5$. Vamos considerar um $k = 5$, dessa forma, teremos um novo conjunto (6, 7, 8, 9) cuja média é $E(X) = \frac{6+7+8+9}{4} = 7,5$. Observamos, então, que $7,5 = 2,5 + 5$.

Propriedade 4: para descrever a quarta propriedade vamos considerar o estudo de duas variáveis aleatórias X e Y , de modo que o Quadro expõe seus elementos.

Quadro 1 - Dados de X e de Y para a propriedade 4 da média

Variável	Elemento			
	E_1	E_2	...	E_N
X	X_1	X_2	...	X_N
Y	Y_1	Y_2	...	Y_N
$X \mp Y$	$X_1 \mp Y_1$	$X_2 \mp Y_2$...	$X_N \mp Y_N$

Fonte: Do autor (2018).

Neste caso, temos, $E(X \pm Y) = E(X) \pm E(Y)$. Morettin & Bussab (2010) descrevem que, por indução essa propriedade pode ser estendida para um número finito de variáveis aleatórias.

Demonstração: $E(X \pm Y) = \frac{1}{N} \sum_{i=1}^{i=N} (X_i \pm Y_i) = \frac{1}{N} (\sum_{i=1}^{i=N} X_i \pm \sum_{i=1}^{i=N} Y_i) = \frac{1}{N} \sum_{i=1}^{i=N} X_i \pm \frac{1}{N} \sum_{i=1}^{i=N} Y_i = E(X) \pm E(Y)$. Tomamos o Quadro para exemplificar a propriedade 4.

Quadro 2 - Dados de X e de Y para o exemplo da propriedade 4 da média

Variável	Elemento			
	E_1	E_2	E_3	E_4
X	1	2	3	4
Y	5	1	6	4
X + Y	6	3	9	8

Fonte: Do autor (2018).

A média da variável X é $E(X) = \frac{1+2+3+4}{4} = 2,5$ e a média da variável Y é $E(Y) = \frac{5+1+6+4}{4} = 4$. Dessa forma, temos que $E(X) + E(Y) = 6,5$. A variável $(X + Y)$, possui média $E(X + Y) = \frac{6+3+9+8}{4} = 6,5$. Assim sendo, $E(X + Y) = 6,5$ que é o mesmo resultado de $E(X) + E(Y)$.

Propriedade 5: ao enunciar a quinta propriedade, vamos observar os dois exemplos a seguir.

Exemplo 1: Sendo as variáveis $X = (1, 2, 1, 5, 1)$, $Y = (2, 1, 1, 2, 3)$ e $X.Y = (2, 2, 1, 10, 3)$. Temos $E(X) = 2$, $E(Y) = 1,8$ e $E(X.Y) = 3,6$. Neste exemplo observamos que $E(X.Y) = E(X).E(Y)$.

Exemplo 2: Sendo as variáveis $X = (2, 3, 1, 1, 5)$, $Y = (1, 2, 3, 4, 5)$ e $X.Y = (2, 6, 3, 4, 25)$. Temos $E(X) = 2,4$, $E(Y) = 3$ e $E(X.Y) = 8$. Neste exemplo observa-se que $E(X.Y) \neq E(X).E(Y)$.

Neste sentido, diferentemente do que acontece com a média da soma ou da diferença de duas variáveis, vistas na quarta propriedade, a média do produto nem sempre coincide com o produto das médias.

Para que a propriedade $E(X.Y) = E(X).E(Y)$ seja verdadeira, é necessário que $\sum_{i=1}^{i=N} X_i.Y_i = \frac{\sum_{i=1}^{i=N} X_i \sum_{i=1}^{i=N} Y_i}{N}$. De fato, se $E(X.Y) = \frac{1}{N} \sum_{i=1}^{i=N} X_i.Y_i = \frac{\sum_{i=1}^{i=N} X_i \sum_{i=1}^{i=N} Y_i}{N} = \frac{1}{N} \sum_{i=1}^{i=N} X_i \cdot \frac{1}{N} \sum_{i=1}^{i=N} Y_i = E(X).E(Y)$.

Morettin & Bussab (2010) salientam que o cálculo da média é uma ótima maneira de resumir um grupo de números, mas, geralmente, não fornece todas as informações tipicamente necessárias para análise de uma pesquisa. Para evitar uma possível perda de informações, Ardilly (2004) recomenda, o uso de uma outra estatística que mede a variação dos elementos. Dessa forma, destaca a variância e o desvio padrão que fazem parte do grupo que a estatística chama de medidas de dispersão.

3.4.6. Variância

Vamos supor que estejamos interessados em apurar algumas características de dois grupos cujos elementos foram apresentados da seguinte forma: Grupo A (52,52,52,52) e Grupo B (54, 48,56,50). Ao calcularmos a média do Grupo A e a média do Grupo B encontraremos, para ambos, o resultado 52. Sob a perspectiva da média, os grupos parecem ser similares. Contudo, é possível perceber que existe diferença entre os dois grupos e, para que essa diferença não resulte em perda de informações, Ardilly (2004) sugere a avaliação da média dos desvios de cada grupo. A média dos desvios de um conjunto numérico é composta pela soma de todas as diferenças entre cada um dos elementos que compõem o conjunto com o valor esperado do conjunto dividido pelo total de elementos. Dessa forma, o desvio de um elemento X_i é $X_i - \bar{X}$ para uma amostra ou $X_i - \mu$ para uma população.

Tabela 5 - Desvios dos Grupos A e B

GRUPO A			GRUPO B		
X_i	$X_i - E(A)$	Desvio	X_i	$X_i - E(B)$	Desvio
52	52 - 52	0	48	48 - 52	-4
52	52 - 52	0	50	50 - 52	-2
52	52 - 52	0	54	54 - 52	2
52	52 - 52	0	56	56 - 52	4

Fonte: Do autor (2018).

De acordo com a Tabela 5, percebemos que a soma dos desvios de ambos os grupos é zero, de onde, também, resulta uma média nula. Morettin e Bussab (2010) indicam que apenas encontrar a média dos desvios não nos ajuda a observar a diferença entre dois grupos, pois a média dos desvios de qualquer grupo sempre será zero, o que também consideram como uma propriedade que define a média. Dessa forma, a Estatística utiliza o desvio quadrado que é a multiplicação do desvio por ele mesmo. A média dos desvios quadrados é chamada de variância. A variância de uma variável aleatória X é uma medida de dispersão dos valores da variável em torno da sua média. Ela é definida como a média aritmética dos quadrados dos desvios dos valores de $X(X_1, X_2, \dots, X_N)$ em relação a média dos próprios valores de X . Nesse sentido, em uma população de tamanho N , a variância de uma variável aleatória X , denotada por $Var(X)$ é dada por $Var(X) = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)^2$, onde μ é a média populacional da variável X .

A fórmula da variância, $Var(X) = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)^2$, pode sofrer um tratamento algébrico e poderá ser escrita da seguinte forma: $Var(X) = \frac{1}{N} \sum_{i=1}^N X_i^2 - \left(\frac{1}{N} \sum_{i=1}^N X_i \right)^2$.

Podemos demonstrar esse fato fazendo:

$$\begin{aligned} Var(X) &= E(X - \mu)^2 = E(X^2 - 2X\mu + \mu^2) = E(X^2) - E(2X\mu) - [E(\mu)]^2 = \\ &= E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - [E(X)]^2. \end{aligned}$$

Logo, temos que $Var(X) = E(X^2) - [E(X)]^2 = \frac{1}{N} \sum_{i=1}^N X_i^2 - \left(\frac{1}{N} \sum_{i=1}^N X_i \right)^2$, ou seja, a variância de uma variável aleatória X , que assume os valores (X_1, X_2, \dots, X_N) , também pode ser obtida calculando-se a média dos quadrados dos valores da variável menos o quadrado da média de tais valores.

Morettin & Bussab (2010), discutem que a primeira expressão, $Var(X) = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)^2$ modela e exprime a própria definição de variância e, portanto, poderá ajudar o aluno a compreender seu significado.

A expressão $Var(X) = \frac{1}{N} \sum_{i=1}^N X_i^2 - \left(\frac{1}{N} \sum_{i=1}^N X_i \right)^2$ não exhibe a definição de variância, pois não é o modelo explícito de seu conceito, mas garante uma simplificação nos cálculos. Importante expormos algumas propriedades da variância que serão úteis para a compreensão da Estimativa Estatística, conforme segue:

Propriedade 1: considerando k uma constante real, temos que $Var(k) = 0$.

Demonstração: temos que $Var(k) = E(k^2) - [E(k)]^2 = k^2 - k^2 = 0$.

Para exemplificarmos, tomamos os valores $(1, 1, 1, 1)$ cuja média é 1, e a $Var(X) = \frac{1^2+1^1+1^2+1^2}{4} - 1^2 = 0$.

Propriedade 2: considerando k uma constante real, expõe que $Var(kX) = k^2 \cdot Var(X)$.

$$\begin{aligned} \text{Demonstração: } Var(kX) &= E[kX^2] - [E(kX)]^2 = E(k^2 \cdot X^2) - [kE(X)]^2 = \\ &= k^2 E(X^2) - k^2 [E(X)]^2 = k^2 \{E(X^2) - [E(X)]^2\} = k^2 Var(X). \end{aligned}$$

Elucidamos essa propriedade considerando os valores $(2, 3, 4, 5)$, onde a média é 3,5 e a $Var(X) = \frac{54}{4} - 3,5^2 = 1,25$. Considerando $k = 2$, temos um novo conjunto de valores $(4, 6, 8, 10)$, cuja média é 7 e a $Var(X) = 5$. Dessa forma, observamos que $5 = k^2 \cdot Var(X) = 2^2 \cdot 1,25$.

Propriedade 3: uma terceira propriedade informa que $Var(X + k) = Var(X)$, onde k é uma constante real.

Demonstração: temos que $Var(X + k) = E[(X + k)^2] - [E(X + k)]^2 = E(X^2 + 2kX + k^2) - [E(X) + k]^2 = [E(X^2) + 2kE(X) + E(k^2)] - \{[E(X)]^2 + 2kE(X) + k^2\} = E(X^2) - [E(X)]^2 = Var(X)$.

Por exemplo, para os valores (1, 2, 3, 4), a média é 2,5 e a $Var(X) = \frac{30}{4} - 2,5^2 = 1,25$. Considerando $k = 3$, temos um novo conjunto (4, 5, 6, 7) onde a média é 5,5 e a $Var(X) = \frac{126}{4} - 5,5^2 = 1,25$, que é, exatamente, a mesma variância do conjunto inicial. Assim, $Var(X + k) = Var(X)$.

A variância, segundo Morettin e Bussab (2010), é uma medida de dispersão que permite identificar o quão distante da média está cada elemento de uma população ou de uma amostra. A variação será grande se os números de um grupo variarem muito com relação à média, e baixa se os números forem similares a média. Assim como a média, as expressões que representam a variância populacional e a variância amostral são distintas, como exibem os quadros a seguir.

Quadro 3 - Variância Populacional

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{i=N} (X_i - \mu)^2$$

σ^2 = variância da população
 N = tamanho da população
 X_i = elemento da população
 μ = média aritmética da população
 $(X_i - \mu)$ = dispersão ou desvio de cada elemento em relação a média

Fonte: Do autor (2018).

Quadro 4 - Variância amostral alterada

$$s^2 = \frac{1}{n - 1} \sum_{i=1}^{i=n} (X_i - \bar{X})^2$$

s^2 = variância da amostra
 n = tamanho da amostra
 X_i = elemento da amostra
 \bar{X} = média aritmética da amostra
 $(X_i - \bar{X})$ = dispersão ou desvio de cada elemento em relação a média

Fonte: Do autor (2018).

Usamos a letra grega σ (sigma) para designar a variância populacional e a letra S para a variância amostral. Como em ambos os casos estamos tratando de quantidades quadradas, utilizamos σ^2 e S^2 . É possível observar que além dessa notação, existe uma diferença entre as expressões da variância populacional e da variância amostral no denominador. Morettin e Bussab (2010) ratificam que quando utilizamos a variância de uma amostra usando $(n - 1)$, o resultado obtido é não tendencioso e com menor erro médio, isso em razão de usarmos a média da amostra \bar{X} como uma estimativa da média da população, a qual não conhecemos. Faremos a demonstração desse fato quando tratarmos sobre as propriedades dos estimadores.

Voltando ao nosso exemplo inicial e supondo estar trabalhando com uma amostra, o Grupo B teria uma variância de, aproximadamente, $S^2 = \frac{(54-52)^2+(48-52)^2+(56-52)^2+(50-52)^2}{4-1} = 13,33$. Esse valor é bem diferente do encontrado no Grupo A, cuja variação é nula.

3.4.7. Desvio padrão

Após calcular a variância em um conjunto de números, obtemos um valor que é medido na unidade quadrada do dado original. Morettin e Bussab (2010) indicam que isso pode trazer uma camuflagem dos valores, pois aumenta a divergência das observações em relação a média. No entanto, como apontam Lethielleux e Chevalier (2016), a estatística postula o uso do desvio padrão, cuja diferença para a variância é a raiz quadrada.

Quadro 5 - Desvio Padrão populacional e amostral

Desvio padrão populacional	Desvio padrão amostral
$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{i=N} (X_i - \mu)^2}{N}}$	$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^{i=n} (X_i - \bar{X})^2}{n - 1}}$

Fonte: Do autor (2018).

Dehon, Droesbeke e Vermandele (2008) anunciam o desvio padrão como uma medida de dispersão que é expressa na mesma unidade de medida da variável em estudo, e por isso mesmo que se enquadra como uma medida absoluta. O desvio padrão é a medida de dispersão que faz relação junto com a variância e com a média aritmética. Um valor baixo para o desvio padrão indica que as observações de um conjunto tendem a estar próximos da média e, um valor alto, indica que as observações estão espalhadas por uma ampla gama de valores.

3.4.8. Cálculo das medidas estatísticas a partir de uma distribuição de frequências

Vamos apresentar o cálculo das medidas estatísticas analisadas (média, variância e desvio padrão) quando os dados forem expostos por meio de tabelas de frequências. A distribuição de frequências mostra a disposição dos valores que uma ou mais variáveis tomam em uma amostra. Dessa forma, cada entrada na tabela contém a frequência de ocorrências de valores absolutos n_i ou relativos f_i dentro de um intervalo, resumindo a distribuição dos valores da amostra. A Tabela 6 resume as fórmulas para as medidas estatísticas analisadas quando os dados estiverem agrupados em uma distribuição de frequências.

Tabela 6 - Medidas estatísticas para distribuição de frequências

	Média	Variância	Desvio Padrão
População	$\mu = \frac{1}{N} \sum_{i=1}^{i=N} n_i X_i$	$\sigma^2 = \frac{1}{N} \sum_{i=1}^{i=N} n_i (X_i - \mu)^2$	$\sigma = \sqrt{\sigma^2}$
Amostra	$\bar{X} = \frac{1}{n} \sum_{i=1}^{i=n} n_i X_i$	$S^2 = \frac{1}{n-1} \sum_{i=1}^{i=n} n_i (X_i - \bar{X})^2$	$S = \sqrt{S^2}$

Fonte: Do autor (2018).

Uma tabela com distribuição de frequências, pode apresentar os dados agrupados com, ou sem, intervalos de classes. A Figura 31 apresenta um exemplo de dados agrupados sem intervalos de classes, onde a tabela exposta foi construída para atender os cálculos das medidas estatísticas em relação a uma amostra. Expõe, também, os cálculos da média amostral, variância amostral e desvio padrão amostral.

Figura 31 - Dados agrupados sem intervalo de classes

X_i	n_i	$n_i X_i$	$n_i (X_i - \bar{X})^2$
2	4	8	9
3	1	3	0,25
4	3	12	0,75
6	2	12	12,5
TOTAL	10	35	22,5

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{i=n} n_i f_i = \frac{35}{10} = 3,5$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{i=N} n_i (X_i - \bar{X})^2 = \frac{22,5}{10-1} = 2,5$$

$$S = \sqrt{2,5} = 1,5811388301$$

Fonte: Do autor (2018).

A Figura apresenta um exemplo de dados agrupados com intervalos de classes, onde a tabela exposta foi construída para atender os cálculos das medidas estatísticas em relação a uma

amostra. Expõe, também, os cálculos da média amostral, variância amostral e desvio padrão amostral. Quando os dados estiverem agrupados com intervalos de classes, X_i assumem o valor representativo do intervalo como vimos anteriormente no texto que tratou de distribuição de frequências.

Figura 32 - Dados agrupados com intervalos de classes

Intervalo	X_i	n_i	$n_i X_i$	$n_i (X_i - \bar{X})^2$
[55,65[60	10	600	1638,4
[65,75[70	3	210	23,52
[75,85[80	7	560	362,88
[85,95[90	5	450	1479,2
TOTAL	---	25	1820	3504

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{i=n} n_i X_i = \frac{1820}{25} = 72,8 \text{ média}$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n n_i (X_i - \bar{X})^2 = \frac{3504}{25-1} = 146$$

$$S = \sqrt{146} = 12,0830459736$$

Fonte: Do autor (2018).

As medidas de tendência central e medidas de dispersão são conceitos fundamentais no estudo da Estatística, abordadas no âmbito da Estatística Descritiva, e possuem uma fundamental relevância no estudo das inferências estatísticas. A média, a variância e o desvio padrão são tratadas na inferência estatística como estimadores que determinam as estimativas de parâmetros populacionais por meio de amostras.

3.4.9. Parâmetros, estimadores e estimativas

Para compreendermos o processo de Estimação Estatística é importante definir o que são parâmetros, estimadores e estimativas. Esses três conceitos arquitetam a estrutura básica da estimação. Os parâmetros estão associados com a população e são descritos por meio de uma medida descritiva como a média, a variância, desvio padrão e a proporção. Nesse sentido, podemos definir um parâmetro como uma grandeza fixa referente a uma população. Na maior parte dos casos, o valor desse parâmetro é desconhecido, pois dificilmente podemos medir toda uma população, e por desconhecermos o valor desse parâmetro, resta tentar estimá-lo.

Para efetivar uma estimação, utilizamos um estimador que se refere a uma fórmula matemática aplicada aos dados de uma amostra. Essa amostra, como já sabemos, advém de um subconjunto representativo retirado de uma população. Os estimadores estão associados a essa amostra e são descritos por meio de uma medida descritiva como a média, a variância, desvio padrão e a proporção.

As estimativas, ou as estatísticas, como descrevem Morettin e Bussab (2010), são os valores numéricos que os estimadores assumem para uma dada amostra específica e, somente

para essa amostra, ou seja, uma estimativa varia de amostra para amostra. Para cada amostra distinta teremos, portanto, diferentes estimativas que são calculadas por meio de uma mesma fórmula que é o estimador.

Lethielleux e Chevalier (2016) ressaltam a presença de notações, para fins de generalizações, em relação a distinção de parâmetros e estimadores. Usa-se a letra grega teta minúsculo θ para a representação de parâmetros e $\hat{\theta}$ para os estimadores. Todavia, os principais parâmetros e estimadores, como relata Ardilly (2004), recebem uma nomenclatura própria que são, justamente, as medidas estatísticas anteriormente visitadas, como apresenta a Tabela 7 abaixo.

Tabela 7 - Principais parâmetros e seus respectivos estimadores

PARÂMETROS	ESTIMADORES
MÉDIA POPULACIONAL: $\mu = \frac{\sum_{i=1}^N X_i}{N}$	MÉDIA AMOTRAL: $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$
VARIÂNCIA POPULACIONAL: $\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$	VARIÂNCIA AMOTRAL: $S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$
DESVIO PADRÃO POPULACIONAL: $\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}}$	DESVIO PADRÃO AMOTRAL: $S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$

Fonte: Do autor (2018).

Temos, ainda, o parâmetro da proporção populacional p , frequência relativa observada na população, e o estimador \hat{p} , frequência relativa observada na amostra.

Figura 33 - Processo da estimação



Fonte: Do autor (2018).

Contudo, podemos entender que a estimação estatística é um processo que consiste na utilização de dados amostrais para estimar os valores de parâmetros populacionais desconhecidos. Qualquer característica de uma população pode ser estimada a partir de uma amostra aleatória. De acordo com a estimativa encontrada, podemos estabelecer uma extrapolação dos dados desconhecidos da população investigada e inferir em conformidade com a proposta investigada.

3.4.10. Modelos de distribuição de probabilidades

Os itens anteriores apresentaram temas envolvendo os tipos de séries de dados, provenientes de amostras e populações, distribuição de frequências e as medidas estatísticas. Todos esses conhecimentos permitem uma análise nas séries de dados para a obtenção de algumas conclusões sobre como tais dados se distribuem em todo seu intervalo de variação ou ao redor de valores centrais. Em síntese, partimos do princípio de que a partir da organização, apresentação e descrição dos dados observados, é possível fazer conjecturas sobre o comportamento da variável em estudo. Esse tipo de raciocínio é denominado de indução, onde, a partir de resultados ou dados observados, lançamos hipóteses sobre o comportamento do fenômeno.

A partir de agora, estamos interessados no raciocínio de forma inversa, ou seja, compreender como poderão ocorrer os resultados de uma variável, a partir de suposições sobre o problema em estudo, caracterizando, assim, o que denominamos de raciocínio dedutivo, com base em hipóteses e conjecturas sobre o comportamento de um fenômeno, tentando prever os resultados. Em suma, estamos interessados em realizar inferências sobre a população de onde foi extraída a amostra. Para tanto, teremos que usar modelos matemático-probabilísticos sobre os quais a estimação estatística adequa-se. Trataremos das características básicas de duas distribuições de probabilidades, quais sejam, a distribuição binomial e a distribuição normal.

3.4.10.1. Distribuição binomial

A distribuição Binomial é uma distribuição de probabilidade discreta do número de sucessos numa sequência de n tentativas. Ela fornece embasamento para inferências sobre proporções. Morettin e Bussab (2010) descrevem que, as tentativas, além de independentes, possuem exclusivamente dois resultados que são, a probabilidade de sucesso p ou a probabilidade de fracasso $q = 1 - p$, a que se chama de tentativa de Bernoulli. A probabilidade

de sucesso, a cada tentativa, permanece constante independente das demais, e a variável de interesse ou pretendida, é o número de sucessos k nas n tentativas.

Seja X o número de sucessos obtidos na realização de n ensaios de Bernoulli independentes, denotamos que X tem distribuição Binomial com parâmetros n e p , da seguinte forma: $X \sim B(n, p)$, em que p é a probabilidade de sucesso em cada ensaio. O Quadro 6 apresenta a expressão geral da distribuição Binomial.

Quadro 6 - Expressão geral da distribuição Binomial

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Fonte: Do autor (2018).

De acordo com a expressão, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ é a combinação que mostra as possíveis ordenações, onde p^k indica o número de sucesso, e $(1 - p)^{n-k}$ a probabilidade de fracassos. Seja X uma variável aleatória com distribuição Binomial (n, p) , as principais características são:

$E(X) = \sum_{k=0}^{k=n} kP(X = k) = \sum_{k=0}^{k=n} k \binom{n}{k} p^k (1 - p)^{n-k} = np$, o valor esperado representa o número médio de sucessos e a variância de X , $Var(X)$ é dada por

$$Var(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2 = np(1 - p).$$

Quadro 7 - Principais características da distribuição Binomial

$$\begin{aligned} \text{Média } \mu_X &= E(X) = np \\ \text{Variância } \sigma_X^2 &= Var(X) = np(1 - p) \end{aligned}$$

Fonte: Do autor (2018).

A distribuição binomial pode ser aproximada pela distribuição Normal que será analisada a seguir.

3.4.10.2. Distribuição normal

A distribuição normal, também conhecida como distribuição de Gauss, Laplace ou Laplace-Gauss, em referência aos matemáticos, físicos e astrônomos franceses Pierre-Simon Laplace (1749-1827) e o alemão Carl Friedrich Gauss (1777-1855), é considerada a distribuição de probabilidades mais importante, permitindo modelar numerosos fenômenos naturais. Também possibilita a realização de aproximações para o cálculo de probabilidades de muitas variáveis aleatórias que possuem outras distribuições, tais como a binomial, e oferece embasamento para as inferências da média \bar{X} e da proporção \hat{p} amostrais sobre a média μ e proporção p populacionais, por exemplo.

Ao citarmos a distribuição normal, estamos nos referindo a uma família de distribuições indexadas pelos parâmetros μ e σ . Dessa forma, para cada par de valores destes parâmetros, temos uma distribuição normal cuja expressão analítica está representada no Quadro 8.

Quadro 8 - Função densidade de probabilidade da distribuição Normal

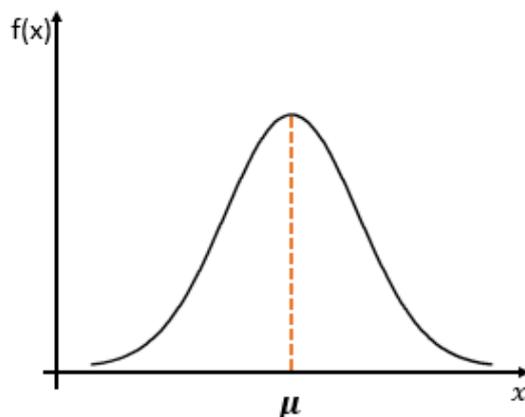
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$\mu = \text{média populacional}$
 $\sigma = \text{desvio padrão populacional}$
 Constantes $e = 2,718 \dots$ e $\pi = 3,1416 \dots$

Fonte: Do autor (2018).

O Gráfico 1, que representa uma distribuição normal, descreve uma curva em forma de sino que evidencia maior probabilidade de uma variável aleatória assumir valores próximos aos valores centrais.

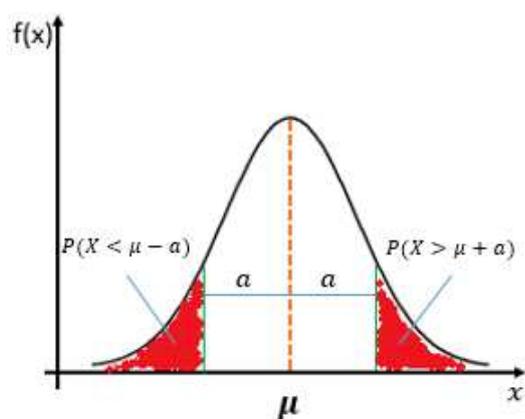
Gráfico 1 - Distribuição normal



Fonte: Do autor (2018).

Teoricamente, a curva prolonga-se de $-\infty$ a $+\infty$, sendo que $\lim_{x \rightarrow \pm\infty} f(x) = 0$ e a área total, sob a curva, é igual a 1, ou seja $\int_{-\infty}^{+\infty} f(x)dx = 1$. A curva é simétrica em torno de μ , onde $P(X < \mu - a) = P(X > \mu + a)$.

Gráfico 2 - Simetria da Distribuição Normal



Fonte: Do autor (2018).

Uma variável aleatória X com distribuição normal de parâmetros μ e σ é representada por $X \sim N(\mu, \sigma)$, onde $E(X) = \mu$ e $Var(X) = \sigma^2$. Considerando a enorme dificuldade de calcular probabilidades pela integração da expressão analítica da distribuição normal para as infinitas combinações de μ e σ , utilizamos a distribuição normal padrão.

A distribuição normal padrão é uma técnica que permite encontrar a área abaixo da curva para qualquer distribuição normal comparando-a com a distribuição normal cuja média seja zero e variância igual a 1, denotada por $N(0,1)$. A conveniência de termos a média igual a zero e desvio padrão igual a um, faz com que esta distribuição passe a representar os valores de Z

como números de desvios em relação à média. Assim, esta distribuição nos permite trabalhar com valores relativos de desvios em relação à média. Qualquer distribuição normal $N(\mu, \sigma^2)$ com média μ e desvio padrão σ pode ser transformada, para efeito de cálculo de probabilidades, na distribuição normal padrão $N(0,1)$, por meio de uma mudança de variável e utilizando a tabela da Curva Normal Reduzida (Anexo C).

Quadro 9 - Transformação de $N(\mu, \sigma^2)$ para $N(0,1)$

$$Z = \frac{X - \mu}{\sigma}$$

Fonte: Do autor (2018).

Um importante teorema diz que se uma variável X tem distribuição normal, uma função linear de X será também normal. Podemos demonstrar da seguinte forma:

Seja X uma variável com distribuição normal e $aX + b$ uma função linear de X com a e b constantes. Se encontrarmos a e b tais que a média de $aX + b$ seja zero e sua variância unitária, teremos uma variável normal padrão. Consideramos um sistema de equações dado por:

$$E(aX + b) = 0 \text{ e } Var(aX + b) = 1.$$

$$\begin{aligned} \text{Sabendo que } Var(aX + b) &= E[(aX + b)^2 - (E[aX + b])^2] = \\ &= E[a^2X^2 + 2abX + b^2] - (aE(X) + b)^2 = \\ &= a^2E(X^2) + 2abE(X) + b^2 - a^2E^2(X) - 2abE(X) - b^2 = \\ &= a^2E(X^2) - a^2E^2(X) = a^2. \end{aligned}$$

Logo, temos que $a\mu + b = 0$ e $a^2\sigma^2 = 1$. Resolvendo esse sistema, encontramos $a = \frac{1}{\sigma}$ e $b = -\frac{\mu}{\sigma}$, onde teremos $aX + b = \frac{(X-\mu)}{\sigma}$. Como $\frac{(X-\mu)}{\sigma}$ tem média zero e variância igual a um, é uma variável normal padrão, isto é, $\frac{(X-\mu)}{\sigma} = Z$. Assim, qualquer variável normal com média μ e variância σ^2 pode ser transformada em uma variável normal padrão.

Vamos considerar $P(x_1 < x < x_2)$, com $x_1 < x_2$. Seja $\frac{(X-\mu)}{\sigma} = Z$, implica que $X = \sigma Z + \mu$.

Portanto, $x_1 = \sigma z_1 + \mu$ e $x_2 = \sigma z_2 + \mu$. Realizando as substituições necessárias, encontramos:

$$P(x_1 < x < x_2) = P(\sigma z_1 + \mu < \sigma z + \mu < \sigma z_2 + \mu)$$

Cancelando os termos positivos comuns do segundo membro da equação temos:

$$P(x_1 < x < x_2) = P(z_1 < z < z_2)$$

Onde, $z_1 = \frac{(x_1 - \mu)}{\sigma}$ e $z_2 = \frac{(x_2 - \mu)}{\sigma}$. Desse modo, a probabilidade de x estar entre x_1 e x_2 é igual a probabilidade de a variável normal padrão estar entre $\frac{(x_1 - \mu)}{\sigma}$ e $\frac{(x_2 - \mu)}{\sigma}$.

3.4.11. Teorema Central Limite

O Teorema Central Limite é fundamental no estudo da inferência estatística pois trata da relação existente entre a média das amostras e a média da população. Quanto maior o tamanho da amostra de uma distribuição, mais as médias da população e da amostra se aproximam. Podemos definir o Teorema do Limite Central como: sejam n variáveis aleatórias $X_1, X_2, X_3, \dots, X_n$ independentes com a mesma distribuição de probabilidade, de média $E(X_i) = \mu$ e variância $V(X_i) = \sigma^2$. Faça-se $S_n = \sum_{i=1}^n X_i$.

Então, se n é suficientemente grande, S_n possui aproximadamente distribuição normal de média $E(S_n) = n\mu$ e variância $V(S_n) = n\sigma^2$. Nas mesmas condições, considere-se a média aritmética das n variáveis aleatórias $X_1, X_2, X_3, \dots, X_n$, temos $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Então, se n é suficientemente grande, \bar{X}_n possui distribuição aproximadamente normal de média $E(\bar{X}_n) = \mu$ e variância $V(\bar{X}_n) = \frac{1}{n} \sigma^2$.

De acordo com o teorema, a distribuição amostral de uma média é aproximadamente uma distribuição normal, caso o tamanho da amostra seja grande o suficiente. Morettin e Bussab (2010) revelam que grande o suficiente significa aproximadamente 30 ou mais observações. A média da distribuição amostral da média é igual a média da população $\mu_{\bar{X}} = E(\bar{X}_n) = \mu$ e, o desvio padrão da distribuição amostral da média, também conhecido como erro padrão da média, é igual ao desvio padrão da população dividido pela raiz quadrada do tamanho da amostra $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$.

3.4.12. Distribuições Amostrais

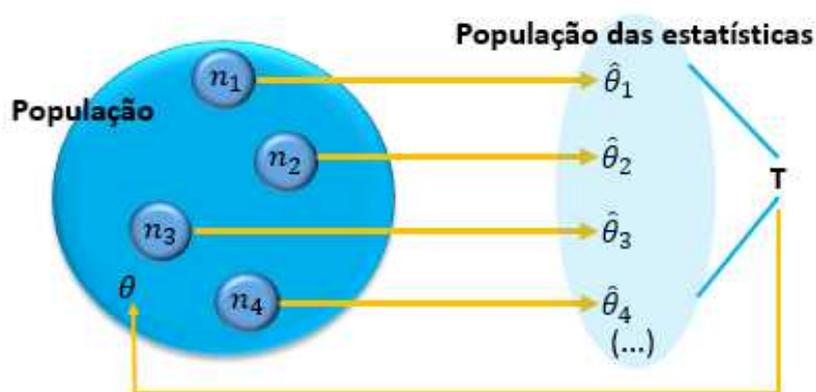
Os parâmetros populacionais θ possuem um valor constante, ou seja, o seu valor não é alterado de amostra para amostra. Contudo, o valor de um estimador $\hat{\theta}$ é dependente da amostra selecionada e, dessa forma, cada amostra poderá revelar uma estatística distinta.

Toda estatística, sendo uma função de uma amostra aleatória $(X_1, X_2, X_3, \dots, X_n)$, é, também, uma variável aleatória e possui uma distribuição. Essa distribuição é o comportamento aleatório dos elementos amostrais do fenômeno estudado. Muito embora existam ocorrências com apenas uma amostra e um único valor correspondente à estatística em relação a várias amostras, a estatística muda de valor segundo a distribuição determinada a partir daquela que controla a amostra aleatória.

Morettin e Bussab (2010) afirmam que nossa resposta seria melhor compreendida se conhecêssemos o que ocorre com as estatísticas quando retiramos todas as amostras de uma população conhecida segundo o plano amostral adotado. O comportamento da estatística pode ser descrito por alguma distribuição de probabilidade. Assim, cada estatística é uma variável aleatória e sua distribuição de probabilidade é chamada de distribuição amostral da estatística.

A Figura apresenta um esquema onde θ é o parâmetro de interesse na população e T é o valor das estatísticas para cada amostra n .

Figura 34 - Inferência sobre o parâmetro populacional θ por meio da distribuição



Fonte: Do autor (2018).

Diante disso, analisamos o comportamento da distribuição amostral da média e da proporção.

3.4.12.1. Distribuição amostral da média: com e sem reposição

Sabemos que a Estatística utiliza a média da amostra \bar{X} para realizar inferências sobre uma população de média μ . Esse procedimento está representado na Figura 35 abaixo.

Figura 35 - Distribuição amostral da média: com e sem reposição



Fonte: Do autor (2018).

Executando esse processo continuamente, iremos encontrar valores distintos para a estimativa de \bar{X} . Por consequência, podemos dizer que a própria estimativa da média de uma variável aleatória X , também é uma variável aleatória. Sendo assim, a distribuição de probabilidade da variável aleatória \bar{X} é chamada de distribuição da média da amostra, ou seja, é a distribuição de todos os valores possíveis da média da amostra. Analisamos o processo de uma distribuição amostral da média, por meio de exemplos, levando-se em consideração amostras com reposição e sem reposição para que possamos entender e conhecer alguns resultados relevantes para a Estatística.

Começamos com um exemplo de uma amostragem com reposição: vamos observar uma população $P = \{2,3,4,5\}$, com média populacional $\mu = 3,5$ e variância populacional $\sigma^2 = 1,25$. Consideramos todas as amostras possíveis de tamanho $n = 2$, ou seja, $k = N^n = 4^2 = 16$, onde k é o número de amostras extraídas com reposição e N o tamanho da população. Dessa forma, o Quadro 10 abaixo apresenta as possíveis amostras com as suas respectivas médias.

Quadro 10 - Amostras extraídas com reposição e suas respectivas médias

Amostras	(2,2)	(2,3)	(2,4)	(2,5)	(3,2)	(3,3)	(3,4)	(3,5)
<i>Média das amostras \bar{x}_i</i>	2	2,5	3	3,5	2,5	3	3,5	4
Amostras	(4,2)	(4,3)	(4,4)	(4,5)	(5,2)	(5,3)	(5,4)	(5,5)
<i>Média das amostras \bar{x}_i</i>	3	3,5	4	4,5	3,5	4	4,5	5

Fonte: Do autor (2018).

Com os dados observados, construímos uma tabela de distribuição de frequências para a distribuição amostral da média e com as informações necessárias para os cálculos da média e da variância das médias amostrais.

Tabela 8 - Distribuição amostral das médias extraídas das amostras com reposição

Média das amostras \bar{x}	n_i	f_i	$\bar{x}_i f_i$	$\bar{x}_i^2 f_i$
2	1	1/16	2/16	4/16
2,5	2	2/16	5/16	12,5/16
3	3	3/16	9/16	27/16
3,5	4	4/16	14/16	49/16
4	3	3/16	12/16	48/16
4,5	2	2/16	9/16	40,5/16
5	1	1/16	5/16	25/16
Total	16	1	56/16	206/16

Fonte: Do autor (2018).

Podemos observar que a probabilidade da média amostral \bar{X} é a média de alguma das amostras \bar{x}_i , indicado por $P(\bar{X} = \bar{x})$. Nesse sentido, quando a amostra selecionada for o par (2,2), a média será 2, então, teremos $P(\bar{X} = 2) = 1/16$. Obteremos a média igual a 4 quando ocorrer (3,5), (4,4) ou (5,3), então $P(\bar{X} = 4) = 3/16$.

De acordo com a Tabela, temos $E(\bar{X}) = \sum \bar{x}_i f_i = \frac{56}{16} = 3,5$. A $Var(\bar{X}) = \sigma_{\bar{X}}^2 = \sum \bar{x}_i^2 f_i - \mu_{\bar{x}}^2 = \frac{206}{16} - 3,5^2 = 0,625$.

Verificamos que a média de todas as médias amostrais extraídas com reposição da população P, é igual a média populacional μ . A variância entre as médias amostrais é "n" vezes menor, em nosso caso particular duas vezes menor que a variância populacional. Podemos obter, ainda, o erro padrão da média $\sigma_{\bar{x}} = 0,79$ que mede a variabilidade entre as médias amostrais e dá uma ideia do erro que se comete ao substituir a média da população pela média da amostra.

Verificamos agora um exemplo de amostragem sem reposição: vamos observar uma população $P = \{2,3,4,5\}$, com média populacional $\mu = 3,5$ e variância populacional $\sigma^2 = 1,25$. Consideramos todas as amostras possíveis de tamanho $n = 2$, ou seja, $k = \binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{4!}{2!(4-2)!} = 6$, onde k é o número de amostras extraídas sem reposição e N o tamanho da população. Dessa forma, o Quadro abaixo apresenta as possíveis amostras com as suas respectivas médias:

Quadro 11 - Amostras extraídas sem reposição e suas respectivas médias

Amostras	(2,3)	(2,4)	(2,5)	(3,4)	(3,5)	(4,5)
Média das amostras \bar{x}	2,5	3	3,5	3,5	4	4,5

Fonte: Do autor (2018).

Com os dados observados, construímos uma tabela de distribuição de frequências para a distribuição amostral da média e com as informações necessárias para os cálculos da média e da variância das médias amostrais.

Tabela 9 - Distribuição amostral das médias extraídas das amostras sem reposição

Média das amostras \bar{x}	n_i	f_i	$\bar{x}_i f_i$	$\bar{x}_i^2 f_i$
2,5	1	1/6	2,5/6	6,25/6
3	1	1/6	3/6	9/6
3,5	2	2/6	7/6	24,5/6
4	1	1/6	4/6	16/6
4,5	1	1/6	4,5/6	20,25/6
Total	6	1	21/6	76/6

Fonte: Do autor (2018).

De acordo com a tabela, temos $E(\bar{X}) = \sum \bar{x}_i f_i = \frac{21}{6} = 3,5 = \mu$, isto é, a média de todas as médias amostrais, extraídas sem reposição da população P, também é igual a média populacional μ , e a variância das médias amostrais $Var(\bar{X}) = \sum \bar{x}_i^2 f_i - \mu_{\bar{X}}^2 = \frac{76}{6} - (3,5)^2 = 0,417 = \frac{\sigma^2}{2} \cdot \frac{N-n}{N-1} = 1,25$. Temos, então, que o valor da variância das médias amostrais é "n" vezes menor que a variância populacional, em nosso caso particular 2 vezes menor, multiplicado por um fator de correção.

O fator de redução da variância na amostragem sem reposição é $\frac{N-n}{N-1}$, onde N é o número populacional e n o amostral. Este fator pode ser considerado como o fator de eficiência da amostragem sem reposição sobre a amostragem com reposição, que, em nosso caso, é $\frac{N-n}{N-1} = \frac{4-2}{4-1} = \frac{2}{3}$. Como na amostragem sem reposição não é possível retirar o mesmo elemento duas vezes, as médias não podem assumir valores tão extremos e, desse modo, o erro padrão na amostragem sem reposição é menor que na amostragem com reposição.

Dessa forma, exposta a distribuição da média amostral com e sem reposição na amostra, construímos a Tabela 10 a seguir:

Tabela 10 - Propriedades da distribuição da média amostral

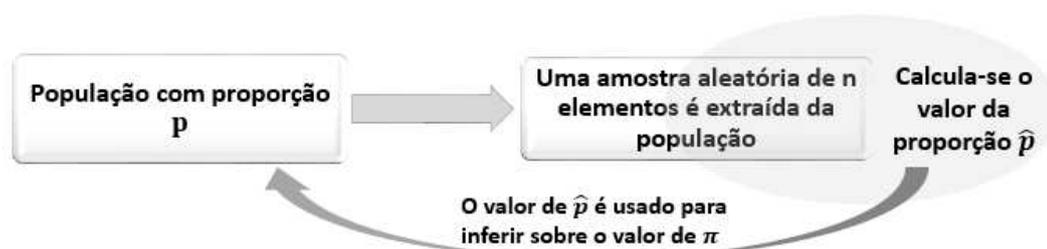
Com reposição na amostra			Sem reposição na amostra		
Média	Variância	Desvio-Padrão (erro padrão da média)	Média	Variância	Desvio-Padrão (erro padrão da média)
$E(\bar{X}) = \mu$	$Var(\bar{X}) = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$	$E(\bar{X}) = \mu$	$Var(\bar{X}) = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$

Fonte: Do autor (2018).

3.4.12.2. Distribuição amostral da proporção: com e sem reposição

Sabemos que a Estatística utiliza a proporção amostral \hat{p} para realizar inferências sobre as proporções p de uma população. Esse procedimento está representado na Figura 36.

Figura 36 - Processo da proporção amostral em relação à proporção populacional



Fonte: Do autor (2018).

Executando esse processo, continuamente, iremos encontrar valores distintos para a proporção da amostra. Por consequência, podemos dizer que a própria estimativa da proporção de uma variável aleatória X , também é uma variável aleatória. Sendo assim, a distribuição de probabilidade da variável aleatória \hat{p} é chamada de distribuição amostral da proporção, ou seja, é a distribuição de todos os valores possíveis da proporção da amostra.

Analisamos o processo de uma distribuição amostral da proporção, por meio de exemplos, levando-se em consideração amostras com reposição e sem reposição para que possamos entender e conhecer alguns resultados relevantes para a Estatística.

Começamos com um exemplo de amostragem com reposição: seja uma população em que a proporção dos elementos que possuem uma determinada característica é p . Dessa forma, definimos uma variável aleatória X como $X = 1$ quando o elemento possui uma determinada

característica, e $X = 0$ quando o elemento não possui uma determinada característica. Diante disso, vamos considerar uma população $P = (2,3,4,5)$ e todas as amostras possíveis de tamanho $n = 2$, ou seja, $k = N^n = 4^2 = 16$, onde k é o número de amostras extraídas, com reposição, e N o tamanho da população. Para cada amostra vamos calcular a proporção P de elementos pares na população. O Quadro 12, a seguir, apresenta as possíveis amostras com as suas respectivas proporções.

Quadro 12 - Amostras extraídas com reposição e suas respectivas proporções

Amostras	(2,2)	(2,3)	(2,4)	(2,5)	(3,2)	(3,3)	(3,4)	(3,5)
<i>Proporção das amostras</i>	1	1/2	1	1/2	1/2	0	1/2	0
Amostras	(4,2)	(4,3)	(4,4)	(4,5)	(5,2)	(5,3)	(5,4)	(5,5)
<i>Proporção das amostras</i>	1	1/2	1	1/2	1/2	0	1/2	0

Fonte: Do autor (2018).

Com os dados dispostos, construímos uma tabela de frequências $f(p)$ da distribuição da proporção amostral, junto com as demais informações necessárias para os cálculos da proporção e da variância das proporções amostrais.

Tabela 11 - Distribuição amostral das proporções extraídas das amostras com reposição

Proporção das amostras \hat{p}_i	f_i	$f_i p_i$	$f_i p_i^2$
0	4/16	0/16	0/16
1/2	8/16	4/16	2/16
1	4/16	4/16	4/16
Total	1	8/16	6/16

Fonte: Do autor (2018).

A Tabela 11 nos ajuda verificar que $E(P) = \sum p f(p) = \frac{8}{16} = 0,5 = p$, ou seja, o valor esperado de todas as proporções amostrais, extraídas com reposição da população, é igual a proporção populacional (parâmetro populacional p). Esse resultado, como iremos observar nas propriedades dos estimadores, mostra que o estimador \hat{p} é um estimador não tendencioso da proporção populacional π , quando as amostras são retiradas com reposição da população. Ainda, de acordo com a Tabela, temos que $Var(P) = \sigma_P^2 = \sum p^2 f(p) - \mu_P^2 = \frac{6}{16} - 0,5^2 =$

$0,125 = \frac{p(1-p)}{n}$, isto é, a variância entre as proporções amostrais é " n " vezes, e, em nosso caso particular, é 2 vezes menor que a variância populacional. O resultado mostra que, quando estamos trabalhando com proporções, podemos afirmar que a variância populacional é igual a $p(1-p)$. O valor de $\sigma_p = \sqrt{\frac{p(1-p)}{n}} = 0,353553 \dots$ é denominado de erro padrão da proporção que mede a variabilidade entre as proporções amostrais e revela a ideia de erro ao substituir a proporção populacional pela proporção amostral.

Verificamos agora um exemplo com amostragem sem reposição: vamos observar uma população $P = \{2,3,4,5\}$, e todas as amostras possíveis de tamanho $n = 2$, ou seja, $k = \binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{4!}{2!(4-2)!} = 6$, onde k é o número de amostras extraídas sem reposição e N o tamanho da população. Para cada amostra vamos calcular a proporção P de elementos pares na população. Dessa forma, o Quadro 13 apresenta as possíveis amostras com suas respectivas proporções.

Quadro 13 - Amostras extraídas com reposição e suas respectivas proporções

Amostras	(2,3)	(2,4)	(2,5)	(3,4)	(3,5)	(4,5)
<i>Proporção das amostras</i>	1/2	1	1/2	1/2	0	1/2

Fonte: Do autor (2018).

Com os dados dispostos construímos uma tabela de frequências $f(p)$ da distribuição da proporção amostral e com as demais informações necessárias para os cálculos da proporção e da variância das proporções amostrais.

Tabela 12 - Distribuição amostral das proporções extraídas das amostras sem reposição

Proporção das amostras	f_i	$f_i p_i$	$f_i p_i^2$
0	1/6	0/6	0/6
1/2	4/6	2/6	1/6
1	1/6	1/6	1/6
Total	1	3/6	2/6

Fonte: Do autor (2018).

A Tabela 12 nos auxilia no cálculo da $E(P) = \sum p f(p) = \frac{3}{6} = 0,5 = \pi$, ou seja, a média de todas as proporções amostrais, retiradas sem reposição da população P , é igual a proporção

populacional (parâmetro populacional π). Novamente, diante desse resultado, verificamos que o estimador \hat{p} é um estimador não tendencioso da proporção populacional π , quando as amostras são extraídas sem reposição. Com a Tabela 12 também conseguimos verificar que $Var(P) = \sum p^2 f(p) - \mu_p^2 = \frac{2}{6} - (0,5)^2 = 0,8333 \dots = \frac{\pi(1-\pi)}{n} \cdot \frac{N-n}{N-1}$, ou seja, a variância entre as populações amostrais é "n" vezes, em nosso caso particular, 2 vezes menor que a variância populacional multiplicada pelo fator de correção. Este fator pode ser considerado como fator de eficiência da amostragem, sem reposição, sobre a amostragem com reposição que, em nosso exemplo, vale $\frac{N-n}{N-1} = \frac{2}{3}$, onde N é o número total da população e n o tamanho da amostra. Dessa forma, exposta a distribuição amostral da proporção com e sem reposição na amostra, construímos a Tabela 13.

Tabela 13 - Propriedades da distribuição amostral

Com reposição na amostra			Sem reposição na amostra		
Média	Variância	Desvio-Padrão (erro padrão da média)	Média	Variância	Desvio-Padrão (erro padrão da média)
$E(\bar{p}) = p$	$Var(\bar{p}) = \sigma_{\bar{p}}^2 = \frac{p(1-p)}{n}$	$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$	$E(\bar{p}) = p$	$Var(\bar{p}) = \sigma_{\bar{p}}^2 = \frac{p(1-p)}{n} \cdot \frac{N-n}{N-1}$	$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n} \cdot \frac{N-n}{N-1}}$

Fonte: Do autor (2018).

3.5. Estimação de parâmetros populacionais

A estimação de parâmetros trata da questão de avaliação de parâmetros populacionais a partir de operações com os dados de uma amostra. É um raciocínio tipicamente indutivo, onde se generalizam resultados obtidos na amostra para toda uma população. Dessa forma, podemos definir a estimação como uma ferramenta estatística que consiste em determinar um valor amostral que substitua o respectivo valor real do parâmetro populacional desconhecido. Observamos que uma Estatística é normalmente uma medida descritiva, que é função dos elementos contidos na amostra. Quando uma Estatística é usada para avaliar algum parâmetro da população, é também chamada de estimador. Dessa forma, pelo fato de depender dos particulares elementos selecionados na amostra, o estimador é também uma variável aleatória. Uma vez realizada uma amostragem, ao valor calculado para o estimador, a esta amostra dá-se o nome de estimativa. Existem dois tipos de estimativas: a estimativa pontual e a estimativa intervalar.

Figura 37 - Tipos de estimação



Fonte: Do autor (2018).

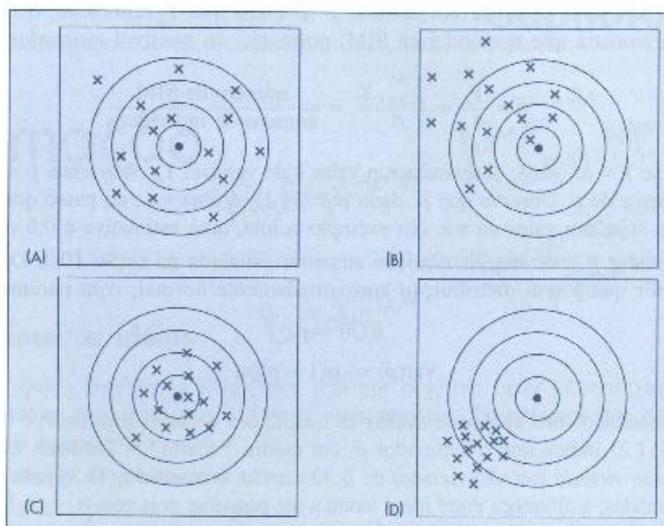
Chamamos de estimação pontual quando, a partir de uma amostra, um único valor é usado para estimar um parâmetro desconhecido. Um estimador pontual para um parâmetro populacional θ , é geralmente representado por $\hat{\theta}$. Em nossa pesquisa observamos os estimadores da média, da variância, do desvio padrão e da proporção populacionais, que são, respectivamente, \bar{X} , S^2 , S e \hat{p} .

Quando encontramos uma estimativa pontual ela raramente coincide com o valor do parâmetro. Uma desvantagem do uso de estimadores pontuais é que, se nenhuma informação adicional for dada, não há maneira de se avaliar o resultado da estimativa, pois não temos nenhuma ideia da sua precisão. Um procedimento mais desejável para estimação é, então, calcular um intervalo que tenha uma probabilidade pré-estabelecida de conter o parâmetro desconhecido. O processo da estimativa intervalar produz uma estimativa que está incluída em um intervalo, dentro de um certo grau de acerto, determinado por um intervalo de confiança que contém a estimativa pontual. Ao definirmos um grau de acerto, automaticamente também fica definido um grau de erro, ao qual chamamos de margem de erro da estimativa e que resulta em um erro de estimativa, ou erro amostral.

3.6. Propriedades dos estimadores

Ao introduzir as propriedades dos estimadores, muitos autores de livros estatísticos como Freund (2006), Martins (2008), Morettin e Bussab (2010), iniciam suas exposições fazendo uma analogia. Sendo assim, Morettin e Bussab (2010), mencionam um teste de pontaria realizado com quatro rifles (A, B, C e D) sobre um alvo onde foram considerados alguns critérios para a escolha do melhor rifle, cujo resultado é ilustrado pela Figura 38.

Figura 38 - Resultado do teste de tiro ao alvo



Fonte: Morettin e Bussab (2010).

Considerando um critério relativo à média de acertos no alvo, os rifles A e C seriam os escolhidos. Poderíamos escolher os rifles C e D caso o critério de escolha fosse o que possui menor dispersão ou pequena variância. Diante dos critérios expostos, o rifle C seria o escolhido por agregar os dois critérios estabelecidos. Essa analogia faz referência a dois conceitos importantes para a escolha de um estimador quais sejam: acurácia e precisão. A acurácia, como expõe Freund (2006), mede a proximidade de cada observação do valor desejado, e a precisão mede a proximidade de cada observação da média de todas as observações.

Dessa forma, podemos dizer que no alvo A as observações são não tendenciosas, pouco acuradas e com baixa precisão. No B as observações são tendenciosas, pouco acuradas e com baixa precisão. Em C as observações são não tendenciosas, muito acuradas e com boa precisão. E, finalmente, em D, as observações são tendenciosas, pouco acuradas e com alta precisão.

De modo geral, a questão da estimação está caracterizada em selecionar uma amostra (X_1, X_2, \dots, X_n) de uma variável aleatória que irá descrever uma característica de interesse de uma população estudada. Um estimador $\hat{\theta}$ do parâmetro θ é qualquer função das observações da amostra, ou seja, $\hat{\theta} = g(X_1, X_2, \dots, X_n)$. Como relata Morettin e Bussab (2010), o problema da estimação é, então, determinar um estimador $\hat{\theta}$ que consiga descrever o parâmetro θ , levando-se em consideração alguns critérios. Esses critérios, que são as propriedades para um bom estimador, estão apresentados na Figura 39.

Figura 39 - Propriedades para um bom estimador



Fonte: Do autor (2018).

3.6.1. Estimador não tendencioso

Essa propriedade, também chamada por Lethielleux e Chevalier (2016) de estimador sem distorção, revela que um estimador é não viciado ou não tendencioso quando o seu valor esperado for igual ao verdadeiro valor do parâmetro.

$$E(\hat{\theta}) = \theta$$

De acordo com a propriedade, para qualquer estimador $\hat{\theta}$, onde $E(\hat{\theta}) = \theta + b(\theta)$, com $b(\theta) \neq 0$, teremos um estimador viciado. Esse valor de $b(\theta)$ é chamado de vício ou viés. Tomamos alguns exemplos para mostrar essa propriedade levando-se em consideração uma amostra aleatória (X_1, X_2, \dots, X_n) de uma variável aleatória X , com determinada distribuição de probabilidades.

Exemplo 1: \bar{X} é um estimador não-tendencioso de μ . Dessa forma, temos $E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \frac{n \cdot \mu}{n} = \mu$.

Exemplo 2: $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ é um estimador tendencioso para σ^2 .

Consideremos uma população com N elementos e a variância populacional $\sigma^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)^2$, onde $\mu = \frac{1}{N} \sum_{i=1}^N X_i$ é a média populacional.

Um possível estimador para σ^2 , baseado numa amostragem aleatória simples é

$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ onde \bar{X} é a média amostral e n é o tamanho da amostra. Vamos

mostrar que esse estimador é viesado.

A esperança do estimador $\hat{\sigma}^2$ é

$$E(\hat{\sigma}^2) = E\left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2\right) = \frac{1}{n} \sum_{i=1}^n E(X_i^2) - E(\bar{X}^2) \quad (1)$$

Sabemos que $Var(X_i) = E(X_i^2) - [E(X_i)]^2$, implicando que $E(X_i^2) = Var(X_i) + [E(X_i)]^2$ (2).

Ainda, temos $Var(\bar{X}) = E(\bar{X}^2) - [E(\bar{X})]^2$, implicando $E(\bar{X}^2) = Var(\bar{X}) + [E(\bar{X})]^2$ (3).

Em relação a variância da média, $Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^n X_i\right)$ e, pelas propriedades apresentadas da variância, temos que $Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n^2}Var(\sum_{i=1}^n X_i) = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$ (4).

Considerando a substituição em (1) por (2), (3) e (4), temos:

$$\begin{aligned} E(\hat{\sigma}^2) &= \frac{1}{n} \sum_{i=1}^n E(X_i^2) - E(\bar{X})^2 = \frac{1}{n} \sum_{i=1}^n \{Var(X_i) + [E(X_i)]^2\} - \left(\frac{\sigma^2}{n} + \mu^2\right) = \\ &= \frac{1}{n} \sum_{i=1}^n (\sigma^2 + \mu^2) - \frac{\sigma^2}{n} - \mu^2 = \frac{1}{n} n(\sigma^2 + \mu^2) - \frac{\sigma^2}{n} - \mu^2 = \sigma^2 - \frac{\sigma^2}{n} = \\ &= \sigma^2 \left(1 - \frac{1}{n}\right) = \sigma^2 \cdot \frac{n-1}{n} \end{aligned}$$

Sendo assim, encontramos a igualdade $E(\hat{\sigma}^2) = \left(\frac{n-1}{n}\right) \sigma^2$.

De acordo com a propriedade do estimador não tendencioso, essa igualdade apresenta um viés que pode ser escrito como $-\frac{\sigma^2}{n}$ e, portanto, $\hat{\sigma}^2$ é viesado para σ^2 . Para obter um estimador não tendencioso, utilizamos, como mostra Morettin & Bussab (2010), um fator $\frac{(n-1)}{n}$.

Dessa forma, temos, então, $E(S^2) = E\left(\frac{n}{n-1} \hat{\sigma}^2\right) = \frac{n}{n-1} \cdot \frac{\sigma^2(n-1)}{n} = \sigma^2$. Assim, obtemos um estimador não tendencioso para a variância da amostra que pode ser definida por $S^2 = \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ ou, ainda, $S^2 = \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2\right) = \frac{1}{n-1} (\sum_{i=1}^n X_i^2 - n\bar{X}^2)$. Morettin & Bussab (2010) previnem que, em casos onde a média populacional é conhecida, o estimador não será dividido por $(n-1)$.

Portanto, teremos $S^2 = \frac{1}{n} [\sum_{i=1}^n (X_i)^2 - n\mu^2]$, então, $E(S^2) = \frac{1}{n} \sum_{i=1}^n E(X_i^2) - E(\mu^2) = \frac{1}{n} \sum_{i=1}^n \{Var(X_i) + [E(X_i)]^2\} - \mu^2 =$
 $= \frac{1}{n} \sum_{i=1}^n (\sigma^2 + \mu^2) - \mu^2 = \frac{1}{n} n(\sigma^2 + \mu^2) - \mu^2 = \sigma^2$.

3.6.2. Estimador consistente

A consistência é uma propriedade que zela pela precisão da estimativa levando em consideração o tamanho da amostra, ou seja, a exatidão de uma estimativa aumenta quando o tamanho da amostra aumenta. A acurácia de uma estimativa aumenta quando o tamanho da amostra aumenta. Em outras palavras, como descrevem Lethielleux e Chevalier (2016), se um estimador é não tendencioso, $E(\hat{\theta}) = \theta$, ele será consistente se o tamanho da amostra aumentar e a variabilidade do estimador diminuir. Dessa forma, as observações ficarão cada vez mais concentradas em torno do parâmetro na medida em que a amostra vai ficando cada vez maior.

Morettin e Bussab (2010) explicitam que um estimador $\hat{\theta}$ é consistente se a sua probabilidade diferir do verdadeiro valor θ em menos do que c , onde c é um número arbitrário positivo e pequeno que tende a 1 quando o tamanho da amostra aumenta, ou seja, $\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta|) = 1$. Podemos interpretar, dessa forma, que quando n aumenta, a estimativa de $\hat{\theta}$ está tendenciosamente mais próxima do verdadeiro parâmetro θ .

Martins (2008) esclarece que essa propriedade é uma propriedade assintótica de um estimador e é aplicada para amostras suficientemente grandes. As condições suficientes para um estimador ser consistente são: $\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta$ e $\lim_{n \rightarrow \infty} Var(\hat{\theta}) = 0$.

Para exemplificar, vamos considerar a distribuição amostral da média baseada em amostras aleatórias simples com reposição de tamanho n , onde $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \frac{Var(X)}{n}$. O crescimento de n faz com que a distribuição de \bar{X} concentre-se mais em torno de μ . Diz-se que \bar{X} é um estimador consistente da média populacional μ . Nesse sentido, o estimador proporcional \hat{p} é tal que $Var(\hat{p}) \rightarrow 0$, ou seja, a sua variância tende a zero, quando $n \rightarrow \infty$. Por esse fato e por $E(\hat{p}) = p$, chamamos \hat{p} um estimador de p .

3.7. Intervalo de Confiança

Observamos que a estimação por ponto, onde a estimativa de um parâmetro populacional é realizada por um único valor, não permite ponderar possíveis erros em relação aos parâmetros desconhecidos da população pesquisada. Nesse sentido, surge a ideia de estabelecer intervalos de confiança que são baseados na distribuição amostral dos estimadores.

Trata-se de uma técnica da estimação estatística, onde a partir de um intervalo de confiança construído com os elementos da amostra, pode-se inferir sobre um parâmetro populacional.

Devido a variabilidade amostral, as possíveis amostras aleatórias, de mesmo tamanho, retiradas da mesma população, terão medidas diferentes, e é nesse momento que surge a dúvida sobre a confiabilidade de uma estimativa pontual. O intervalo de confiança foi instituído para definir, de forma objetiva, a credibilidade da estimativa por meio de um intervalo de valores que contém o parâmetro da população, com uma determinada probabilidade de acerto, e é construído a partir de uma amostra aleatória retirada da população.

A estimativa de um parâmetro populacional, dada por dois valores **a** e **b**, onde consideramos o valor **b** maior que o valor de **a**, entre os quais podemos dizer que o parâmetro esteja contido, é denominada estimativa por intervalo. Dessa forma, ao invés de propormos um único valor como estimativa de um parâmetro, podemos obter, com base na amostra, um intervalo de valores ao qual atribuímos um certo grau ou coeficiente de confiança que contém o verdadeiro valor desconhecido desse parâmetro.

Para exemplificar, tomamos o valor $1,75m$ que corresponde à média da altura de certos indivíduos, ou seja, estamos fazendo uso de uma estimativa pontual. Por outro lado, se falarmos que a medida da altura é $1,75 \pm 0,05m$ a estimativa é por intervalo, isto é, estamos afirmando que a altura dos indivíduos está em um intervalo situado entre $1,70m$ e $1,80m$. De forma formal, seja X_1, X_2, \dots, X_n uma amostra aleatória de tamanho n e θ um parâmetro populacional desconhecido dessa população. Um intervalo de confiança para θ é estabelecido observando a amostra de modo que ele contenha o verdadeiro e desconhecido valor de θ , com uma particular e elevada probabilidade.

A notação dessa probabilidade é $1 - \alpha$ e indicamos por $P(a < \theta < b) = 1 - \alpha$. Dessa forma, o intervalo $]a, b[$ é chamado de intervalo com $100(1 - \alpha)\%$ de confiança para o parâmetro θ , onde $1 - \alpha$ é o nível de confiança associado ao intervalo **a** e **b** que são os limites de confiança, inferior e superior, respectivamente, do intervalo. Geralmente, trabalha-se com um nível e confiança entre 90% e 99%. Valores menores que 90% não são utilizados pois possuem baixa precisão, ou seja, a confiabilidade é pequena. Valores acima de 99%, embora consistam em uma confiança elevada, acarretam problemas de cálculo pois obtemos intervalos muito grandes ou necessitamos de amostras muito grandes, o que pode inviabilizar uma investigação. Analisamos em nossa pesquisa a estimação por intervalo levando-se em consideração a distribuição amostral da média e da proporção.

3.7.1. Estimaco intervalar da mdia populacional

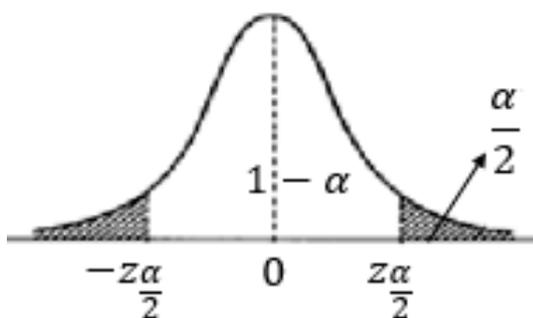
Para realizar a anlise da estimaco intervalar da mdia populacional, vamos considerar dois casos: quando o desvio padro populacional σ  conhecido e quando o desvio padro populacional σ  desconhecido.

3.7.1.1. Desvio Padro populacional σ conhecido

O intervalo de confiana para a mdia populacional μ  estabelecido considerando a estimativa pontual \bar{X} . De acordo com a distribuico amostral da mdia, temos que $\bar{X} \equiv N\left(\mu, \frac{\sigma^2}{n}\right)$ para populao infinita ou com reposio na amostra, e $\bar{X} \equiv \left(\mu, \frac{\sigma^2}{n} \frac{N-n}{N-1}\right)$ para populao finita ou sem reposio na amostra.

Para construir este intervalo, fixamos uma probabilidade $1 - \alpha$ de modo que o intervalo contenha o parmetro populacional. Sendo assim, α ser a probabilidade de que o intervalo obtido no contenha o valor do parmetro, isto , α ser a probabilidade de erro. Nosso desejo  construir um intervalo que contenha o parmetro populacional μ com probabilidade $1 - \alpha$, dessa forma temos, $P\left(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}\right) = 1 - \alpha$, onde $z_{\frac{\alpha}{2}}$  o valor normal padro com rea a direita igual a $\frac{\alpha}{2}$.

Grfico 3 - Curva Normal Padro para o intervalo de confiana da mdia populacional com σ conhecido



Fonte: Do autor (2018).

No caso de populaes infinitas ou com reposio na amostra, temos que a varivel padronizada de \bar{X}  $Z = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}$ e, substituindo em $P\left(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}\right) = 1 - \alpha$, encontramos $\left(-z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$, desenvolvendo essa desigualdade chegamos ao intervalo

procurado $P\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$. Assim, o Quadro mostra o intervalo de confiança de $1 - \alpha$ para a média populacional quando σ for conhecido em populações infinitas ou com reposição na amostra.

Quadro 14 - Intervalo de confiança para a média populacional com σ conhecido em populações infinitas ou com reposição na amostra

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

Fonte: Do autor (2018).

De acordo com o intervalo encontrado, temos que \bar{X} é a estimativa pontual da média populacional, σ é o desvio padrão da populacional, $z_{\frac{\alpha}{2}}$ indica o valor da distribuição normal cuja área à direita é igual a $\frac{\alpha}{2}$, ou seja, é o valor de Z tal que $P\left(Z > z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$. Admitindo uma probabilidade $1 - \alpha = 0,90$, teríamos $\alpha = 0,10$ e $\frac{\alpha}{2} = 0,05$ que devemos buscar na tabela da distribuição Normal. Elaboramos o Quadro 15 abaixo com os principais níveis de interesse:

Quadro 15 - Principais níveis de interesse

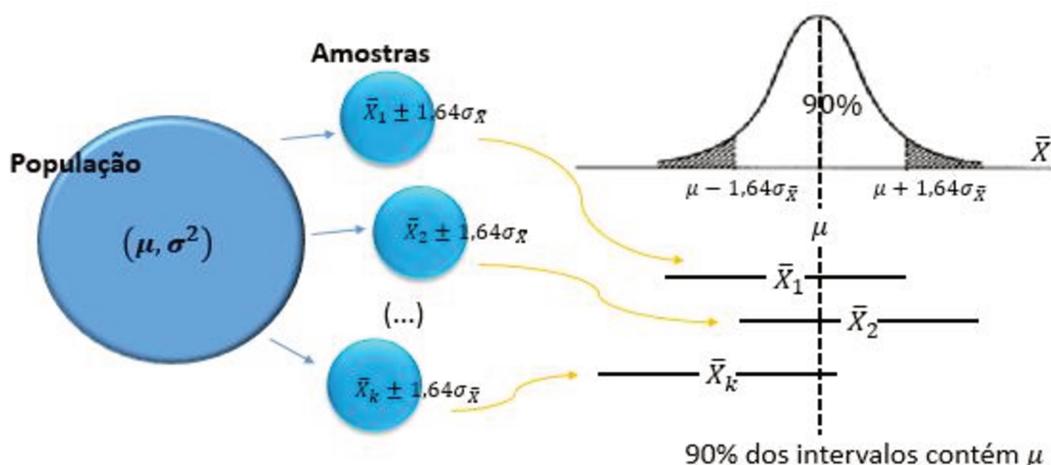
IC (1 - α)	90,00%	95,00%	95,50%	97,50%	99,00%
α	10,00%	5,00%	4,50%	2,50%	1,00%
(1 - $\alpha/2$)	95,00%	97,50%	97,75%	98,75%	99,50%
$Z_{\alpha/2}$	1,64	1,96	2,00	2,24	2,58

Fonte: Do autor (2018).

Temos, então $P\left(\bar{X} - 1,64 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1,64 \frac{\sigma}{\sqrt{n}}\right) = 0,90$.

A Figura apresenta uma construção de intervalos da forma $\bar{X} \pm 1,64\sigma_{\bar{X}}$, onde 90% contém μ que não é uma variável aleatória, mas um parâmetro, o que não significa dizer que μ tem 90% de probabilidade de estar entre os limites indicados.

Figura 40 - Significado de um intervalo de confiança para μ com $(1 - \alpha) = 0,90$ e σ conhecido



Fonte: Do autor (2018).

No caso de populações finitas ou sem reposição na amostra, temos $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}}$ e,

portanto, o intervalo de confiança será:

$$P\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}\right) = 1 - \alpha.$$

Quadro 16 - Intervalo de confiança para a média populacional com σ conhecido em populações finitas ou sem reposição na amostra

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \right]$$

Fonte: Do autor (2018).

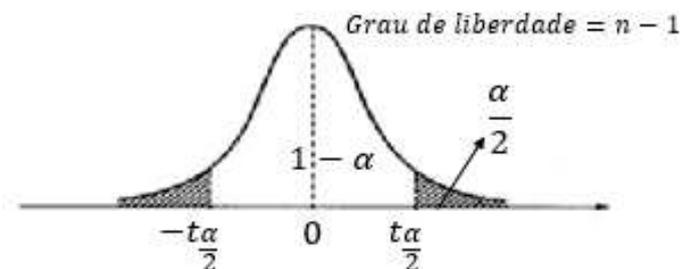
3.7.1.2. Desvio Padrão populacional σ desconhecido

O processo de obtenção do intervalo de confiança, neste caso, é bastante semelhante ao caso anterior. A grande diferença é que, ao substituirmos a variância populacional pelo seu estimador que é a variância amostral, a variável normalizada passa a ser construída pelo quociente entre duas variáveis aleatórias: \bar{X} e S , o que implica que a distribuição da variável

$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$ para populações infinitas ou com reposição na amostra e $t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}}$ para populações

finitas ou sem reposição na amostra passe a ter uma distribuição de t Student com $n - 1$ graus de liberdade. Dessa forma, temos:

Gráfico 4 - Curva Normal Padrão para o intervalo de confiança da média com σ desconhecido



Fonte: Do autor (2018).

A notação t é requerida por que a variável aleatória S , no denominador, aumenta a variância de t para um valor maior do que um, de modo que a razão não é padronizada. A designação $n - 1$ graus de liberdade (G.L.) é necessária por que para cada diferente tamanho de amostra n ou valor $n - 1$, existe uma diferente distribuição t . O grau de liberdade é conceituado como o número de valores independentes de uma estatística, tomando como exemplo o estimador S^2 de σ^2 , onde a quantidade $n - 1$ é o divisor que aparece na expressão de S^2 . Não faremos uma análise sobre a distribuição t Student pois na realidade existem infinitas distribuições t , uma para cada tamanho de amostra. Estas distribuições a exemplo da Normal Padrão, encontram-se tabeladas. Considerando que os valores das abscissas $t_{\frac{\alpha}{2}}$ da distribuição t dependem do tamanho da amostra n e do desvio padrão amostral, podemos calcular referido valor para alguns níveis de confiabilidade de interesse e para alguns tamanhos usuais de amostra, conforme o Quadro 17.

Quadro 17 - Principais níveis de interesse

IC ($1 - \alpha$)	90,00%	95,00%	95,50%	97,50%	99,00%
α	10,00%	5,00%	4,50%	2,50%	1,00%
$t_{\alpha/2}, n=10, gl=9$	1,83	2,26	2,33	2,69	3,25
$t_{\alpha/2}, n=20, gl=19$	1,73	2,09	2,15	2,43	2,86
$t_{\alpha/2}, n=30, gl=29$	1,70	2,05	2,10	2,36	2,76
$t_{\alpha/2}, n=50, gl=49$	1,68	2,01	2,06	2,31	2,68
$t_{\alpha/2}, n=100, gl=99$	1,66	1,98	2,03	2,28	2,63

Fonte: Do autor (2018), com base em uma tabela t student.

Contudo, queremos encontrar $P\left(-t \frac{\alpha}{2} \leq t \leq t \frac{\alpha}{2}\right) = 1 - \alpha$, substituindo as inequações, encontraremos os seguintes intervalos:

Quadro 18 - Intervalo de confiança para a média populacional com σ desconhecido em populações finitas ou sem reposição na amostra

$$P\left(\bar{X} - t \frac{\alpha S}{2 \sqrt{n}} \leq \mu \leq \bar{X} + t \frac{\alpha S}{2 \sqrt{n}}\right) = 1 - \alpha$$

Fonte: Do autor (2018).

Quadro 19 - Intervalo de confiança para a média populacional com σ desconhecido em populações finitas ou sem reposição na amostra

$$P\left(\bar{X} - t \frac{\alpha S}{2 \sqrt{n}} \sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{X} + t \frac{\alpha S}{2 \sqrt{n}} \sqrt{\frac{N-n}{N-1}}\right) = 1 - \alpha$$

Fonte: Do autor (2018).

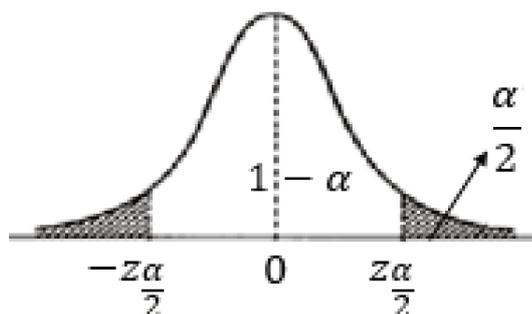
3.7.2. Intervalo de confiança para a proporção

O intervalo de confiança para a proporção populacional p é estabelecido considerando a estimativa pontual \hat{p} . De acordo com a distribuição amostral da proporção, temos $\hat{p} \equiv N\left(p, \frac{p(p-1)}{n}\right)$ para população infinita ou com reposição na amostra e $\hat{p} \equiv \left(p, \frac{p(p-1)N-n}{N-1}\right)$ para população finita ou sem reposição na amostra.

Para construir este intervalo, fixamos uma probabilidade $1 - \alpha$ de modo que o intervalo contenha o parâmetro populacional. Sendo assim, α será a probabilidade de que o intervalo obtido não contenha o valor do parâmetro, isto é, α será a probabilidade de erro.

Nosso desejo é construir um intervalo que contenha o parâmetro populacional p com probabilidade $1 - \alpha$, dessa forma temos, $P\left(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}\right) = 1 - \alpha$, onde $z_{\frac{\alpha}{2}}$ é o valor normal padrão com área a direita igual a $\frac{\alpha}{2}$.

Figura 41 - Curva Normal Padrão para o intervalo de confiança da proporção



Fonte: Do autor (2018).

No caso de populações infinitas ou com reposição na amostra, temos que a variável padronizada de \hat{p} é $Z = \frac{(\hat{p}-p)}{\sqrt{\frac{p(1-p)}{n}}}$ e, substituindo em $P\left(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}\right) = 1 - \alpha$, e desenvolvendo essa desigualdade chegamos ao intervalo procurado $P\left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right) = 1 - \alpha$.

Quadro 20 - Intervalo de confiança para a proporção de populações infinitas ou com reposição na amostra

$$\left[\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$$

Fonte: Do autor (2018).

No caso de populações finitas ou sem reposição na amostra, temos que a variável padronizada de \hat{p} é $Z = \frac{(\hat{p}-p)}{\sqrt{\frac{p(1-p)(N-n)}{n(N-1)}}}$ e, portanto, o intervalo de confiança será

$$P\left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)(N-n)}{n(N-1)}} < p < \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)(N-n)}{n(N-1)}}\right) = 1 - \alpha.$$

Quadro 21 - Intervalo de confiança para a proporção de populações finitas ou sem reposição na amostra

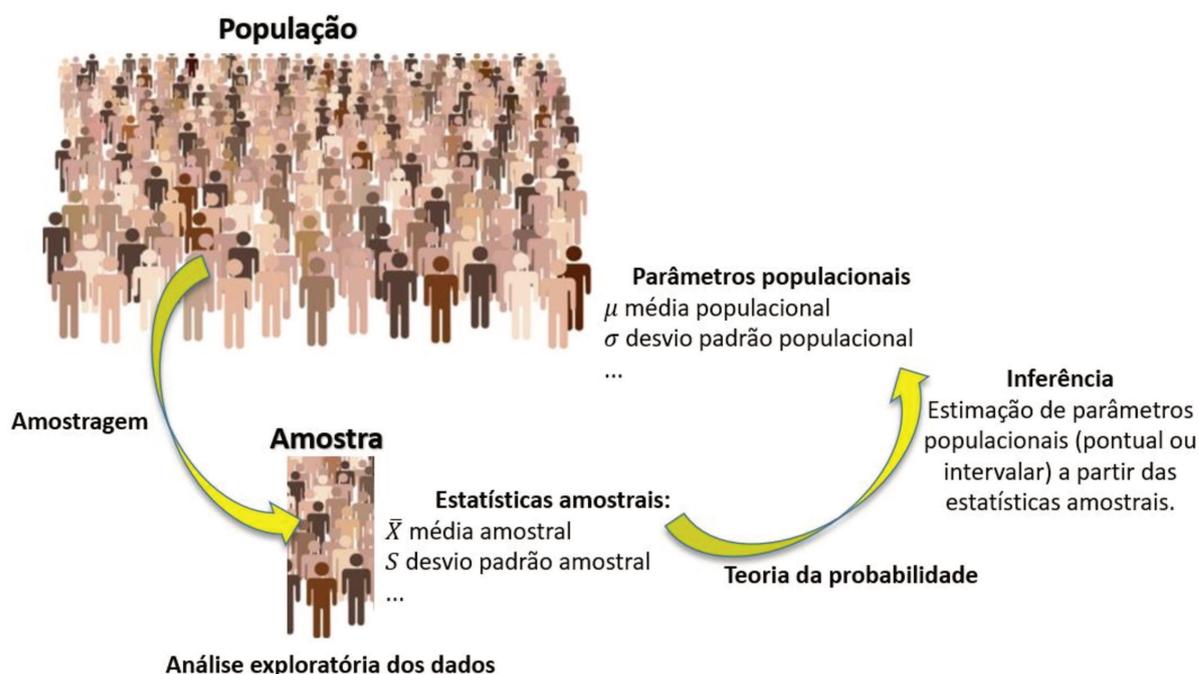
$$\left[\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} \left(\frac{N-n}{N-1} \right)}; \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} \left(\frac{N-n}{N-1} \right)} \right]$$

Fonte: Do autor (2018).

3.8. Fechamento sobre a estimação de parâmetros populacionais

De um modo conciso, procuramos explorar algumas das principais definições que promovem uma estruturação básica para um bom entendimento sobre a estimação de parâmetros populacionais. A teoria da probabilidade e conceitos estatísticos como o de população, amostra, média, desvio padrão e outros, são alguns dos fundamentos para a implementação de uma estatística de inferência que ocupa a estimação de parâmetros populacionais para tomadas de decisões. Diante disso, podemos definir a estimação de parâmetros populacionais como um método de decisão que permite estimar características populacionais a partir de dados amostrados de uma determinada população, considerando os diversos fatores que caracterizam os casos a serem observados.

Figura 42 - Estimação de parâmetros populacionais



Fonte: Do autor (2018).

Para que pudéssemos melhor compreender o modo como esse conhecimento é construído, por meio da implementação ativa dos estudantes na resolução de situações que envolvem a estimação, evocamos as contribuições da Teoria dos Campos Conceituais. Essa teoria enfatiza a importância fundamental da atividade efetiva produzida pelos estudantes no processo de aprendizagem. Para Vergnaud (1990a), um conceito não pode ser reduzido a uma única definição, deve, então, ser estudado por meio de diversas situações que caracterizam esse conceito. Não obstante a uma ampla variedade de situações relacionadas a um conceito, a condução das atividades deve ser organizada de acordo com características comuns, recebendo um tratamento idêntico. Desse modo, as atividades formam classes similares que irão relacionar outros conceitos para a formação de conjuntos e subconjuntos que irão compor um campo conceitual para a construção de um novo conceito. Esses são os fundamentos da Teoria dos Campos Conceituais, que será descrito no capítulo seguinte.

4. Alguns elementos sobre a Teoria dos Campos Conceituais

Pertencente à tradicional escola piagetiana, Gérard Vergnaud, atualmente professor emérito do Centro Nacional de Pesquisa Científica de Paris, é um psicólogo, matemático e filósofo francês que desenvolveu a teoria dos campos conceituais investigando o sujeito do conhecimento em resposta a uma situação de ensino. O autor procurou, em sua teoria, redirecionar o foco piagetiano do sujeito epistêmico³⁹ para o sujeito-em-situação. Esse movimento do objeto central da análise tem como propósito a busca pela compreensão de como o sujeito aprende em uma determinada situação, observando a conexão necessária de conceitos para um domínio epistemológico específico. A teoria dos campos conceituais nasce com o intuito de investigar, na didática da matemática, as estruturas multiplicativas, e vem sendo difundida e aplicada em diversas áreas do conhecimento. Em nossa pesquisa, apuramos questões pertinentes da teoria para investigar o ensino da Estatística.

Para a teoria dos campos conceituais, um bom desenvolvimento cognitivo depende densamente da situação e da conceitualização particulares. Entende-se que a “situação” é qualquer tarefa, teórica ou empírica, a ser realizada pelo sujeito. Segue Vergnaud (1990b, p. 52):

³⁹ O termo “sujeito epistêmico” surgiu da epistemologia genética de Piaget que evidencia os processos cognitivos da espécie humana diferenciando-os dos processos cognitivos do indivíduo (sujeito psicológico). O interesse de Piaget em modelos mais gerais e descontextualizados acabou ocultando informações que revelassem as ações humanas em relação uma proposta de análises sobre o desenvolvimento de conceitos específicos.

[...] o saber se forma a partir de problemas para resolver, quer dizer, de situações para dominar. [...] Por ‘problema’ é preciso entender, no sentido amplo, toda situação na qual é preciso descobrir relações, desenvolver atividades de exploração, de hipótese e de verificação, para produzir uma solução.

A proposta do autor, em analisar o sujeito-em-situação, propicia uma melhor compreensão da evolução temporal dos sujeitos à medida em que aprendem, bem como nos faz refletir sobre as escolhas de intervenções didáticas situadas nas características dos conteúdos a serem abordados. O autor explica que o desenvolvimento cognitivo está fortemente calcado no conteúdo a ser ensinado, onde o ponto fundamental da cognição é o processo de conceitualização do real, atividade psicológica interna ao sujeito, que não pode ser reduzida a operações lógicas gerais e nem em operações puramente linguísticas.

Vergnaud (1990b) esclarece que o desenvolvimento cognitivo não pode ser compreendido por simples modelos que recorrem para ideias de reprodução social mediante estruturas inatas do sujeito, ou por meio de representações mentais como processamento de informação. No entanto, Piaget (1985) nos revela dispositivos para uma análise de mecanismos mais gerais do desenvolvimento do sujeito que podem conduzir às aprendizagens. A Teoria da Equilibração, com seus conceitos de assimilação, acomodação, perturbação, compensação e equilíbrio majorante, nos fornece uma extensa base para explicar o surgimento de novas informações em relação as operações utilizadas por um sujeito frente a um objeto de conhecimento. Nessa perspectiva, a teoria piagetiana surge como a essência preservada na teoria de campos conceituais que acolhe, como base, o conceito piagetiano de esquema. A teoria dos campos conceituais tenta explicar o desenvolvimento dos processos de conceitualização, partindo do princípio que a maior parte dos nossos conhecimentos são formados por competências (informações e habilidades) que estão disponíveis sob a forma de esquemas.

De acordo com a teoria muitos fatores estão conectados para a aquisição de um novo conhecimento, como as experiências vividas no mundo físico e as influências culturais. O conhecimento, segundo Vergnaud (1990a), está organizado em campos conceituais que são conjuntos informais e heterogêneos de problemas, situações, conceitos, relações, estruturas e operações de pensamento, unidos uns aos outros e, provavelmente, atrelados durante todo o processo de aprendizagem.

Vergnaud (1990b, p. 78) comunica que o objetivo da teoria é o de

[...] fornecer um quadro para a investigação sobre as atividades cognitivas complexas, principalmente na aprendizagem científica e técnica. É uma teoria psicológica do conceito, ou melhor ainda, a conceitualização da realidade: pode-se identificar e estudar as filiações e rupturas entre o conhecimento do ponto de vista do seu conteúdo

conceitual; ele também permite analisar a relação entre conceitos como conhecimento explícito e, invariantes operacionais que estão implícitos nas situações e aprofundar a análise da relação entre significado e significante. Os exemplos são tomados a partir de vários campos conceituais: estruturas aditivas, estruturas multiplicativas, aulas de lógica, álgebra. (Vergnaud 1990, tradução nossa⁴⁰)

Por se tratar de uma teoria complexa que lida com o desenvolvimento cognitivo e com a aprendizagem, levando em consideração os próprios conteúdos do conhecimento e a análise conceitual do seu domínio, vislumbra-se nela uma teoria que nos rende um terreno fértil para implementar em pesquisas na área da Educação.

Revelando o uso de sua teoria, Vergnaud (1990a), elenca três justificativas como forma de análise para a questão da obtenção de conhecimento:

- a) Um conceito não se forma dentro de um só tipo de situação, sugerindo uma quantidade diversificada de atividades de ensino, onde o sujeito pode constatar a aplicação de um conceito em inúmeras situações, fazendo a integração entre as partes e o todo;
- b) Uma situação não se analisa com um só conceito, o que sugere uma visão integradora do conhecimento. Entende-se que o aprendiz deve ser submetido a atividades didáticas onde tenha tempo de construir, testar e validar seus próprios modelos explicativos;
- c) A construção e apropriação de todas as propriedades de um conceito ou todos os aspectos de uma situação é um processo longo, ou seja, uma progressão dos modelos pessoais para os modelos científicos.

Compreende-se que a teoria dos campos conceituais não está direcionada à construção de uma teoria geral para o desenvolvimento, mas sim preocupada em relacionar o desenvolvimento do sujeito com as tarefas a que serão submetidos. Dessa forma, a cognição está fortemente elencada nas situações.

Vergnaud (1996), afirma que o processo de desenvolvimento cognitivo, por ser dependente das situações a serem enfrentadas pelo sujeito, possui, como essência, a construção de conceitos, ou seja, a conceitualização. Dessa forma, esta torna-se um processo longo que

⁴⁰ L'objet de la théorie des champs conceptuels est de fournir un cadre aux recherches sur les activités cognitives complexes, principalement sur les apprentissages scientifiques et techniques. C'est une théorie psychologique du concept, ou mieux encore, de la conceptualisation du réel: elle permet de repérer et d'étudier les filiations et les ruptures entre connaissances du point de vue de leur contenu conceptuel; elle permet également d'analyser la relation entre les concepts comme connaissances explicites, et les invariants opératoires qui sont implicites dans les conduites des sujets en situation, ainsi que d'approfondir l'analyse des relations entre signifiés et signifiants. Les exemples sont pris dans plusieurs champs conceptuels: les structures additives, les structures multiplicatives, la logique des classes, l'algèbre.

requer uma variedade diversificada de situações. A Figura 43 apresenta um mapa conceitual referente a alguns aspectos da teoria dos campos conceituais segundo nossa interpretação.

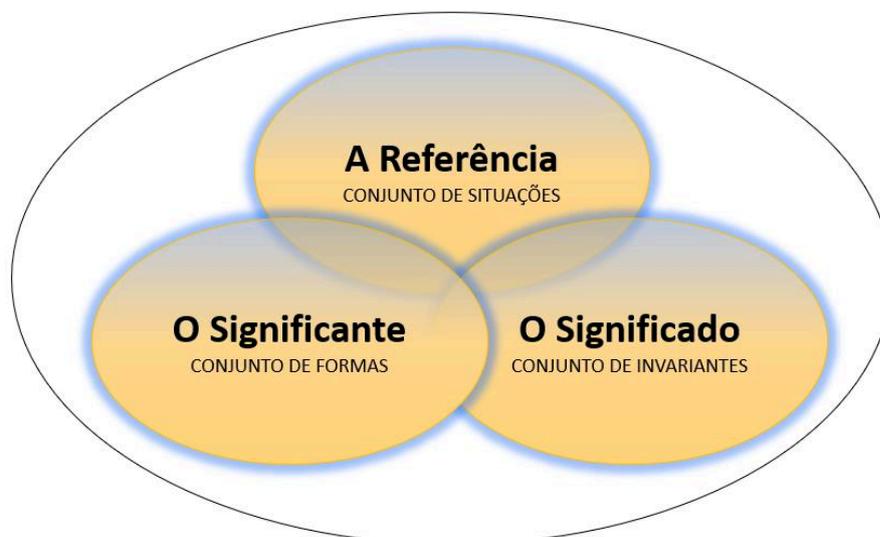
Figura 43 - Mapa conceitual da teoria dos campos conceituais



Fonte: Do autor (2018).

Como a teoria entende que a conceitualização é a base para o desenvolvimento cognitivo e que é por meio de situações e problemas a serem resolvidos que um conceito adquire sentido, devemos, então, explorar o que se entende por conceito na teoria dos campos conceituais. Segundo Vergnaud (1990a) o conceito é formado por um conjunto interligado de três elementos como mostra a Figura 44, a seguir:

Figura 44 - Conjunto do conceito

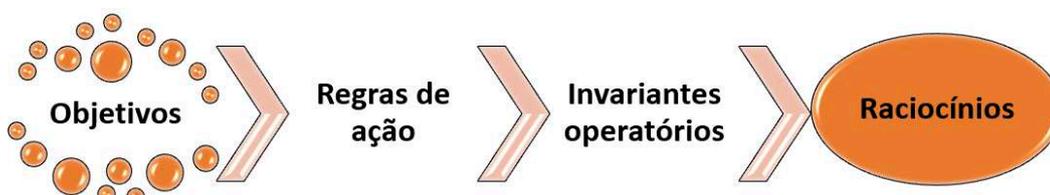


Fonte: Do autor (2018).

A referência é o conjunto de situações que dão sentido ao conceito. É por meio de uma variedade de situações que o sujeito compõe significado a um conceito. Postula-se aqui, o coração da teoria, uma vez que o foco da análise é o sujeito-em-situação, e o conjunto das situações é reconhecido como o referente do conceito. O significado é um conjunto de invariantes operatórios sobre os quais repousa a operacionalidade dos conceitos. Esses invariantes representam aquilo que se preserva nos conceitos e que permite serem reconhecidos como tais nas situações, em outras palavras, os invariantes concebem o significado do conceito. O significante é um conjunto de formas linguísticas e não linguísticas que permitem representar, simbolicamente o conceito e suas propriedades.

Ao buscarmos a expansão de diversas situações que envolvem um conceito, exige-se a construção de novos esquemas e de outros conhecimentos e conceitos conexos, dessa forma, estaremos contribuindo efetivamente para o desenvolvimento de um novo conceito. Para Vergnaud (1998), o conceito de esquema é muito frutífero, tanto para descrever comportamentos familiares, como para descrever e compreender os processos de resolução de problemas. O autor indica quatro elementos para a formação de um esquema, como mostra a Figura 45, a seguir:

Figura 45 - Elementos de um esquema



Fonte: Do autor (2018).

Vergnaud (1990a) esclarece que os objetivos ou metas devem estar sempre presentes, pois um esquema é orientado para a resolução de uma determinada classe de situações. As regras de ação buscam por informações e controles que são os elementos que dirigem a sequência de ações do sujeito-em-situação. Os invariantes operatórios (teoremas-em-ação e conceitos-em-ação) que dirigem o reconhecimento, por parte do sujeito-em-situação, dos elementos pertinentes à situação e, portanto, guiam a construção dos modelos mentais. A possibilidade de inferência, ou raciocínios, permitem determinar as regras e antecipações a partir das informações e dos invariantes operatórios dos quais dispõe o sujeito.

Diante desses elementos, Vergnaud (1996) pondera que nos invariantes operatórios está a base conceitual implícita que permite obter as informações necessárias para alcançar os objetivos e inferir as regras de ação mais pertinentes para a resolução de uma situação. Dessa

forma, é nos esquemas que devemos deter a observação para compreender a constituição do desenvolvimento cognitivo do sujeito, uma vez que é aí que podemos encontrar os elementos que fazem com que a ação seja operatória. O esquema é um referente do sujeito do conhecimento e a situação é a circunstância e o contexto em que o objeto a ele se apresenta.

A organização do pensamento, referente ao conceito de esquema, pode nos conduzir a uma análise dos conhecimentos que estão em ação no sujeito frente a uma determinada situação. Essa análise permite compor um quadro em que se verifica o desenvolvimento temporal dos modelos explicativos dos sujeitos, induzida a partir dos invariantes operatórios (conceito-em-ação e teoremas-em-ação) utilizados ao longo de uma atividade de ensino.

Vergnaud (1998) assinala que os conceitos-em-ação estão relacionados a objetos, predicados, classes, condições, etc. Diante de uma ampla quantidade de conceitos que podem estar disponíveis no repertório dos sujeitos, é selecionada uma pequena parte para cada ação. Dessa forma, os conceitos-em-ação podem ser adequados ou inadequados para uma classe de situações. Esses conceitos-em-ação permanecem, em sua maioria, implícitos ao longo da ação do sujeito. Sendo assim, as análises das respostas dadas pelos sujeitos-em-situação são de grande importância para averiguarmos se os conceitos utilizados estão de acordo com a situação proposta. Os conceitos-em-ação são articulados por meio dos teoremas-em-ação, que são proposições verdadeiras ou falsas e, de maneira análoga aos conceitos-em-ação, permanecem implícitas na maioria das vezes durante as ações dos sujeitos, podendo se tornar explícitas.

Sendo o sujeito-em-situação a peça principal na teoria, Vergnaud (1990) nos faz refletir sobre a importância de acompanharmos os estudantes enquanto aprendem, procurando nos conceitos e teoremas em ação, a evolução temporal de seu conhecimento. Essa tarefa exige uma imersão profunda no universo da pesquisa, fazendo com que o pesquisador fique atento ao contexto de cada enunciação feita pelo aluno.

Por se tratar de uma teoria complexa, onde diversos conceitos são levados em consideração para que o sujeito determine uma certa situação, a teoria dos campos conceituais admite que o professor possa pensar sobre seu objeto de ensino de uma forma mais global. Cabe salientar que essa teoria trabalha não apenas com conceitos já formalizados e consolidados pelo sujeito, mas também e sobretudo com conhecimentos em via de formalização. Esse aspecto garante que possamos identificar, nas situações-problema, os modos de compreensão dos alunos em processo de formação.

4.1. A estimação de parâmetros populacionais enquanto campo conceitual

Diante da breve explanação de alguns pontos fundamentais sobre a teoria dos campos conceituais, observamos em Oriol (2007) e Diaz (2016) demonstrações da aplicação dessa teoria no ensino da Estatística. Os autores citados utilizam a noção de média aritmética de uma série de dados, para observar a formação do conceito.

Figura 46 - Conjunto do conceito da média



Fonte: Do autor (2018).

A média é inegavelmente um conceito, pois se refere a um conjunto de situações (a referência), reúne também um conjunto de invariantes em que se baseia a operacionalidade do conceito (o significado) e, finalmente, a média admite um conjunto de formas linguísticas e não linguísticas que permitem representar simbolicamente suas propriedades, situações e procedimentos de tratamento (o significante).

Dessa forma, estabelecemos relações entre a teoria dos campos conceituais com a estimação de parâmetros, onde o conceito de estimação agrega várias situações passíveis de aplicabilidade (a referência), englobando um conjunto de invariantes que operacionalizam sua prática (o significado) e admite um conjunto de formas linguísticas e não linguísticas que representam simbolicamente suas propriedades e seus processos de tratamento (o significante).

Quadro 22 - Conjunto do conceito da estimação pontual da média populacional

Conceito: estimativa intervalar de uma média populacional		
Situações (A Referência)	Invariantes Operatórios (O significado)	Representações simbólicas (O signifiante)
Estimação pontual da média populacional μ a partir de dados amostrais n .	Medida de tendência central	$\bar{X} = \frac{x_1 + x_2 + \dots + x_n}{n}$
Descobrir por meio de uma amostra o tempo médio que uma criança leva para montar um brinquedo.	Soma de todos os elementos de uma amostra dividido pelo número total de elementos.	$\frac{1}{n} \sum_{i=1}^n x_i$
	Estimador \bar{X} do parâmetro μ .	

Fonte: Do autor (2018).

Quadro 23 - Conjunto do conceito da estimação intervalar da média populacional

Conceito: estimativa intervalar de uma média populacional		
Situações (A Referência)	Invariantes Operatórios (O significado)	Representações simbólicas (O signifiante)
Estimação intervalar da média populacional μ a partir de dados amostrais n .	Medida de tendência central.	$\bar{X} = \frac{x_1 + x_2 + \dots + x_n}{n}$
Descobrir por meio de uma amostra o tempo médio que uma criança leva para montar um brinquedo.	Medida de dispersão.	$\frac{1}{n} \sum_{i=1}^n x_i$
	Estimador \bar{X} do parâmetro μ .	$\left[\bar{X} \mp z_{\alpha} \frac{\sigma}{\sqrt{n}} \right]$
	Intervalo de confiança.	
	Distribuição de probabilidades.	
	Erro amostral.	

Fonte: Do autor (2018).

Salientamos que os quadros apresentados estão descrevendo conceitos científicos preparados para a mediação didática visando sua apropriação por parte dos estudantes, tais conceitos não podem ser considerados conceitos-em-ação (articulados por teoremas-em-ação), visto que não são categorias de entendimento de um sujeito em ação frente a uma dada situação.

Nossa intenção foi de contextualizar a teoria utilizando alguns possíveis conceitos que julgamos necessários a um dado sujeito para dar conta das situações.

Percebemos, então, que um conceito, aqui especificamente o de estimação, não é uma estrutura isolada, cabe em um campo conceitual que está relacionado a diversos outros conceitos, como o de amostra, média, desvio padrão, para citar apenas os mais comuns. Vergnaud (1990a) afirma que para o estudo de um conceito é necessário estar presente um campo de problemas distintos, mas que tenham uma base idêntica. Nossa pesquisa observou os estudantes na resolução de situações que convocaram os mesmos conceitos estatísticos. Dessa forma, apuramos como os conceitos estão em tensão entre si, a fim de formar um campo conceitual.

PARTE 2:

METODOLOGIA DA CONSTRUÇÃO E TRATAMENTO DOS DADOS

Descreveremos agora os métodos escolhidos para a construção e para o tratamento dos dados que promoveram reflexões para as nossas análises em relação a questão norteadora da pesquisa. A partir dos dados que foram gerados durante o desenvolvimento da tese, extraímos informações para o levantamento de hipóteses e interpretações que acreditamos serem colaborativas para a Educação Estatística e para a oferta de novas demandas em futuras pesquisas da área. Barbetta (2006) assinala que uma pesquisa científica é um processo iterativo da evolução do conhecimento, onde as hipóteses levantadas servirão de suporte para outras pesquisas e, assim, contribuirão na construção de novos saberes.

Figura 47 - Processo iterativo da construção do conhecimento



Fonte: Adaptado de Barbetta (2006).

Optamos em realizar uma pesquisa com metodologia mista, buscando explorar os dados coletados por meio de métodos qualitativos e quantitativos. De acordo com Gonsalves (2003), durante muitos anos as pesquisas qualitativas e quantitativas foram tratadas como paradigmas, onde as de cunho quantitativo transformaram os fenômenos investigados em números, fazendo o uso de ferramentas Estatísticas para explicar suas causas e testar suas hipóteses. Por outro lado, as pesquisas qualitativas preocuparam-se com a compreensão e com a interpretação dos fenômenos. Entendemos a necessidade de superar essa dicotomia, articulando técnicas quantitativas e qualitativas em nossa pesquisa para podermos acompanhar uma tendência atual de pesquisas que estão acontecendo nas mais variadas áreas do conhecimento. Régnier (2007) observa que não devemos desvincular o qualitativo do quantitativo ao longo de uma pesquisa. Essa complementaridade é fundamental, tendo em vista os vários e distintos desideratos da pesquisa, cujos propósitos não podem ser alcançados por uma única abordagem metodológica.

Para tanto, dispomos de um dispositivo pedagógico que deu conta de levantar informações que pudessem ser observadas qualitativa e quantitativamente. Esse dispositivo, tratado no item 2.5 da Parte 1, foi composto por dois questionários que trouxeram situações de

investigação abordando possíveis conhecimentos específicos da estimação estatística e possíveis relações de sentimento do estudante sobre a Estatística.

Nosso intuito foi de investigar a construção do raciocínio estatístico por meio da estimação estatística, explorando as respostas dadas nos questionários por um grupo de estudantes matriculados em uma disciplina de Estatística ofertada no Ensino Superior, antes e depois de passarem por uma formação sobre o conteúdo de estimação estatística. Dessa forma, os mesmos questionários foram aplicados duas vezes: a primeira aplicação ocorreu antes da formação e, a segunda, depois da formação.

Com a finalidade de obtermos subsídios para inferirmos sobre as informações que foram coletadas na aplicação dos questionários e tratadas por meio de softwares estatísticos, observamos, em nosso referencial teórico, as eminentes reflexões da Educação Estatística, como o letramento, o pensamento e o raciocínio estatísticos. Além disso, buscamos na Teoria dos Campos Conceituais de Gérard Vergnaud os argumentos necessários para observar a Estimação Estatística enquanto campo conceitual e melhor entender o processo de construção desse conceito. A seguir, descrevemos as escolhas que nos guiaram ao longo da pesquisa.

1. Contexto e sujeitos da pesquisa

A coleta dos dados, para a nossa pesquisa, ocorreu nos dois semestres letivos do ano de 2016 no Centro Universitário Franciscano, atualmente Universidade Franciscana - UFN, instituição de natureza privada localizada na cidade de Santa Maria, interior do Estado do Rio Grande do Sul. A instituição oferta trinta e quatro cursos de graduação, vinte e dois cursos de especialização, oito residências, seis mestrados e dois doutorados, atendendo diversas áreas do conhecimento.

Nossa pesquisa investigou uma amostra de oitenta e um estudantes oriundos dos cursos de graduação das áreas de Ciências Sociais⁴¹ e Ciências Tecnológicas⁴², que estavam cursando a disciplina de Estatística que possui caráter obrigatório e estava sendo ofertada no 3º semestre de acordo com a grade curricular dos cursos. Além disso, a escolha dos sujeitos da pesquisa se deveu ao fato de que o autor da pesquisa foi o professor regente da disciplina.

⁴¹ Administração, Ciências Contábeis e Ciências Econômicas.

⁴² Engenharia Ambiental e Sanitária, Engenharia Química e Matemática.

Gráfico 5 - Número de estudantes nos cursos de graduação participantes da pesquisa



Fonte: Do autor (2018).

Algumas turmas investigadas na pesquisa eram compostas por mais de um curso de graduação, dessa forma, foram examinadas quatro turmas, sendo, a primeira, com trinta e dois alunos do curso de Administração; a segunda com vinte e nove alunos do curso de Ciências Contábeis e seis alunos do curso de Ciências Econômicas; a terceira com cinco alunos do curso de Licenciatura em Matemática e, a quarta, com seis alunos do curso de Engenharia Química e três alunos do curso de Engenharia Ambiental e Sanitária. A identificação dos estudantes foi realizada como mostra a Figura 48.

Figura 48 - Identificação dos estudantes

Grupo 1 – ADMINISTRAÇÃO – {E01_001, E01_002, E01_003, ..., E01_032}
Grupo 2 – MATEMÁTICA – {E02_001, E02_002, E02_003, E02_004, E02_005}
Grupo 3 – CIÊNCIAS CONTÁBEIS – {E03_001, E03_002, E03_003, ..., E03_029}
Grupo 4 – CIÊNCIAS ECONÔMICAS – {E04_001, E04_002, E04_003, ..., E04_006}
Grupo 5 – ENGENHARIA QUÍMICA – {E05_001, E05_002, E05_003, ..., E05_006}
Grupo 6 – ENGENHARIA AMBIENTAL E SANITÁRIA – {E06_001, E06_002, E06_003}

Fonte: Do autor (2018).

Observamos que a faixa etária dos estudantes variou dos 17 aos 45 anos, com uma média geral de 24 anos e um desvio padrão de 5 anos. O gênero dos estudantes foi praticamente igualitário, sendo quarenta e um estudantes do sexo feminino e quarenta estudantes do sexo masculino.

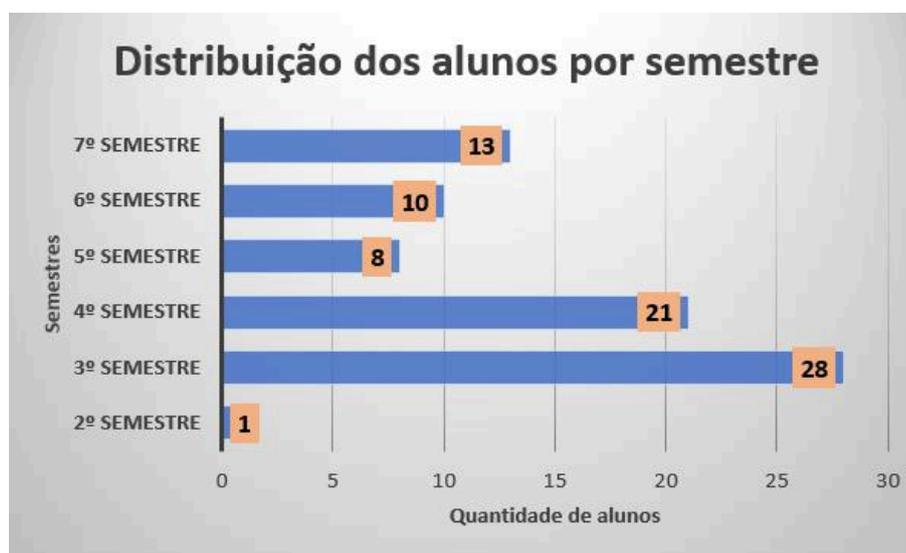
Gráfico 6 - Gênero dos estudantes participantes da pesquisa



Fonte: Do autor (2018).

A grade curricular dos cursos envolvidos na pesquisa oferecia a disciplina de Estatística a partir do 3º semestre do curso, apesar disso, houve uma heterogeneidade entre os estudantes em relação aos semestres, tendo em vista que muitos optam em não realizar a disciplina nos primeiros semestres do curso ou por terem sido reprovados nessa disciplina.

Gráfico 7 - Distribuição dos alunos por semestre



Fonte: Do autor (2018).

Por se tratar de áreas que formam usuários de Estatística, muitos pesquisadores, como Carzola (2002), Batanero (2001) e outros, entendem que a Educação Estatística deve possuir um olhar atento na formação acadêmica desses profissionais que nem sempre é satisfatória no que diz respeito ao letramento estatístico. Uma das premissas apontadas pelos pesquisadores da

área é a da importância de compreender o processo de construção do raciocínio e dos conceitos estatísticos para contribuir, de forma relevante, na aprendizagem da Estatística durante a formação acadêmica desses profissionais. Sendo assim, nossos sujeitos de pesquisa estão contextualizados dentro da Educação Estatística como um público que possui informações pertinentes que podem colaborar com a melhoria do ensino.

2. Ferramentas computacionais

Diante da posse dos questionários já respondidos pelos estudantes, reunimos todos os dados em planilhas eletrônicas do software EXCEL de modo a não negligenciar nenhuma informação que foi produzida. Esses dados brutos permitiram que pudéssemos consultá-los a qualquer instante de modo rápido a fim de verificar alguma informação que desejássemos analisar. Além disso, justifica-se o uso do software Excel, por ele servir como base de dados para a exploração dos softwares CHIC e SPAD que foram utilizados em nossa pesquisa e que serão apresentados a seguir.

2.1. Software CHIC

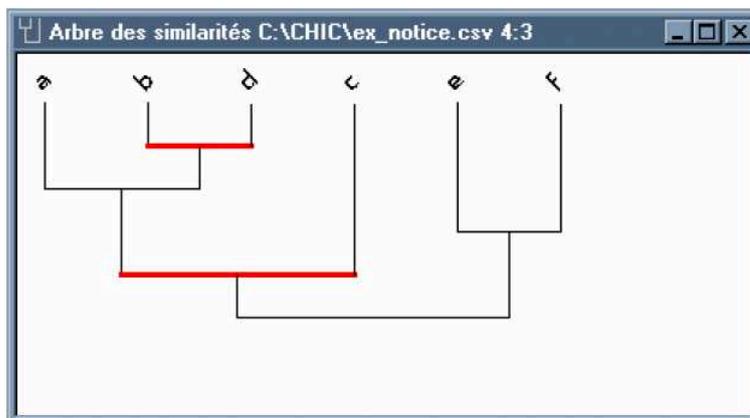
Para analisar alguns dados que foram levantados dos questionários, utilizamos a versão 7.0 do software CHIC – Classificação Hierárquica, Implicativa e Coesiva, com o algoritmo originalmente concebido por Régis Gras na década de 1980 e aperfeiçoado ao longo dos anos por Saddo Ag Almouloud, Harrison Ratsimba-Rajon e, atualmente, sob a responsabilidade de Raphaël Couturier. O algoritmo do software recai para análise de dados baseados em pesquisas de Análise Estatística Implicativa – ASI, que é um método de análise e de classificação de dados multidimensionais, um método quantitativo desenvolvido no âmbito da Didática da Matemática francesa a partir de situações didáticas.

De acordo com Régnier *et al.* (2014), Gras *et al.* (2009, 2013, 2017) o quadro teórico da Análise Estatística Implicativa tem como objetivo principal a estruturação de dados no cruzamento de indivíduos e variáveis, a partir da contingência de regras e determinando os conceitos de intensidade e envolvimento, a coesão de classe, o significado dos níveis hierárquicos, a contribuição de adicionais etc. Essa análise permite extrair, de um conjunto de dados categóricos, relações entre sujeitos e variáveis, além de associações entre variáveis por meio de um índice de qualidade dessa associação. Analogamente, o tratamento de variáveis binárias é adicionado ao de variáveis modais, frequenciais e intervalo.

Segundo Gras e Almouloud (2005) a ASI possibilita não somente classificar elementos de um banco de dados como é feito pelos métodos de Análise Multivariada de Dados, mas, também, determinar o grau de implicação por meio de um índice, isto é, em que medida uma variável a segue uma variável b , em outras palavras, de acordo com Régnier *et al.* (2012), quando uma variável a é escolhida, temos uma tendência a escolher a variável b . A implicação é uma relação lógica pela qual uma coisa se conclui de outra, ou de uma forma mais rigorosa, é toda subjunção logicamente verdadeira, isto é, uma subjunção ($a \rightarrow b$) na qual a cada coleção de a e b corresponde ao valor verdade.

Em conformidade com o artigo de Couturier (2001) o software CHIC conta com uma série de funções, entre elas a de calcular parâmetros da estatística básica, como a média, o desvio padrão e o coeficiente de correlação. Outra função é a construção de uma árvore de similaridade, como mostra a Figura 49, baseada na teoria de Lerman (1981), onde são empregados índices de probabilidade.

Figura 49 - Exemplo de uma árvore de similaridade



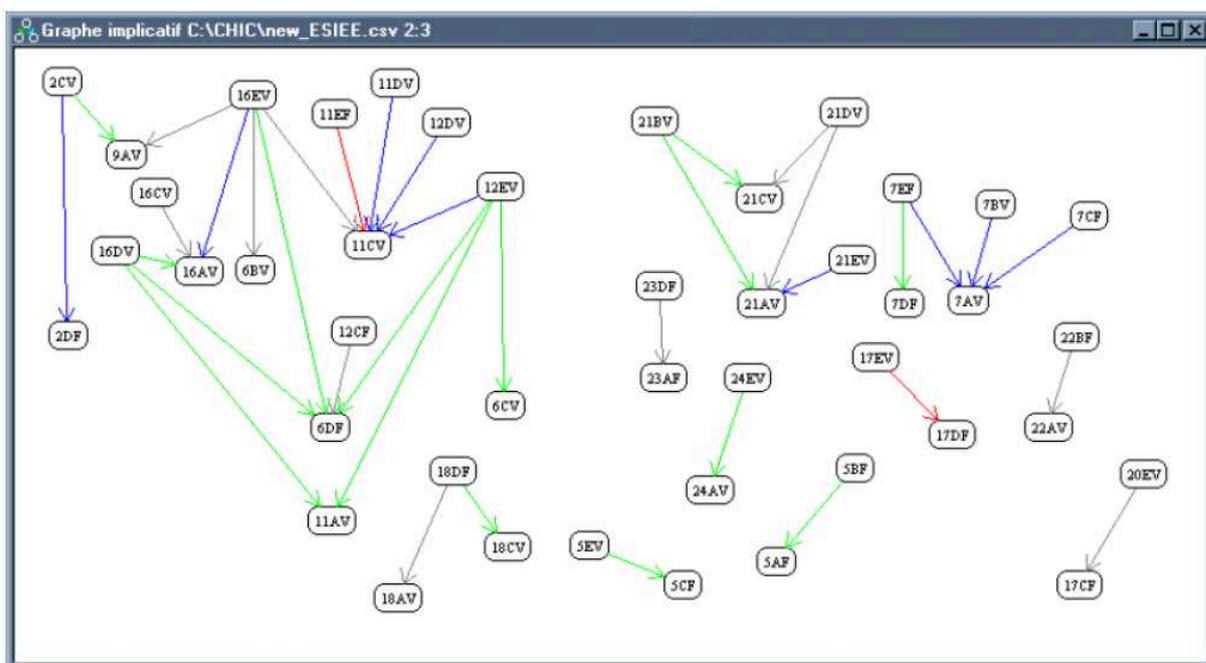
Fonte: Couturier (2001).

Couturier (2001) explica que as variáveis b e d são, a princípio, as variáveis mais similares. Então, o algoritmo escolhe associar a variável a com a classe (b, d) , formando uma nova classe $(a, (b, d))$. No terceiro nível é formada a classe (e, f) , no quarto nível existe uma associação da variável c com a classe $(a, (b, d))$ e, finalmente, no último nível temos uma única classe, onde (a, b, d, c) e (e, f) são distintas. Como podemos observar, as variáveis são agrupadas duas a duas de acordo com o nível de semelhança e os pares que apresentam um maior “significado” são destacados por uma linha vermelha em negrito.

O software CHIC permite o cálculo da intensidade da implicação entre as variáveis que, segundo a definição de Gras *et al.* (2002), de uma variável a para uma variável b a medida $j(a,$

$b) = 1 - Pr(Q(a, b) \notin q(a, b))$ para $nb \neq n e$, a quase-implicação a fi b é admissível ao nível de confiança $1 - \alpha$ se e somente se $j(a, b) \geq 1 - \alpha$. Essa intensidade entre os pares de variáveis é representada por um grafo implicativo que mostra os resultados obtidos por meio de flechas coloridas. As cores das flechas são determinadas de acordo com os quatro limites diferentes que o software disponibiliza e que poderão ser escolhidos antes dos cálculos. Por convenção, é comumente aceito uma ligação estatisticamente significativa no limiar de risco $\alpha = 0,05$, ou seja, assume-se como margem de segurança 5% de chances de erro, ou, olhando por outro ângulo, 95% de chances de estar certo.

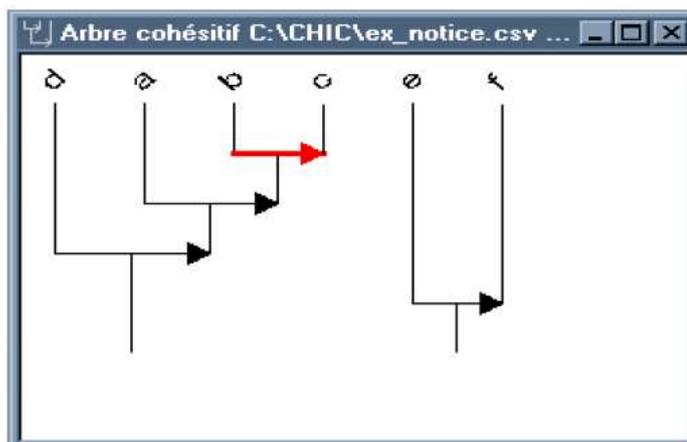
Figura 50 - Exemplo de um grafo implicativo



Fonte: Couturier (2001).

O software também permite o cálculo da coesão entre as classes construindo uma árvore hierárquica coesiva, que possui semelhanças com a árvore de similaridade. Entre as semelhanças, os dois gráficos são lidos horizontalmente, mostrando as linhas com correspondências mais acentuadas e destacadas por flechas vermelhas mais grossas. Podemos, também, analisar a árvore coesiva por meio da leitura vertical.

Figura 51 - Exemplo de árvore coesiva



Fonte: Couturier (2001).

A árvore coesiva permite revelar a hierarquia descendente entre as classes coesas por meio de um algoritmo que reúne variáveis de acordo com uma classificação decrescente. Gras (2004) esclarece que a árvore possibilita identificar os níveis mais relevantes e é interrompida onde a ligação é muito fraca. Esse arranjo resulta em linhas contínuas que terminam com uma seta indicativa da direção da conexão.

A Figura 51 mostra os mesmos dados da Figura 49, onde, dessa vez, é a classe (b, c) que aparece no primeiro nível de hierarquia, possuindo uma linha vermelha que indica uma ligação significativa. Couturier (2001) esclarece que a variável b implica na variável c com maior intensidade em relação aos outros pares de variáveis. Em seguida, temos o segundo nível $(a, (b, c))$ à qual a variável d se une no terceiro nível para formar um novo grupo $(d, (a, (b, c)))$. Finalmente é gerado um quarto e último nível formado pelo par (e, f) . As duas últimas classes não estão conectadas entre si, formando, então, duas associações distintas $(d, (a, (b, c)))$ e (e, f) . Acioly-Régner e Régner (2005) informam que apesar dessa aparente simplicidade da leitura da árvore coesa, dificuldades aparecem tanto na leitura vertical como na horizontal, devendo o pesquisador fazer a interpretação necessária para a sua investigação.

2.2. Software SPAD

O software SPAD – Système Portable pour l'Analyse des Données foi desenvolvido pela empresa Coheris⁴³ e classificado dentro da família de softwares estatísticos como o SAS (desenvolvido pelo SAS Institute⁴⁴), o SPSS (distribuído pela IBM⁴⁵), XLSTAT (desenvolvido

⁴³ Disponível em: www.coheris.com/produits/analytics/logiciel-data-mining/.

⁴⁴ Disponível em: www.sas.com/pt_br/home.html.

⁴⁵ Disponível em: www.ibm.com/br-pt/products/spss-statistics.

pela Addinsoft⁴⁶), Tanagra e R (plataformas de experimento de código aberto). Essas citações não abrangeram toda a família de softwares estatísticos, mas contemplam os programas mais conhecidos que permitem a exploração de grandes volumes de dados. Tufféry (2010) salienta que diante dessa oferta abundante, o SPAD se destaca dos concorrentes por sua importante difusão no mundo universitário francês, devido a qualidade de seus algoritmos para o processamento de dados.

O software implementa uma série de métodos, tais como análise descritiva e exploratória dos dados: classificação plana, tabulação cruzada, análise bivariada, etc. Permite, em particular, produzir a análise em componentes principais (ACP) e análise de correspondência simples (AFC), classificando sujeitos a partir das modalidades escolhidas. Esses dois métodos são usados para descrever a integridade das informações, resumir e visualizar todas as inter-relações existentes, fornecendo, assim, uma estrutura sinóptica da população estudada. Além disso, é um dos pioneiros em análise quantitativa de dados textuais, possuindo muitas outras funções que abrangem inúmeros processos estatísticos, permitindo que um grande volume de informações seja processado.

Figura 52 - Interface do software SPAD 9.1

Index	Etudiants	[aV01] Sexo	[aV02] Idade	[aV03] Voc...	[aV04] Qua...	[aV05] Sem...	[aV06Ta] Ho...
1	E_001	1	23.619	3_Alguas ...	3_Alguas ...	3	Matemática
2	E_002	2	24.731	4_Frequent...	5_Muito Fre...	5	Números
3	E_003	1	25.520	5_Muito Fre...	5_Muito Fre...	3	Matemática
4	E_004	2	22.653	2_Raramente	3_Alguas ...	4	Gráficos

Fonte: Do autor (2018).

⁴⁶ Disponível em: www.xlstat.com/en/.

Utilizamos a versão SPAD 9.1 que apresenta uma interface estética intuitiva e de fácil utilização. Essa versão não é gratuita e nossa licença de uso foi cedida pelo professor Doutor Jean-Claude Régnier que incentiva e ensina o manuseio do software para pesquisas científicas em seus seminários para mestrandos e doutorandos na Universidade de Lyon 2. A interface do software é dividida em quatro partes como mostra a Figura 52: 1 – é o relatório de todos os arquivos de projetos que já foram executados; 2 – mostra o projeto que está sendo executado e estruturado de acordo com uma arquitetura que agrupa elementos que compuseram o processamento dos dados a partir de uma base (em nossa pesquisa usamos a base de dados elaborada no Excel); 3 – permite visualizar todos os métodos estatísticos contidos no software (importação e exportação de arquivos, estatística descritiva, análise fatorial, etc.); 4 – resultados dos processos executados.

3. Construção dos dados: questionários da pesquisa

A coleta das informações, que nos encaminharam para as reflexões acerca do propósito da pesquisa, contou com a aplicação de dois questionários. O uso de questionários, segundo Gil (1999), pode ser definido como uma técnica de investigação composta por um número mais ou menos elevado de questões apresentadas por escrito às pessoas, tendo por objetivo o conhecimento de opiniões, crenças, sentimentos, interesses, expectativas, situações vivenciadas etc. Para Gray (2012), os questionários são uma das técnicas mais usadas de coleta de dados, permitindo uma abordagem analítica que explora as relações entre as variáveis. Gil (1999) aponta que a preservação do anonimato e a não influência do pesquisador sobre as respostas dos respondentes, estão entre as vantagens do uso dessa técnica. O mesmo autor ainda indica que existem pontos negativos, entre eles é que a técnica não oferece garantia de que as pessoas irão devolvê-lo devidamente preenchido, o que pode implicar a significativa diminuição da representatividade da amostra. A construção de um questionário, segundo Aaker (2001), é considerada uma “arte imperfeita”, pois não existem procedimentos exatos que garantam que seus objetivos de medição sejam alcançados com boa qualidade. No entanto, o pesquisador deve procurar compor um questionário que atenda suas expectativas, observando o contexto e os sujeitos da pesquisa.

A aplicação dos nossos questionários, como já mencionado, foi durante o desenvolvimento de uma disciplina de Estatística no Ensino Superior. Para que pudessemos melhor entender a formação do raciocínio estatístico por meio da estimação estatística, ambos os questionários foram entregues aos estudantes, matriculados na disciplina, antes da exposição

do conteúdo de estimação e uma segunda aplicação dos mesmos questionários depois que os estudantes receberam o conteúdo por meio de aulas expositivas. Quatro turmas foram observadas, alcançando seis cursos de graduação, perfazendo um total de 24 horas-aula (6 horas-aula em cada turma) entre as aplicações e o desenvolvimento do conteúdo. Vale salientar que todos os estudantes já haviam sido expostos aos conteúdos que contemplaram a estatística descritiva, de acordo com programa da disciplina. Para o conteúdo de estimação estatística, foram abordados a estimação de parâmetros para a média e para a proporção.

3.1. Questionário Q1

Intitulado de Q1, o primeiro questionário aplicado foi uma tradução e adaptação de um questionário que é empregado pelo professor Doutor Jean-Claude Régnier em suas inúmeras pesquisas realizadas na França. Régnier (2003) observa que esse questionário pode servir de apoio para o estudo das dificuldades e das facilidades na aprendizagem da Estatística. Ajustado para o contexto da nossa pesquisa, o questionário contou com dois segmentos: 1 – Informações sobre a formação do estudante; 2 – Relação do estudante com a Estatística.

Para cada item que compôs os segmentos, usamos um código de identificação que serviu para o reconhecimento da questão durante o tratamento que foi realizado utilizando o software SPAD. Dessa forma, como aplicamos o mesmo questionário duas vezes, em todos os itens analisados, para a nossa diferenciação durante as análises, utilizamos a letra “a” para identificar o “antes” e a letra “d” para identificar o “depois” na frente de cada código. Sendo assim, para cada item teremos duas variáveis a serem analisadas, como por exemplo: **aV03** (identificando a respostas antes) e **dV03** (identificando a resposta depois).

Expomos agora o segmento que tratou sobre as informações relativas a formação do estudante, como mostra a Figura 53.

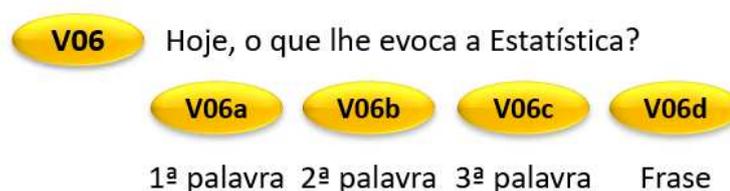
Figura 53 - Q1: informações sobre a formação do estudante

- V01** Sexo (Masculino/Feminino)
- V02** Data de nascimento
- V03** Você já enfrentou situações que exigiram conhecimentos em Estatística?
- V04** Quando os conhecimentos em Estatística são requisitados, isso lhe causa problemas?
- V05** Semestre

As variáveis V01 e V05, tratadas como variáveis nominais, identificaram, respectivamente, o sexo e o semestre dos estudantes. A variável V02, tratada como variável contínua, identificou a faixa etária dos estudantes. As variáveis nominais V03 e V04 foram respondidas de acordo com uma escala composta por seis opções: 1) Nunca; 2) Raramente; 3) Algumas vezes; 4) Frequentemente; 5) Muito frequentemente e 6) Quase todos os dias. Em V03 tivemos a possibilidade de averiguar se o estudante possui enfrentamentos estatísticos no seu cotidiano, tendo em vista que muitos estudantes que participaram da pesquisa, trabalham no turno inverso da universidade. Dessa forma, podemos entender um pouco sobre a atmosfera dos estudantes e chamar a atenção sobre a importância do letramento estatístico que está inserido nos mais variados ramos e atividades da vida cotidiana. A variável V04 sinalizou a existência de possíveis dificuldades que os estudantes encontram ao enfrentarem situações onde uma leitura estatística se faz necessária.

Dando continuidade ao questionário Q1, o segmento que tratou da relação dos estudantes com a Estatística foi composto por 11 itens que serão descritos a seguir.

Figura 54 - Q1: V06



Fonte: Do autor (2018).

A fim de observarmos as possíveis associações que os estudantes possuem em relação a Estatística, a variável V06, tratada e analisada por meio de quatro variáveis textuais (V06a, V06b, V06c e V06d), solicitou ao estudante a escrita de três palavras e uma frase que remetesse a Estatística. As variáveis textuais serão tratadas no SPAD por meio da elaboração de um vocabulário contendo as palavras encontradas com maior frequência e expostas pelo método “Nuage de mots”, ou seja, a construção de nuvens de palavras que mostra o grau de frequência das palavras em um texto. Quanto mais a palavra é utilizada, mais chamativa é a representação dessa palavra no gráfico. As palavras aparecem em fontes de vários tamanhos e em diferentes cores, indicando o que é mais relevante e o que é menos relevante no contexto.

Figura 55 - Q1: V07 e V08

- V07** Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação?
- V08** Qual nível de utilidade você atribui para uma formação em Estatística na sua graduação?

Fonte: Do autor (2018).

As variáveis contínuas V07 e V08 foram respondidas por meio de um nível de importância de zero a dez. Em V07, o estudante expôs o interesse pessoal em relação a Estatística, exprimindo um sentimento que atinge diretamente a pessoa em questão. Na V08 o estudante manifestou um sentimento de utilidade, indicando se os conhecimentos estatísticos possuem relevância. Entendemos que o estudante pode enxergar inúmeras utilidades advindas das ferramentas estatísticas e, ao mesmo tempo, não ter interesse pessoal em dominá-las.

A variável V09 solicitou que os estudantes escrevessem o entendimento pessoal sobre alguns termos empregados na Estatística. Dessa forma, foram analisadas seis variáveis textuais como mostra a Figura 56.

Figura 56 - Q1: V09

- V09** Atualmente o que significa para você os termos a seguir?
- V09a** Representativo
 - V09b** Significativo
 - V09c** Estimação
 - V09d** Teste de Hipóteses
 - V09e** Probabilidade
 - V09e** Sondagem

Fonte: Do autor (2018).

O termo representativo (V09a) aparece no sentido baseado em representação, onde algo responde por um todo. Esse termo pode ser encontrado na Estatística quando estivermos nos referindo a amostra, como por exemplo, uma amostra corresponde a um grupo representativo da população. O termo significativo (V09b) desponta um sentido de algo que contém alguma

informação expressiva. Esse termo pode ser encontrado no cálculo do p-valor⁴⁷, como por exemplo, o resultado do p-valor é significativo. Em V09c é solicitado a compreensão do que é estimação, que podemos definir como o processo de mensurar hipoteticamente algum valor, ou, ainda, um processo que consiste em utilizar dados amostrais para estimar parâmetros populacionais desconhecidos. Essa variável, ainda, nos ajudará a perceber a evolução dos estudantes em relação a compreensão do conceito de estimação. Teste de Hipóteses (V09d) é uma metodologia estatística que auxilia na tomada de decisões sobre uma ou mais populações baseadas nas informações obtidas da amostra. Probabilidade (V09e) é o estudo sobre experimentos que, mesmo realizados em condições bastante parecidas, apresentam resultados que não são possíveis de prever. A última variável, V09e, apresenta o termo sondagem, que é um estudo estatístico de uma população, feito por meio de uma amostra, destinada a estudar uma ou mais características tal como elas se apresentam nessa população.

Figura 57 - Q1: V10 e V11

- V10** Atualmente o que significa para você uma sondagem de opinião?
- V11** Nos dias de hoje, qual a utilidade você atribui para a Estatística?

Fonte: Do autor (2018).

As variáveis V10 e V11, tratadas como variáveis textuais, foram perguntas abertas que evocaram, respectivamente, o significado de sondagem de opinião e sobre a utilidade da Estatística. Em relação a sondagem de opinião, esperava-se que os estudantes escrevessem sobre a coleta de uma amostra para medir opiniões. Quanto a variável V11, esperava-se observar as associações que os estudantes percebem para as ferramentas estatísticas.

Figura 58 - Q1: V12

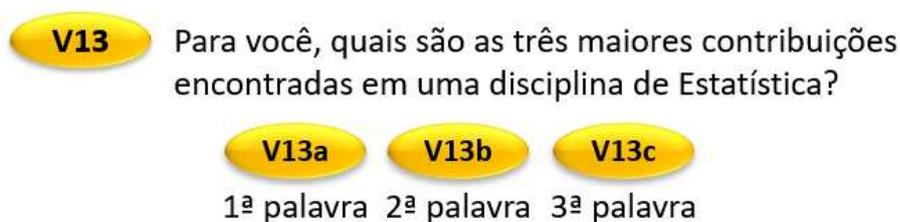
- V12** Para você, quais são as três maiores dificuldades encontradas em uma disciplina de Estatística?
- V12a** 1ª palavra **V12b** 2ª palavra **V12c** 3ª palavra

Fonte: Do autor (2018).

⁴⁷ É a probabilidade de se obter uma estatística de teste igual ou mais extrema que a estatística observada a partir de uma amostra de uma população.

A fim de observarmos as possíveis dificuldades encontradas em uma disciplina de Estatística, a variável V12, tratada e analisada por meio de três variáveis textuais (V12a, V12b e V12c), solicitou a escrita de três palavras que os estudantes julgaram ser os obstáculos para a aprendizagem.

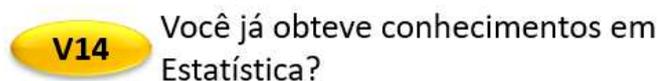
Figura 59 - Q1: V13



Fonte: Do autor (2018).

A variável V13, tratada e analisada por meio de três variáveis textuais (V13a, V13b e V13c), solicitou aos estudantes a escrita de três palavras que consideram ser relevantes quanto a aprendizagem da Estatística. Dessa forma, poderemos observar quais as associações que os estudantes estabelecem para uma formação estatística.

Figura 60 - Q1: V14



Fonte: Do autor (2018).

A variável nominal V14 tratou sobre o nível de conhecimento dos estudantes em relação a Estatística e foi respondida de acordo com uma escala composta por seis opções: 1) Nunca; 2) Em um nível muito superficial; 3) Em um nível superficial; 4) Introdução ao básico; 5) Introdução ao avançado e 6) Eu tenho uma formação em Estatística ou quase. Nosso contexto está inserido em uma disciplina de Estatístico no Ensino Superior, dessa forma, acreditamos que grande parte dos estudantes já haviam recebido instruções elementares de estatística no ensino básico. Cabe salientar, novamente, que todos os estudantes da pesquisa já haviam passado por uma formação básica sobre a estatística descritiva, ocorrida nessa mesma disciplina.

Figura 61 - Q1: V15

V15 Atualmente qual o seu sentimento em relação a Estatística?

v15a	1) Eu não tenho um bom conhecimento em Estatística.	1	2	3	4
v15b	2) A Estatística é algo que aprecio muito.	1	2	3	4
v15c	3) Eu me sinto inseguro quando preciso utilizar a Estatística.	1	2	3	4
v15d	4) Eu tenho uma relação muito boa com a Estatística: eu gosto e aprecio essa disciplina.	1	2	3	4
v15e	5) Quando estudo Estatística, não consigo pensar claramente pois minha cabeça fica bloqueada.	1	2	3	4
v15f	6) Quando eu escuto a palavra Estatística eu sinto um sentimento de aversão.	1	2	3	4
v15g	7) Eu realmente não gosto de matemática.	1	2	3	4
v15h	8) Pensar em resolver um problema de Estatística me deixa nervoso.	1	2	3	4
v15i	9) A Estatística é fascinante e divertida.	1	2	3	4
v15j	10) Eu nunca gostei de Estatística, ela é a disciplina que me causa mais medo.	1	2	3	4
v15k	11) O sentimento que experimento em relação a Estatística é bom.	1	2	3	4
v15l	12) Eu tenho medo de Estatística pois tenho medo de Matemática.	1	2	3	4
v15m	13) Eu acho a Estatística muito interessante e gosto de ir na aula dessa disciplina.	1	2	3	4
v15n	14) A Estatística me dá segurança e me estimula.	1	2	3	4
v15o	15) A Estatística me dá a impressão que estou perdido em uma sala cheia de números na qual eu não acho saída.	1	2	3	4
v15p	16) Eu me sinto bem com a Estatística e gosto muito dessa disciplina.	1	2	3	4
v15q	17) Eu estou sempre estressado durante as aulas de Estatística.	1	2	3	4
v15r	18) A Estatística é uma das matérias que eu realmente gosto de estudar na universidade ou na escola.	1	2	3	4
v15s	19) Eu não gosto de Estatística e nem de ir nas aulas de Estatística.	1	2	3	4
v15t	20) A Estatística me provoca um sentimento de indecisão que resulta no meu medo de não ser capaz de ter êxito na disciplina.	1	2	3	4
v15u	21) Eu me sinto mais feliz com a Estatística do que com as outras disciplinas.	1	2	3	4
v15v	22) A Estatística me deixa inquieto, descontente, irritado e impaciente.	1	2	3	4
v15w	23) Eu sou competente em Matemática.	1	2	3	4
v15x	24) Eu realmente gosto de Estatística.	1	2	3	4

Fonte: Do autor (2018).

A variável V15 foi observada por meio de vinte e quatro variáveis (V15a, V15b, V15c, ..., V15x) que expuseram atitudes positivas e negativas em relação a Estatística e a Matemática. Foi solicitado que os estudantes assinalassem o grau de adesão conforme uma escala composta por quatro opções: 1) Absolutamente contra; 2) Não estou bem de acordo; 3) Estou de acordo e 4) Concordo plenamente. Em V15, poderemos encontrar indícios de que as atitudes podem comprometer na formação do raciocínio estatístico. Muitos estudos como os de Gal e Ginsburg (1994), Sgoutas-Emch e Johnson (1998), Gonzalez (2002) e outros, apontam que analisar componentes afetivos é fundamental para o estudo de qualquer disciplina, pois podem indicar as causas de certos comprometimentos no desenvolvimento cognitivo dos estudantes. Entendemos que a afetividade representa um processo interno que engloba outros processos (psicológicos) como são as da motivação e das atitudes. Esses processos exercem um papel

importante na cognição dos estudantes, onde as atitudes têm uma forte carga afetiva na sua manifestação, uma vez que é uma resposta dos estudantes a partir de experiências e aprendizados com o objeto da atitude, em nosso caso, a Estatística.

Entre as vinte e quatro variáveis que compuseram V15, treze expressaram atitudes negativas (V15a, V15c, V15e, V15f, V15g, V15h, V15j, V15l, V15o, V15q, V15s, V15t e V15v) e onze atitudes positivas (V15b, V15d, V15i, V15k, V15m, V15n, V15p, V15r, V15u, V15w e V15x). Assim sendo, as repostas foram observadas de acordo com a atitude envolvida por meio de uma expressão que padronizou o grau de adesão escolhido pelos estudantes e de onde extraímos o que foi chamado de índice de positividade (Anexo D).

Figura 62 - Índice de positividade

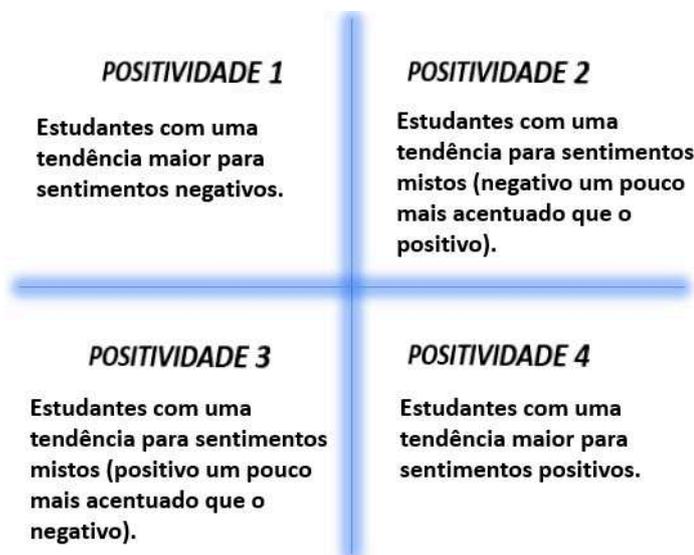
$$\text{índice de positividade} = \frac{65 - (\sum \text{ grau atitudes negativas}) + (\sum \text{ grau atitudes positivas}) - 24}{72}$$



Fonte: Do autor (2018).

Descobertos todos os índices de positividade (antes e depois), utilizamos, de forma arbitrária, os quartis para estabelecer quatro zonas de comportamento como mostra a Figura 63. Dessa forma, construímos novas variáveis para a realização das análises (a_Positividade_1, d_Positividade_1, a_Positividade_2, d_Positividade_2, a_Positividade_3, d_Positividade_3, a_Positividade_4 e d_Positividade_4).

Figura 63 - Zonas de comportamento em relação aos índices de positividade



Fonte: Do autor (2018).

A primeira zona, chamada de “Positividade 1”, com índices menores ou iguais ao primeiro quartil, são os estudantes que apresentaram uma tendência maior para sentimentos negativos em relação as respostas dadas (a_Positividade_1; d_Positividade_1). A segunda zona, chamada de “Positividade 2”, com índices maiores que o primeiro quartil e menores ou iguais ao segundo quartil, são os estudantes que apresentaram uma tendência para sentimentos mistos, mostrando uma intensidade maior para atitudes negativas (a_Positividade_2; d_Positividade_2). A terceira zona, “Positividade 3”, com índices maiores que o segundo quartil e menores ou iguais ao terceiro quartil, são estudantes que apresentaram uma tendência para sentimentos mistos, manifestando uma intensidade maior para atitudes positivas (a_Positividade_3; d_Positividade_3). A última zona, “Positividade 4”, com índices maiores que o terceiro quartil, são os estudantes que apresentaram uma tendência maior para sentimentos positivos em relação as respostas dadas (a_Positividade_4; d_Positividade_4). Essas zonas de comportamento nos ajudaram nas interpretações durante as análises para verificarmos possíveis relações das atitudes com as repostas das demais questões da pesquisa.

Figura 64 - Q1: V16

V16 Atualmente, qual sua definição de Estatística?

Fonte: Do autor (2018).

O último item, V16, tratada como variável textual, é uma pergunta aberta onde foi solicitado como os estudantes definem, atualmente, a Estatística.

Diante dos dados de Q1, todas as informações obtidas foram tratadas descritivamente por meio do software SPAD, e realizados alguns cruzamentos que foram pertinentes para as análises em resposta às hipóteses da pesquisa.

3.2. Questionário Q2

O segundo questionário, Q2, trouxe algumas situações-problema que entendemos serem propícias para o reconhecimento das noções que os estudantes já carregam sobre os procedimentos estatísticos e para a introdução do conteúdo que essa pesquisa propõe, a estimação estatística. As situações propostas não possuem associação com nenhum conhecimento específico dos cursos de graduação pesquisados, onde, de acordo com Gal

(2002), uma das principais características da Estatística é sua ancoragem em situações do cotidiano.

Com a finalidade de observar a construção de um campo conceitual em relação a estimação, as situações propostas evocaram os mesmos conceitos estatísticos apoiados na afirmação de Vergnaud (1990b) que entende o estudo de um conceito por meio de um campo de problemas distintos, mas com uma base idêntica. A primeira e a segunda situação-problema do questionário procuraram observar nos estudantes os componentes cognitivos, por meio da mobilização dos invariantes operatórios (conceitos-em-ação e teoremas-em-ação) advindos da estatística descritiva, como a noção de amostra, média, desvio padrão e a relações matemáticas envolvidas. A terceira situação procurou observar componentes de disposição, analisando a postura crítica por meio da capacidade de gerar questionamentos diante das informações quantitativas. Nesse sentido, as variáveis foram determinadas seguindo a proposta de Gal (2002) que aponta a necessidade de uma correlação entre componentes cognitivos e de disposição para o desenvolvimento do letramento estatístico.

Assim como em Q1, as variáveis observadas em Q2 foram identificadas por códigos para o reconhecimento da questão durante o tratamento que foi realizado utilizando o software CHIC. Salientamos, novamente, que o mesmo questionário foi aplicado duas vezes, dessa forma, utilizamos na frente de cada código a letra “a” para identificar a variável “antes”, e a letra “d” para identificar a variável “depois”. Descreveremos, agora, as situações encontradas em Q2.

Quadro 24 - Q2: Situação 1

Situação-problema 1	
S1_1)	Vamos supor que você esteja interessado em investigar a quantidade de ocupantes dentro dos automóveis de passeio que transitam por dia em uma avenida movimentada de sua cidade. Como você procederia para dar início a essa investigação?
S1_2)	Frente as informações do item anterior, que tipo de medidas estatísticas você utilizaria para estabelecer algum tipo de padrão? Faça uma simulação.

Fonte: Do autor (2018).

A primeira situação procurou observar se os estudantes percebessem o uso de alguns procedimentos da estatística descritiva, como a noção de amostra, média e desvio padrão.

Figura 65 - Q2: variáveis da primeira situação

ANTES	DEPOIS
aS1_1 AMOSTRA	dS1_1 AMOSTRA
aS1_2A MÉDIA	dS1_2A MÉDIA
aS1_2B VARIÂNCIA/DESVIO PADRÃO	dS1_2B VARIÂNCIA/DESVIO PADRÃO
aS1_2C CÁLCULO	dS1_2C CÁLCULO
	dS1_2D NOÇÕES DE ESTIMAÇÃO

Fonte: Do autor (2018).

Em S1_1 esperava-se que os estudantes reconhecessem o processo da amostragem {aS1_1; dS1_1} para começar a investigação pretendida pelo problema. Em S1_2 esperava-se que os estudantes recorressem para o uso da média {aS1_2A; dS1_2A} e do desvio padrão {aS1_2B; dS1_2B} para estabelecer algum padrão retirado da possível amostra construída. Além disso, o problema solicitava uma simulação {aS1_2C; dS1_2C} por meio de um cálculo que utilizaria valores determinados pelos estudantes. Observamos também a variável dS1_2D que pretendia analisar se os estudantes, após o conteúdo de estimação, revelassem alguns procedimentos utilizados desse novo conteúdo.

Quadro 25 - Q2: Situação 2

Situação 2
<p>S2_1) Em um supermercado o funcionário responsável pelo hortifrutigranjeiro recebeu a incumbência de empacotar as maçãs em pacotes de dois quilos. O que o funcionário deverá levar em consideração?</p>
<p>S2_2) Após empacotar todas as maçãs, num total de 150 pacotes, o funcionário ficou curioso em saber sobre a distribuição das frutas nos pacotes. Não possuindo tempo para fazer uma contagem sobre todos os pacotes, que procedimento ele poderia realizar para fazer essa descoberta?</p>
<p>S2_3) Que tipos de medidas estatísticas podemos proceder para tentarmos estabelecer algum padrão em relação aos pacotes de maçãs? Faça uma simulação.</p>

Fonte: Do autor (2018).

A segunda situação procurou observar se os estudantes perceberam, além do uso de alguns procedimentos da estatística descritiva, a noção sobre a variabilidade do elemento que o problema apresenta.

Figura 66 - Q2: variáveis da segunda situação

ANTES	DEPOIS
aS2_1 VARIABILIDADE	dS2_1 VARIABILIDADE
aS2_2 AMOSTRA	dS2_2 AMOSTRA
aS2_3A MÉDIA	dS2_3A MÉDIA
aS2_3B VARIÂNCIA/DESVIO PADRÃO	dS2_3B VARIÂNCIA/DESVIO PADRÃO
aS2_3C CÁLCULO	dS2_3C CÁLCULO
	dS2_3D NOÇÕES DE ESTIMAÇÃO

Fonte: Do autor (2018).

Em S2_1 esperava-se que os estudantes identificassem a variabilidade {aS2_1; dS2_1} do elemento em questão e, pudessem supor que haverá diferentes quantidades de maçãs⁴⁸ nos pacotes. Em S2_2 esperava-se que os estudantes reconhecessem o processo da amostragem {aS2_2; dS2_2} para começar a investigação pretendida pelo problema. Em S2_3 esperava-se que os estudantes recorressem para o uso da média {aS2_3A; dS2_3A} e do desvio padrão {aS2_3B; dS2_3B} para estabelecer algum padrão retirado da possível amostra construída. Além disso, o problema solicitava uma simulação {aS2_3C; dS2_3C} por meio de um cálculo que utilizaria valores determinados pelo estudante. Observamos também a variável dS2_3D que pretendia analisar se os estudantes, após o conteúdo de estimação, revelassem alguns procedimentos utilizados desse novo conteúdo.

Quadro 26 - Q2: Situação 3

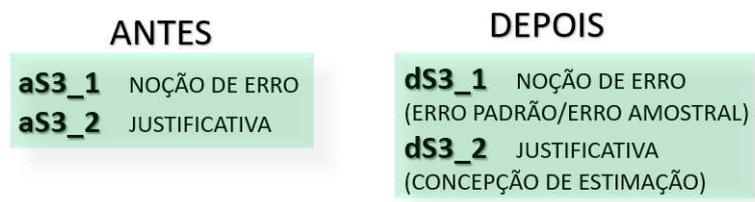
Situação 3
S3_1 Em ambas as situações anteriores (Situação 1 e Situação 2), você acredita que possam haver erros em relação aos padrões estabelecidos? Justifique sua resposta.

Fonte: Do autor (2018).

A terceira situação procurou analisar a percepção dos estudantes em relação a noção de erro.

⁴⁸ A literatura apresenta um peso médio de aproximadamente 154g para uma maçã.

Figura 67 - Q2: variáveis da terceira situação

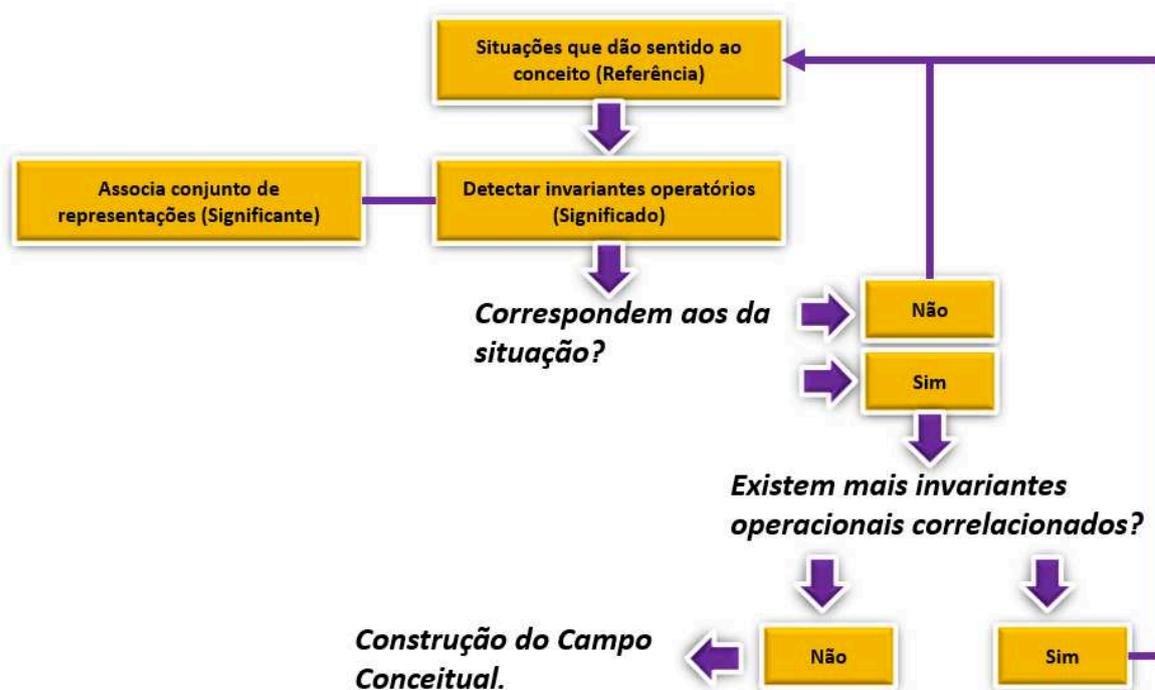


Fonte: Do autor (2018).

Em aS3_1 foi observado se o estudante concordava, ou não, com a existência de possíveis erros diante dos padrões que eles estabeleceram nas situações anteriores. Para dS3_1 esperava-se que os estudantes já mencionassem sobre erro padrão, ou erro amostral. Em aS3_2 os estudantes já poderiam revelar uma justificativa que conduzisse para o entendimento da estimação. Em dS3_2 esperava-se por uma justificativa que fosse o conceito de estimação, ou muito próximo a ele.

Para explorarmos as respostas à luz da Teoria dos Campos Conceituais e analisá-las fazendo uso da Análise Estatística Implicativa por meio do software CHIC, estabelecemos um algoritmo como critério que está representado na Figura 68.

Figura 68 - Critério para analisar as respostas de Q2



Fonte: Do autor (2018).

Esse esquema mostra, de forma simplificada, as etapas fundamentais apoiadas em Vergnaud (1990a) para a construção de um campo conceitual. As situações propostas em Q2

dão sentido aos conceitos pretendidos. Dessa forma, buscamos pelos invariantes operatórios (teoremas-em-ação e conceitos-em-ação) que correspondam com as variáveis analisadas em cada questão. Verificamos, também, o conjunto de representações que estão associados com as variáveis e ligadas horizontalmente com os invariantes. Nesse sentido, para cada resposta coerente com a intenção das variáveis analisadas, inserimos o número um em uma planilha do Excel como resposta coerente para a variável analisada, caso contrário, zero.

Figura 69 - Planilha com as variáveis binárias

A	B	C	D	E	F	G	H	I	J	K	L
	aS1_1	aS1_2A	aS1_2B	aS1_2C	aS2_1	aS2_2	aS2_3A	aS2_3B	aS2_3C	aS3_1	aS3_2
E01_001	0	1	0	0	0	1	1	0	0	1	0
E01_002	1	0	0	1	0	0	1	0	1	0	0
E01_003	1	0	0	1	1	1	0	0	0	0	0
E01_004	1	1	0	0	1	1	0	0	1	0	0
E01_005	0	1	1	0	0	1	1	0	1	0	0
E01_006	0	0	0	1	0	0	0	0	0	0	0
E01_007	1	0	0	0	0	1	0	0	0	1	0
E01_008	1	0	0	0	0	1	1	1	1	1	0
E01_009	1	0	0	0	0	1	0	0	0	0	0
E01_010	1	0	0	1	0	1	0	0	1	1	0
E01_011	1	1	1	0	1	1	1	0	0	1	0
E01_012	0	0	0	0	0	1	0	0	0	0	0
E01_013	1	0	0	0	0	1	1	0	1	1	0
E01_014	1	0	0	1	0	0	0	0	0	1	0
E01_015	1	0	0	0	1	0	1	0	1	0	0
E01_016	1	0	0	0	0	1	0	0	0	1	0
E01_017	1	0	0	0	0	1	1	1	0	0	0
E01_018	1	0	0	0	0	1	0	0	1	1	0

Fonte: Do autor (2018).

Desse modo, construímos uma planilha com variáveis binárias, presença ou ausência do atributo analisado, que servirá de base para o tratamento dos dados com o software CHIC por meio da construção da árvore de similaridades, do grafo implicativo e da árvore coesiva, observando a Análise Estatística Implicativa.

Figura 70 - Síntese das ferramentas de construção, tratamento e análise dos dados



Fonte: Do autor (2018).

As reflexões sobre a Educação Estatística realizadas em nosso referencial teórico e o estudo da Teoria dos Campos Conceituais, nos ajudaram ao encaminhamento de discussões que foram geradas durante a observação dos dados coletados dos questionários aplicados em nossa pesquisa. Esses dados foram tratados por meio dos softwares EXCEL, SPAD e CHIC que expuseram informações estatísticas a fim de colaborar com as análises realizadas.

PARTE 3:

RESULTADOS, DISCUSSÕES E PROLONGAMENTOS

Depois de descrevermos as escolhas metodológicas, indicando o contexto do estudo e especificando as ferramentas utilizadas para a coleta e tratamento dos dados, iremos, agora, detalhar os resultados obtidos. Diante da grande quantidade de dados gerados por meio dos dois questionários aplicados, selecionamos as informações que foram relevantes para a geração de subsídios que atendessem na verificação ou não das hipóteses levantadas:

H1: Existem limitações na formação do raciocínio estatístico, na estimação estatística, ligadas aos procedimentos e conceitos básicos da Estatística Descritiva;

H2: A relação afetiva do estudante com a Estatística pode influenciar na formação do raciocínio estatístico;

H3: A relação afetiva do estudante com a Matemática pode influenciar na formação do raciocínio estatístico.

Dessa forma, dividimos essa parte em três capítulos. O primeiro capítulo mostra os resultados e as discussões realizadas observando o primeiro questionário (Q1). No segundo capítulo estão os resultados e as discussões em relação ao segundo questionário (Q2). O último capítulo divulga as considerações finais em relação as hipóteses levantadas na pesquisa e um prolongamento das discussões que serão realizadas para futuras contribuições da tese.

1. Resultados e discussões do Q1

Analizamos descritivamente o questionário Q1 por meio do software Excel e SPAD. Realizamos alguns cruzamentos que foram medidos pelo software SPAD utilizando o teste não paramétrico Qui-Quadrado. O Teste t Student também foi empregado pelo software para a verificação da existência de diferença significativa entre as médias das variáveis contínuas. Em todos os testes realizados foi considerado um nível de significância de 5%.

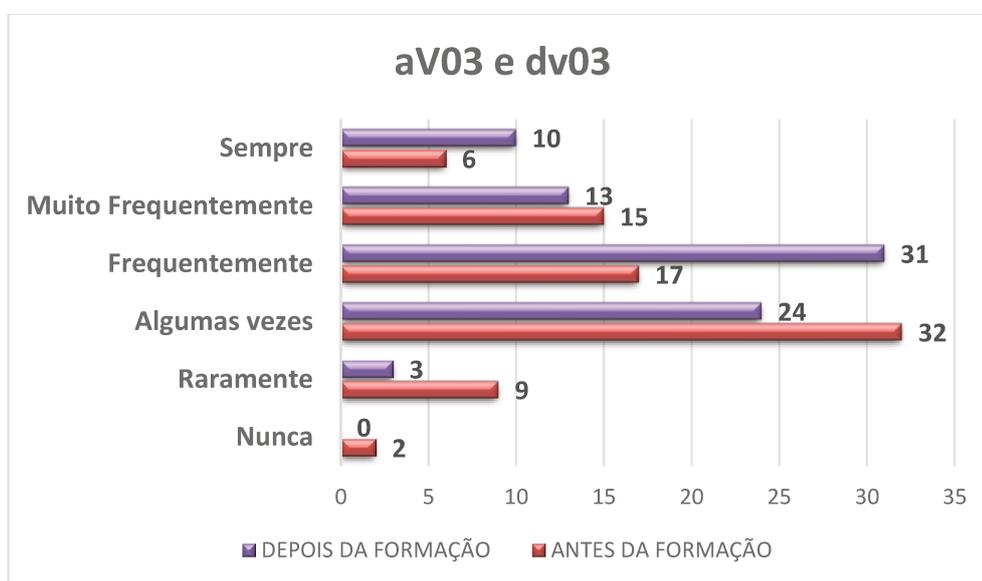
As primeiras informações geradas pelo questionário Q1, revelaram que entre os participantes da pesquisa, 51% são do sexo feminino e 49% do sexo masculino (variável V01). Suas idades (variável V02) variam de 17 a 45 anos com média de 24,19 anos e desvio padrão de 5,60. O Instituto Nacional de Estudos e Pesquisas Educacionais (INEP)⁴⁹, aponta que 58,5% dos estudantes que frequentam o Ensino Superior no Brasil estão na faixa etária dos 18 a 24

⁴⁹ Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Sinopse Estatística da Educação Superior 2017. Disponível em: <http://portal.inep.gov.br/web/guest/educacao-superior>.

anos com uma média global de 21 anos e tendo o sexo feminino como o gênero majoritário entre os matriculados.

A variável V03, que interpelou os estudantes sobre o enfrentamento de situações que exigiram conhecimentos estatísticos, apontou que boa parte dos respondentes já reconheceu situações onde a Estatística se fez necessária para possíveis interpretações. Isso mostra indícios que os estudantes possuem, segundo Campos (2007), um dos estágios para o desenvolvimento do letramento estatístico que é, justamente, o reconhecimento da linguagem estatística inserida num contexto de alguma discussão social.

Gráfico 8 - V03: Você já enfrentou situações que exigiram conhecimentos estatísticos?

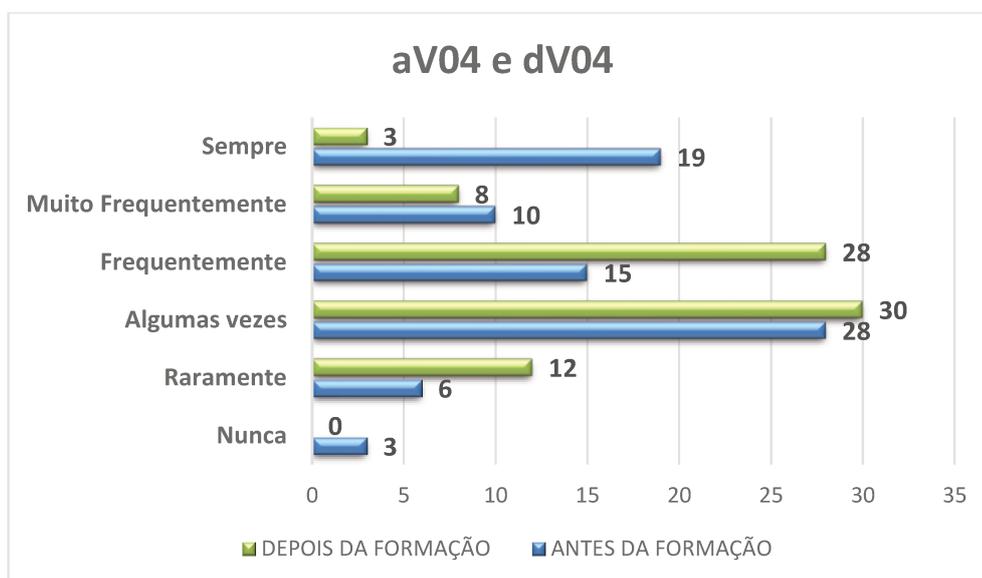


Fonte: Do autor (2018).

Dos 81 respondentes, 38 (antes da formação sobre estimação) afirmaram o enfrentamento frequente de alguma situação. Logo após a formação sobre estimação, esse número passou para 54, conduzindo para uma interpretação de que o desenvolvimento de novos conceitos estatísticos promoveu nos estudantes uma amplitude de percepções na compreensão de informações que utilizam a Estatística para serem expressas. Verificamos, também, que o nível de positividade dos estudantes não possui associação ($p\text{-valor} \geq 0,05$ antes e depois) com o reconhecimento de situações estatísticas (Anexo E).

A variável V04 buscou reconhecer a existência de possíveis obstáculos que os estudantes possuem quando existe a necessidade do uso de conhecimentos estatísticos.

Gráfico 9 - V04: Quando os conhecimentos em Estatística são requisitados, isso lhe causa problemas?



Fonte: Do autor (2018).

Antes da formação sobre estimação, mais de 50% ($n = 44/81$) dos respondentes confirmaram possuir dificuldades frequentes em relação aos conhecimentos estatísticos. Esse percentual diminuiu um pouco ($n = 39/81$) para as respostas obtidas depois da formação. Concordamos com Diaz (2016) que entende esses dados como indicativos de que as percepções negativas sobre a capacidade de resolver problemas estatísticos podem ser uma das causas de resistência que dificultam a aquisição de conhecimentos e habilidades. A relação da variável V04 com o nível de positividade (Anexo F) antes da formação mostrou a existência de associação ($p\text{-valor} < 0,05$) o que não aconteceu depois da formação ($p\text{-valor} \geq 0,05$). Acreditamos que houve uma atenuação em relação a alguns obstáculos que os estudantes encontram, visto que a estimação requer o uso de conceitos e procedimentos básicos da Estatística e que foram revisitados durante as aulas de estimação. Esse é um indício de que nossa primeira hipótese que admite existir limitações no processo de formação do raciocínio estatístico por meio da estimação ligadas ao processo cognitivo dos procedimentos e conceitos básicos da estatística descritiva está coerente.

A variável V05 apresentou a disposição dos estudantes por semestre. Em nosso contexto, a disciplina de Estatística é ofertada no 3º semestre dos cursos que participaram da pesquisa, onde encontramos a maior frequência de estudantes ($n = 28/81$). Porém, a maioria dos estudantes estão localizados em semestres distintos pelo de fato de que muitos optam em cursar essa disciplina mais tarde, independentemente do nível de positividade (Anexo G). Nossa experiência docente observa essa ocorrência com muita regularidade, onde os estudantes

justificam essa opção por considerarem uma disciplina difícil. Um outro motivo, mas que não foi apurado nessa pesquisa, é o da reprovação. Dessa forma, a variável V05 revela indícios de uma relação que nem sempre é tão cordial entre os estudantes e a disciplina de Estatística.

Quadro 27 - Frequência dos estudantes por semestre

Semestres	Frequência	%
2º Semestre	1	1,2 %
3º Semestre	28	34,6 %
4º Semestre	21	25,9 %
5º Semestre	8	9,9 %
6º Semestre	10	12,3 %
7º Semestre	13	16,0 %
Total	81	100,0 %

Fonte: Do autor (2018).

A variável V06 solicitou que os estudantes escrevessem três palavras que remetesse a Estatística, além da elaboração de uma frase que explicitasse essas escolhas. Dessa forma, foi possível observar as associações que os estudantes expressaram em relação a disciplina.

Figura 71 - aV06 e dv06: hoje, o que lhe evoca a Estatística – Palavras



Fonte: Do autor (2018).

De acordo com a nuvem de palavras acima, o termo “Matemática” foi o mais demandado antes e depois da formação. Observamos que antes da formação existiu uma forte tendência para associações com elementos da matemática, seguindo Régnier (2005) que também mostrou em sua pesquisa um alto índice de respostas relacionadas com a matemática ou com elementos da matemática.

Quadro 28 - aV06 e dV06: cinco palavras com maior frequência

ANTES		DEPOIS	
Palavra	Frequência	Palavra	Frequência
MATEMÁTICA	30	MATEMÁTICA	30
NÚMEROS	22	AMOSTRA	19
GRÁFICOS	21	GRÁFICOS	18
CÁLCULOS	13	PESQUISA	17
DADOS	11	INTERPRETAÇÕES	16

Fonte: Do autor (2018).

É possível observar que depois da formação os estudantes manifestaram, com maior frequência, elementos mais próximos da Estatística. Esse dado nos revela um possível desenvolvimento do letramento estatístico, onde os termos como “Amostra”, “Pesquisa” e “Interpretações” estão mais contextualizados com a terminologia estatística.

Verificamos, também, as palavras com maior frequência de acordo com os grupos formados pelos índices de positividade.

Figura 72 - aV06: Palavras por níveis de Positividade "Antes"



Fonte: Do autor (2018).

Antes da formação sobre estimação estatística, o termo “Matemática” possui o maior destaque nos grupos com os menores índices de positividade. Ao verificarmos as frases que foram desenvolvidas, observamos que o termo “Matemática” aparece quase sempre com uma conotação negativa (Anexo H).

Figura 73 - Frases “Antes”

a_Positividade_1

**Não sei matemática e não sei ler gráficos.
 Não entendo por que tem matemática.
 Interpretações dos gráficos e matemática.
 Interessante, porém tenho muita dificuldade.
 Não consigo entender matemática.
 Muitas interpretações com números.
 Matemática com muitos cálculos e fórmulas.**

a_Positividade_2

**Matemática com aplicações.
 Matemática com exercícios mais práticos.**

a_Positividade_3

**Não gosto por que tem matemática e muita conta.
 Quando envolve números fica muito complicado.**

a_Positividade_4

**Muito útil para compreender muitas informações que lemos.
 Muita dificuldade, mas muito útil.**

Fonte: Do autor (2018).

Depois da formação sobre estimação estatística, o termo “Matemática” aparece com destaque em todos os níveis de positividade.

Figura 74 - dV06: Palavras por níveis de positividade “Depois”

d_Positividade 1



d_Positividade 2



d_Positividade 3



d_Positividade 4



Fonte: Do autor (2018).

Acreditamos que esse fato tenha ocorrido por que antes da exposição do conteúdo sobre estimação foi realizado uma revisão dos conteúdos relativos a estatística descritiva, atentando para as operações matemáticas envolvidas. Essa retomada foi observada como necessária após o exame da primeira aplicação do questionário e, também, por solicitação dos estudantes que manifestaram dificuldades nos procedimentos envolvendo cálculos. Entretanto, as frases que foram expostas depois da estimação, revelaram um contexto mais voltado para os processos estatísticos (Anexo I).

Figura 75 - Frases “Depois”

d_Positividade_1
Usada para a realização de pesquisas por meio de uma amostra.
Tomar decisões por meio de dados coletados, amostra.
Ciência que analisa os dados de uma amostra para uma pesquisa.
Tira conclusões com base em uma amostra.

d_Positividade_2
Coleta e interpreta dados para um trabalho de pesquisa.
É uma ciência capaz de coletar e interpretar dados.
Informações dos dados de uma pesquisa.
Desvio padrão, média, mediana.
Interpretações de resultados.
Interpretar dados numéricos.
Analisa dados coletados de uma pesquisa.

d_Positividade_3
Interpretar dados.
Ciência que domina a área de pesquisas.
Analisa uma amostra.
Coleta e interpreta dados.

d_Positividade_4
Descreve e analisa dados.
Muito útil para coletar e interpretar dados.
Análise de dados para uma pesquisa.

Fonte: Do autor (2018).

Entendemos que houve uma compreensão que a matemática, como é exposto por Hand (1998), ajuda na formalização das ideias estatísticas, mas que a concepção estatística envolve outros elementos. Silva *et al.* (2002) compreendem que o reforço em fundamentos matemáticos é necessário quando o ensino é voltado para a formação de estatísticos, mas deve ser reduzido quando os estudantes serão apenas usuários das ferramentas estatísticas. Embora compactuando com essa alegação, percebemos que o reforço matemático possibilitou aos estudantes um importante avanço no letramento estatístico. Ao proporcionar uma atenuação das dificuldades advindas das operações matemáticas, os estudantes vislumbraram, independentemente do nível de positividade, os aspectos práticos da estatística. Essas reflexões reforçaram nossas três hipóteses da pesquisa. A retomada dos conceitos básicos da estatística com um melhor auxílio

nas operações matemáticas, que foi vista como um limitador para o desenvolvimento do letramento estatístico, ajudou os estudantes no progresso do pensamento e raciocínio estatísticos, atenuando suas atitudes negativas em relação à estatística e à matemática.

A variável V07 solicitou que os estudantes declarassem o nível de interesse pessoal que eles atribuíam para uma formação em estatística na sua graduação.

Gráfico 10 - V07: Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação?



Fonte: Do autor (2018).

Antes da formação, aproximadamente 67% ($n = 54/81$) dos estudantes consideraram um nível igual ou superior a 7 e, depois da formação, aproximadamente 95% ($n = 77/81$), mostrando uma tendência positiva no interesse pessoal com a estatística para a formação profissional. De um modo geral, houve uma melhora significativa (Anexo J) entre as médias encontradas para os níveis atribuídos de antes e depois. Esse fato pode ser observado, também, quando o software SPAD fez o cruzamento entre V07 com os níveis de positividade, como mostra o Quadro 29.

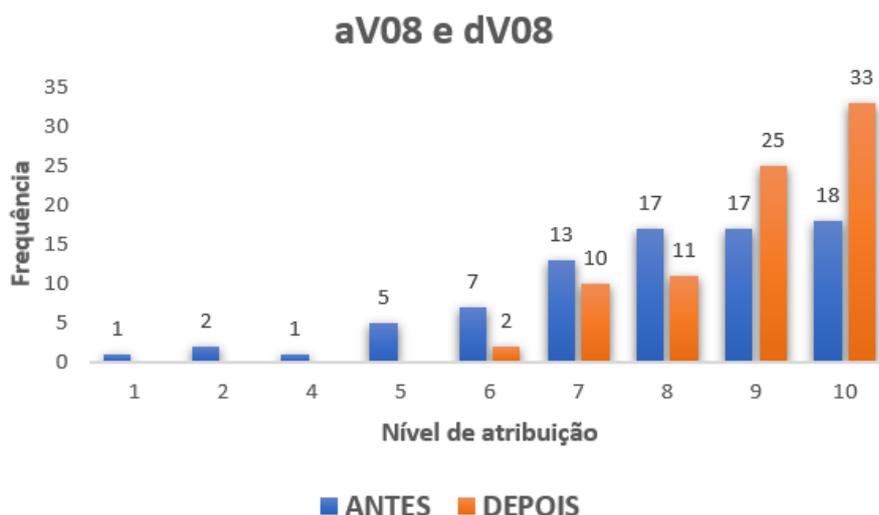
Quadro 29 - Relação V07 X índice de positividade “Antes e Depois”

Média	[aV07]	Média	[dV07]
Desvio-Padrão		Desvio-Padrão	
Total		Total	
a_Positividade 1	5,682	d_Positividade 1	8,273
	1,585		1,279
	22		22
a_Positividade 2	7,269	d_Positividade 2	8,200
	1,430		1,322
	26		20
a_Positividade 3	8,214	d_Positividade 3	8,238
	1,424		1,136
	14		21
a_Positividade 4	8,158	d_Positividade 4	8,278
	1,740		1,320
	19		18
Total	7,210	Total	8,247
	1,828		1,240
	81		81

Fonte: Do autor (2018).

A variável V08 solicitou que os estudantes declarassem o nível de utilidade que eles atribuíam para uma formação Estatística na sua graduação.

Gráfico 11 - V08: Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação?



Fonte: Do autor (2018).

Antes da formação, aproximadamente 80% ($n = 65/81$) dos estudantes consideraram um nível igual ou superior a 7 e, depois da formação, aproximadamente 97% ($n = 79/81$), mostrando uma tendência positiva na utilidade das ferramentas estatística aplicadas na

formação profissional. De um modo geral, houve uma melhora significativa (Anexo K) entre as médias encontradas para os níveis atribuídos de antes e depois. Esse fato pode ser observado, também, quando o software SPAD fez o cruzamento entre V08 com os níveis de positividade, como mostra o Quadro 30.

Quadro 30 - Relação V08 X índice de positividade “Antes e Depois”

Média	[aV08]	Média	[dV08]
Desvio-Padrão		Desvio-Padrão	
Total		Total	
a_Positividade 1	6,727	d_Positividade 1	8,773
	2,434		1,152
	22		22
a_Positividade 2	8,077	d_Positividade 2	9,000
	1,412		1,076
	26		20
a_Positividade 3	8,357	d_Positividade 3	9,381
	1,737		0,805
	14		21
a_Positividade 4	8,474	d_Positividade 4	8,611
	1,775		1,378
	19		18
Total	7,852	Total	8,951
	1,969		1,128
	81		81

Fonte: Do autor (2018).

Nas variáveis aV07 e aV08 as zonas de comportamento com os menores índices de positividade (a_Positividade_1 e a_Positividade_2), apresentaram as menores médias para os níveis atribuídos, assim como as zonas de comportamento com os maiores índices de positividade (a_Positividade_3 e a_Positividade_4), exibiram as maiores médias. Dessa forma, podemos entender que houve uma disposição das influências dos estudantes em relação à estatística sobre os níveis escolhidos. Esse fato não ficou acentuado nas variáveis dV07 e dV08, onde a apresentação das médias foi mais homogênea em todas as zonas de comportamento. Entendemos que os estudantes poderiam enxergar inúmeras utilidades advindas das ferramentas estatísticas e, ao mesmo tempo, não ter interesse pessoal em dominá-las, o que não foi apurado na análise, existindo uma correlação positiva antes e depois entre as variáveis V07 e V08 (Anexo L).

A variável V09 interpelou os estudantes sobre o significado de alguns termos que fazem parte do contexto estatístico.

Figura 76 - V09: Atualmente o que significa para você os termos a seguir?

- V09a Representativo
- V09b Significativo
- V09c Estimação
- V09d Teste de Hipóteses
- V09e Probabilidade
- V09e Sondagem

Fonte: Do autor (2018).

Nessa questão os estudantes não responderam de forma assídua, nem antes e nem depois da formação sobre estimação. No entanto, as repostas encontradas possuem relações com os termos (Anexo M). Nos detemos na variável V09c, onde está o termo estimação.

Figura 77 - aV09c e dV09c: Estimação



Fonte: Do autor (2018).

Os estudantes, antes da formação sobre estimação, expressaram a sua concepção usando alguns elementos que possuem relações com o contexto do termo. Verificando por zonas de comportamento, os estudantes situados em a_Positividade_1 e a_Positividade_2 foram os menos frequentes nas respostas (Anexo N).

Figura 78 - aV09c: Estimação

a_Positividade_1
Tem relações com previsões.
Um recorte da população.
Previsões.

a_Positividade_2
São avaliações baseadas em dados.
Previsão.
Proximidade.
Tudo que faz previsões na estatística.

a_Positividade_3
Fazer previsões com a estatística.
Fazer uma suposição.
Estima dados.
Valor aproximado.

a_Positividade_4
Representa uma previsão, tendência significativa.
Estimar uma população.
Tendência.
Previsão da realidade.

Fonte: Do autor (2018).

Depois da formação, os estudantes usaram termos mais técnicos relativos a estimação estatística, como “parâmetros”, “pontual” e “intervalar”, por exemplo. No entanto, percebemos que as zonas de comportamento com os maiores índices de positividade foram as que mais usaram esses termos (Anexo O).

Figura 79 - dV09c: Estimação

d_Positividade_1
Inferência estatística.
Previsões de parâmetros.
Interpretações de uma amostra.
Cálculo com dados para inferir o geral.

d_Positividade_2
Uma tendência de valor de uma população.
Previsões que podem ter erros.
Cálculo com uma amostra para estimar uma população.

d_Positividade_3
É uma previsão da pesquisa de população por meio de uma amostra.
Usar uma amostra para fazer previsões.
Nível de confiança de um intervalo.
Previsão, pontual, intervalar.

d_Positividade_4
Cálculo do parâmetro por meio de uma amostra.
Estimar parâmetros.
São previsões através de cálculos, pontuais, intervalos de confiança.
Previsão com nível de confiança.

Fonte: Do autor (2018).

Dessa forma, constatamos uma tendência: apesar das respostas ocorrerem de forma mais assídua em todas as zonas de comportamento, os estudantes situados nas de maior positividade (d_Positividade_3 e d_Positividade_4) foram os que mais usaram termos condizentes com a estimação estatística.

Na variável V10 os estudantes foram questionados sobre o significado da expressão “sondagem de opinião”.

Figura 80 - Sondagem de opinião: aV10 e dV10



Fonte: Do autor (2018).

As respostas encontradas (antes e depois) em todas as zonas de comportamento foram coerentes com a proposta da questão, que foi a construção de uma amostra para a realização de uma pesquisa. Na variável V06, analisada anteriormente, os estudantes já haviam relacionado a “amostra” como uma das palavras que evocavam a estatística, verificando aqui, mais uma vez, essa relação.

Em V11 os estudantes dissertaram sobre as utilidades que eles atribuem para a estatística.

Figura 81 - aV11 e dV11: Nos dias de hoje, qual utilidade você atribui para a Estatística?



Fonte: Do autor (2018).

De modo geral, as respostas em todas as zonas de comportamento (antes e depois) foram equivalentes. Os estudantes entendem a utilidade da estatística aplicada nas relações de interpretações de informações e na elaboração de pesquisas (Anexo P). Esse contexto foi

observando, também, em V06, onde estudantes identificaram “interpretação” e “pesquisa” como palavras que estão relacionadas com a estatística.

A variável V12 solicitou que os estudantes escrevessem três palavras que remetesse sobre as dificuldades encontradas numa disciplina de estatística.

Figura 82 - aV12 e dV12: Para você, quais são as três maiores dificuldades encontradas numa disciplina de Estatística?



Fonte: Do autor (2018).

De acordo com a nuvem de palavras acima, as dificuldades mais recorrentes invocam (antes e depois) as operações de cálculo matemático e as interpretações, o que também foi examinado na pesquisa de Diaz (2016).

Quadro 31 - aV12 e dV12: cinco palavras com maior frequência

ANTES		DEPOIS	
Palavra	Frequência	Palavra	Frequência
Matemática	31	Matemática	49
Cálculo	24	Interpretação	19
Prova	9	Cálculo	15
Interpretação	22	Fórmulas	13
Gráficos	15	Gráficos	13

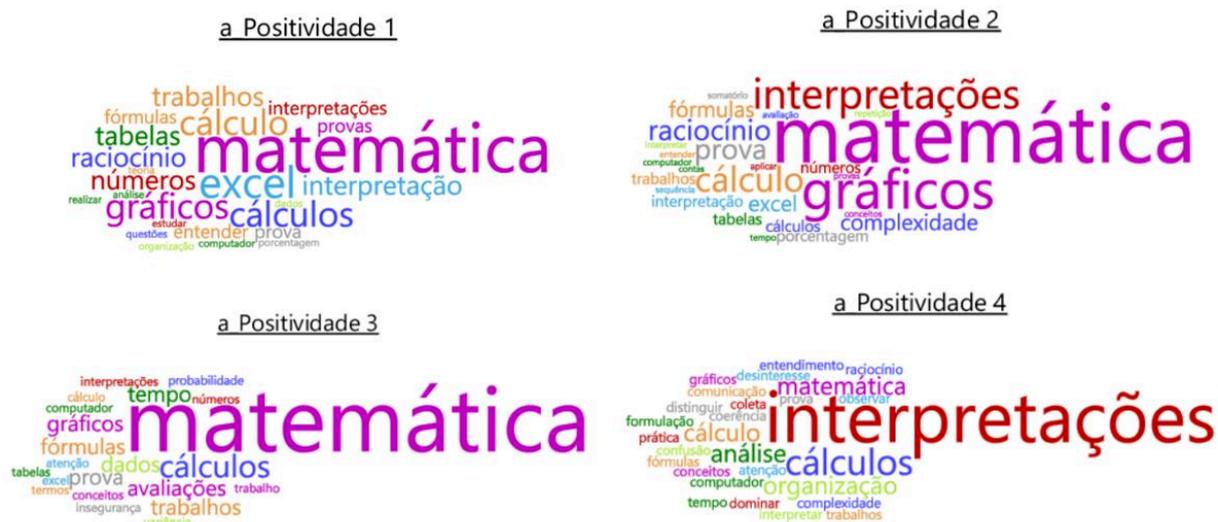
Fonte: Do autor (2018).

Segundo Vergnaud (2001) as dificuldades percebidas estão atribuídas a conhecimentos diferentes, pois uma evoca a forma predicativa e, a outra, a forma operativa do conhecimento. A forma predicativa está ligada ao que pode ser expresso verbalmente, o qual nos permite explicar e descrever os processos cognitivos que ocorrem durante uma ação, denominada, também, de conhecimento explícito, formalizado e articulado. A forma operatória está relacionada ao discernimento de fórmulas e operações matemáticas.

As dificuldades apontadas foram examinadas por zonas de comportamento, indicando que a forma operativa, antes da formação, não está evidente apenas na zona a_Positividade_4.

As zonas com os menores índices de positividade demonstram uma tendência negativa para os cálculos estatísticos.

Figura 83 - aV12 X Zonas de comportamento “Antes”



Fonte: Do autor (2018).

Como apontado durante a análise da variável V06, antes da exposição do conteúdo de estimação, foi realizada uma retomada dos conceitos e procedimentos básicos da estatística, trabalhando as operações matemáticas envolvidas. Apesar disso, constatamos que as dificuldades operativas permaneceram sinalizadas entre todas as zonas de comportamento como mostra a Figura 85.

Figura 84 - dV12 X Zonas de comportamento "Depois"



Fonte: Do autor (2018).

Entendemos que esse obstáculo leva os estudantes a permanecerem com atitudes negativas em relação à estatística, justificando e confirmando nossa terceira hipótese da pesquisa que revela a influência da matemática no desenvolvimento do letramento estatístico.

Em V13 os estudantes escreveram suas opiniões por meio de três palavras que expressassem as maiores contribuições relativas a uma disciplina de estatística.

Figura 85 - aV13 e dV13: Para você, quais são as três maiores contribuições encontradas numa disciplina de Estatística?



Fonte: Do autor (2018).

É possível observar que os estudantes relacionaram, em todas as zonas de comportamento, assim como na análise de V11, a elaboração de pesquisas como a principal contribuição da estatística.

Quadro 32 - aV13 e dV13: cinco palavras com maior frequência

ANTES		DEPOIS	
Palavra	Frequência	Palavra	Frequência
Pesquisa	24	Pesquisa	35
Gráficos	10	Gráficos	14
Raciocínio	9	Interpretação	12
Interpretação	10	Raciocínio	11
Informações	7	Excel	8

Fonte: Do autor (2018).

O termo “Intepretação”, apesar de ser apontado como uma das dificuldades na disciplina de estatística (V12), aparece, também, como contribuição. Dessa forma, entendemos que os estudantes reconhecem no estudo da estatística, elementos que podem suprir os obstáculos para o seu desenvolvimento profissional. Essa argumentação fica evidente, também, com a frequência das palavras “Gráficos”, “Raciocínio”, “Informações” e “Excel” que expressam possíveis situações do emprego da estatística nas práticas cotidianas dos estudantes.

Em V14 os estudantes definiram o seu nível de conhecimento estatístico. Antes da formação, aproximadamente 37% ($n = 30/81$) dos estudantes entendiam ter um nível de conhecimento estatístico muito superficial ou não sem nenhum. Essa ocorrência não condiz com a realidade dos grupos pesquisados, pois todos os estudantes já possuíam conhecimentos relativos a estatística descritiva e uma introdução dos elementos que circundam a estatística de inferência, que são conteúdos ofertados na disciplina de estatística antes da formação sobre estimação. Esse fato nos levou a entender que, mesmo possuindo alguns conhecimentos estatísticos, os estudantes não se sentem preparados para o enfrentamento de situações onde a estatística se faz presente.

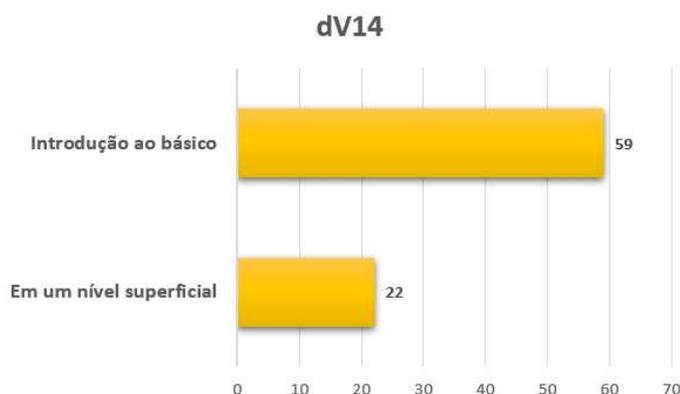
Gráfico 12 - aV14: Você já obteve conhecimentos em Estatística?



Fonte: Do autor (2018).

O nível de positividade antes da formação influenciou nessas escolhas (Anexo Q), onde os grupos com índices mais elevados reconheceram possuir competências estatísticas. Depois da formação, houve uma certa homogeneidade entre as escolhas, onde mais de 72% ($n = 59/81$) dos estudantes permaneceram na mesma região.

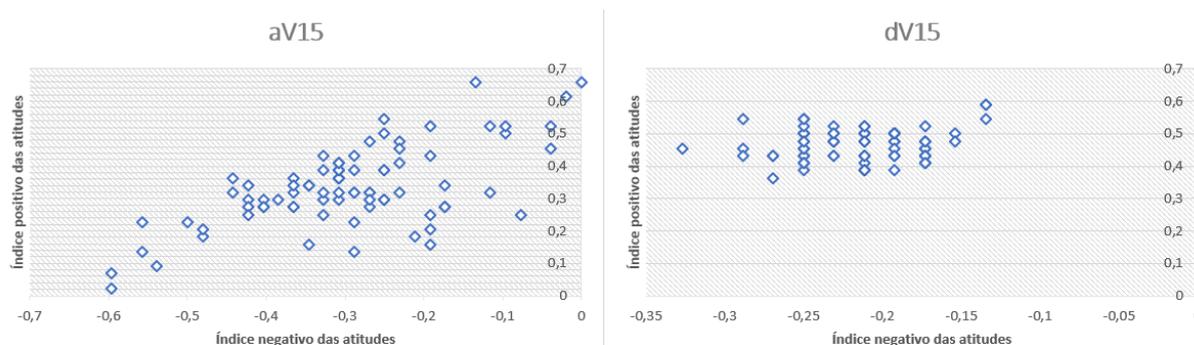
Gráfico 13 - dV14: Você já obteve conhecimentos em Estatística?



Fonte: Do autor (2018).

A variável V15, que tratou sobre o sentimento dos estudantes em relação a estatística, foi utilizada na construção das zonas de comportamento que foram analisadas com as demais variáveis do questionário. A metodologia da construção das zonas examinadas foi apresentada na Parte 2 (Anexo D) da pesquisa. Contudo, apresentamos algumas análises extraídas da variável (Anexo R). Houve uma melhora nos níveis de positividade dos estudantes em relação à estatística depois da formação sobre o conteúdo de estimação. A média do índice de positividade “antes” foi de 0,53 com um desvio padrão de 0,15 (aproximadamente 28% de variação), contra 0,67 de média com um desvio padrão de 0,04 (aproximadamente 6% de variação) “depois”.

Gráfico 14 - aV15 e dV15: dispersão dos índices



Fonte: Do autor (2018).

Observamos uma melhora significativa (Anexo S) entre as médias, onde, de um modo geral, tornou os grupos mais homogêneos. Compreendemos que uma das justificativas foi pelo fato de ter ocorrido uma retomada dos conceitos básicos da estatística antes da exposição do conteúdo, motivada pela análise preliminar da primeira aplicação do questionário. Encontramos também um outro argumento no livro de Selz e Maillachon (2009) onde é apontado uma mudança de comportamento dos estudantes quando o conteúdo da estatística inferencial é abordado. Segundo os autores, os estudantes começam a compreender melhor as aplicações da estatística, facilitando o processo de ensino e aprendizagem. Pactuamos com esse argumento diante das manifestações dos estudantes durante as aulas. Entretanto, encontramos transições dos estudantes nas quatro zonas de comportamento. O estudante E03_006, por exemplo, que estava situado em a_Positividade_1 passou para d_Positividade_4, assim como o estudante E05_005 passou de a_Positividade_4 para d_Positividade_1.

Descrevemos algumas tendências das zonas de comportamento por meio de uma classificação tipológica realizada pelo software SPAD, onde foi agrupado as frases das atitudes mais características de cada grupo. Não foi realizada uma análise detalhada das atitudes que

compreenderam os grupos, para essa pesquisa, elas apenas serviram de apoio para a construção dos índices de positividade.

Quadro 33 - Atitudes características: a_Positividade_1

Eu me sinto inseguro quando preciso utilizar a Estatística.	Concordo plenamente
Eu tenho medo de Estatística pois tenho medo de Matemática.	Concordo plenamente
Eu tenho uma relação muito boa com a Estatística : eu gosto e aprecio esta disciplina.	Absolutamente contra
O sentimento que experimento em relação a Estatística é bom.	Absolutamente contra
A Estatística é fascinante e divertida.	Absolutamente contra
A Estatística me dá segurança e me estimula.	Absolutamente contra
A Estatística me provoca um sentimento de indecisão que resulta no meu medo de não ser capaz de ter êxito na disciplina.	Concordo plenamente
Pensar em resolver um problema de Estatística me deixa nervoso.	Concordo plenamente

Fonte: Do autor (2018).

A zona a_Positividade_1 foi considerado o grupo com uma tendência de comportamento mais negativo, ocupando as extremidades para responder as atitudes de cunho afetivo, ou seja, concordou plenamente com as atitudes negativas e foi absolutamente contra as atitudes positivas.

Quadro 34 - Atitudes características: a_Positividade_2

Eu acho a Estatística muito interessante e gosto de ir na aula dessa disciplina.	Estou de acordo
Eu nunca gostei de Estatística, ela é a disciplina que me causa mais medo.	Não estou bem de acordo
Eu tenho medo de Estatística pois tenho medo de Matemática.	Estou de acordo
Quando estudo Estatística, não consigo pensar claramente pois minha cabeça fica bloqueada.	Estou de acordo
A Estatística me dá a impressão que estou perdido em uma sala cheia de números na qual eu não acho saída.	Não estou bem de acordo
A Estatística me dá segurança e me estimula.	Não estou bem de acordo
Eu me sinto inseguro quando preciso utilizar a Estatística.	Estou de acordo
Eu realmente não gosto de matemática.	Não estou bem de acordo
Eu sou competente em Matemática.	Não estou bem de acordo

Fonte: Do autor (2018).

A zona a_Positividade_2 foi considerado o grupo com uma tendência de comportamento misto, com uma disposição maior para atitudes negativas, não ocupando as extremidades para responder as atitudes de cunho afetivo e cognitivo.

Quadro 35 - Atitudes características: a_Positividade_3

Eu realmente não gosto de matemática.	Absolutamente contra
Eu não tenho um bom conhecimento em Estatística.	Não estou bem de acordo
Eu nunca gostei de Estatística, ela é a disciplina que me causa mais medo.	Absolutamente contra
Eu sou competente em Matemática.	Estou de acordo
A Estatística me dá a impressão que estou perdido em uma sala cheia de números na qual eu não acho saída.	Absolutamente contra
Eu tenho medo de Estatística pois tenho medo de Matemática.	Absolutamente contra
Quando estudo Estatística, não consigo pensar claramente pois minha cabeça fica bloqueada.	Absolutamente contra
A Estatística me provoca um sentimento de indecisão que resulta no meu medo de não ser capaz de ter êxito na disciplina.	Absolutamente contra

Fonte: Do autor (2018).

A zona a_Positividade_3 foi considerado o grupo com uma tendência de comportamento misto, com uma disposição maior para atitudes positivas, ocupando as extremidades para responder algumas as atitudes de cunho afetivo e cognitivo.

Quadro 36 - Atitudes características: a_Positividade_4

Pensar em resolver um problema de Estatística me deixa nervoso.	Absolutamente contra
Eu realmente gosto de Estatística.	Concordo plenamente
Eu não tenho um bom conhecimento em Estatística.	Absolutamente contra
Eu tenho uma relação muito boa com a Estatística: eu gosto e aprecio esta disciplina.	Concordo plenamente
Eu acho a Estatística muito interessante e gosto de ir na aula dessa disciplina.	Concordo plenamente
A Estatística é fascinante e divertida.	Concordo plenamente
O sentimento que experimento em relação a Estatística é bom.	Concordo plenamente
Eu tenho medo de Estatística pois tenho medo de Matemática.	Absolutamente contra

Fonte: Do autor (2018).

A zona a_Positividade_4 foi considerado o grupo com uma tendência de comportamento mais positivo, ocupando as extremidades para responder as atitudes de cunho afetivo e

cognitivo, ou seja, concordou plenamente com as atitudes positivas e foi absolutamente contra as atitudes negativas.

Depois da formação sobre estimação, não houve um grupo com uma tendência totalmente negativa como em a_Positividade_1, o que nos revela uma melhora nas atitudes em relação a estatística e um possível avanço no desenvolvimento do letramento estatístico. Entretanto, dos grupos analisados, d_Positividade_1 foi o menos positivo.

Quadro 37 - Atitudes características: d_Positividade_1

Pensar em resolver um problema de Estatística me deixa nervoso.	Absolutamente contra
O sentimento que experimento em relação a Estatística é bom.	Absolutamente contra
A Estatística me dá segurança e me estimula.	Concordo plenamente
Eu tenho medo de Estatística pois tenho medo de Matemática.	Concordo plenamente
A Estatística me dá a impressão que estou perdido em uma sala cheia de números na qual eu não acho saída.	Não estou bem de acordo
A Estatística me deixa inquieto, descontente, irritado e impaciente.	Não estou bem de acordo
Eu nunca gostei de Estatística, ela é a disciplina que me causa mais medo.	Estou de acordo
Eu tenho uma relação muito boa com a Estatística: eu gosto e aprecio esta disciplina.	Estou de acordo

Fonte: Do autor (2018).

É possível observar que d_Positividade_1 possui uma tendência não favorável com a estatística, o que pode ser justificado pelo componente cognitivo das operações matemáticas. No entanto, existe uma disposição positiva no aspecto afetivo com a disciplina.

Quadro 38 - Atitudes características: d_Positividade_2

Eu não gosto de Estatística e nem de ir nas aulas de Estatística.	Não estou bem de acordo
Eu tenho uma relação muito boa com a Estatística: eu gosto e aprecio esta disciplina.	Concordo plenamente
Eu estou sempre estressado durante as aulas de Estatística.	Não estou bem de acordo
Eu tenho medo de Estatística pois tenho medo de Matemática.	Não estou bem de acordo
A Estatística me deixa inquieto, descontente, irritado e impaciente.	Não estou bem de acordo
A Estatística me dá segurança e me estimula.	Estou de acordo
Eu me sinto mais feliz com a Estatística do que com as outras disciplinas.	Estou de acordo
Quando eu escuto a palavra Estatística eu sinto um sentimento de aversão.	Não estou bem de acordo

Fonte: Do autor (2018).

A zona d_Positividade_2 é intermediária, com uma tendência em concordar parcialmente com atitudes positivas e negativas de cunho afetivo e cognitivo.

Quadro 39 - Atitudes características: d_Positividade_3

Eu não gosto de Estatística e nem de ir nas aulas de Estatística.	Absolutamente contra
A Estatística me deixa inquieto, descontente, irritado e impaciente.	Absolutamente contra
Eu estou sempre estressado durante as aulas de Estatística.	Absolutamente contra
Eu me sinto mais feliz com a Estatística do que com as outras disciplinas.	Não estou bem de acordo
Eu tenho medo de Estatística pois tenho medo de Matemática.	Estou de acordo
Eu realmente gosto de Estatística.	Concordo plenamente
Eu acho a Estatística muito interessante e gosto de ir na aula dessa disciplina.	Concordo plenamente
A Estatística me dá a impressão que estou perdido em uma sala cheia de números na qual eu não acho saída.	Absolutamente contra

Fonte: Do autor (2018).

A zona d_Positividade_3 é intermediária, apresentando uma tendência em concordar plenamente com atitudes positivas e absolutamente contras atitudes negativas de cunho afetivo.

Quadro 40 - Atitudes características: d_Positividade_4

Eu realmente gosto de Estatística.	Concordo plenamente
Eu tenho medo de Estatística pois tenho medo de Matemática.	Absolutamente contra
Eu sou competente em Matemática.	Concordo plenamente
Eu me sinto mais feliz com a Estatística do que com as outras disciplinas.	Concordo plenamente
A Estatística é fascinante e divertida.	Concordo plenamente
Eu nunca gostei de Estatística, ela é a disciplina que me causa mais medo.	Absolutamente contra
Eu me sinto bem com a Estatística e gosto muito dessa disciplina.	Concordo plenamente
Quando estudo Estatística, não consigo pensar claramente pois minha cabeça fica bloqueada.	Absolutamente contra

Fonte: Do autor (2018).

A zona d_Positividade_4 foi considerado o grupo com uma tendência de comportamento mais positivo, ocupando as extremidades para responder as atitudes de cunho afetivo e cognitivo, ou seja, concordou plenamente com as atitudes positivas e foi absolutamente contra as atitudes negativas.

Ainda em relação a análise da variável V15, constatamos que não existe diferença significativa entre os gêneros, os níveis de positividade (antes e depois) foram bem distribuídos entre homens e mulheres (Anexo T), assim como nos cursos de graduação envolvidos na pesquisa (Anexo U).

A última variável de Q1, V16, solicitou que os estudantes escrevessem a sua atual definição para estatística.

Figura 86 - aV16 e dV016: Atualmente, qual sua definição de Estatística?



Fonte: Do autor (2018).

De modo geral, todos os estudantes, em todas as zonas de comportamento (antes e depois), entendem a estatística como uma ciência que coleta, organiza e analisa os dados para uma pesquisa (Anexo V). Abaixo, organizamos uma síntese das investigações realizadas nas variáveis de Q1:

V01: dos participantes da pesquisa, 51% são do sexo feminino e 49% do sexo masculino, não houve uma diferença significativa entre o gênero e as zonas de comportamento.

V02: a média da idade dos participantes da pesquisa é de aproximadamente 24 anos.

V03: os estudantes reconhecem situações onde a estatística se faz necessária, independentemente do índice de positividade.

V04: os estudantes com os menores índices de positividade são mais propensos a encontrarem dificuldades em situações que exigem conhecimentos estatísticos, fato que ficou atenuado após a retomada dos conceitos e procedimentos operatórios básicos da estatística descritiva.

V05: existe uma tendência entre os estudantes de deixar a disciplina de estatística para ser cursada com mais tempo no curso, independentemente no nível de positividade.

V06: a estatística é lembrada, principalmente, pela matemática ou por elementos da matemática, independentemente do nível de positividade.

V07: os estudantes possuem uma tendência em atribuir um bom nível de interesse pessoal em relação a estatística, em todas as zonas de comportamento.

V08: os estudantes possuem uma tendência em atribuir um bom nível de utilidade em relação a estatística, em todas as zonas de comportamento, ocorrendo uma correlação positiva entre os níveis de interesse e os níveis de utilidade.

V09: os estudantes reconhecem os termos estatísticos, mas as zonas com os maiores índices de positividade possuem uma tendência em utilizar termos mais técnicos nas definições.

V10: os estudantes possuem uma tendência em reconhecer o termo estatístico, em todas as zonas de comportamento.

V11: os estudantes possuem uma tendência em reconhecer a utilidade da estatística nas relações de pesquisas e interpretações de informações, em todas as zonas de comportamento.

V12: a maior tendência encontrada nas dificuldades dos estudantes em relação a estatística é a matemática, seguido das interpretações, em todas as zonas de comportamento.

V13: as maiores tendências encontradas para a utilidade da estatística estão relacionadas a interpretação de informações e na elaboração de pesquisas, em todas as zonas de comportamento.

V14: os estudantes possuem uma tendência em reconhecer o seu nível de conhecimento em estatística entre um nível superficial e de introdução ao básico.

V15: os índices de positividade obtiveram um aumento depois da formação sobre estimação, estimuladas pela retomada dos conteúdos da estatística descritiva e das operações matemáticas.

V16: os estudantes possuem uma tendência em reconhecer a estatística como uma ciência que coleta, analisa e interpreta os dados para uma pesquisa.

2. Resultados e discussões do Q2

Analizamos Q2 por meio do software CHIC, observando a Análise Estatística Implicativa, considerando as respostas dos estudantes antes e depois da formação sobre estimação. A análise teve como base a procura pelos invariantes operatórios (conceitos-em-ação e teoremas-em-ação) utilizados pelos estudantes durante a resolução das situações-problema, seguindo o algoritmo apresentado na Figura 68 que verificou nas respostas a formação de um possível campo conceitual para a estimação. Vergnaud (1990) demonstra na Teoria dos Campos Conceituais uma preocupação com o sujeito-em-situação. Essa característica torna a teoria útil no planejamento e na análise de situações de ensino, uma vez que temos a necessidade de acompanhar os estudantes enquanto aprendem, procurando, nos conceitos e teoremas em ação, a evolução temporal de seu conhecimento. Nesse sentido, relacionamos as variáveis envolvidas em Q2 com as variáveis de V15 do questionário Q1, que mostrou as zonas de comportamento dos estudantes em relação a estatística. Dessa forma, observamos se a evolução dos conhecimentos estão relacionadas com as atitudes dos estudantes.

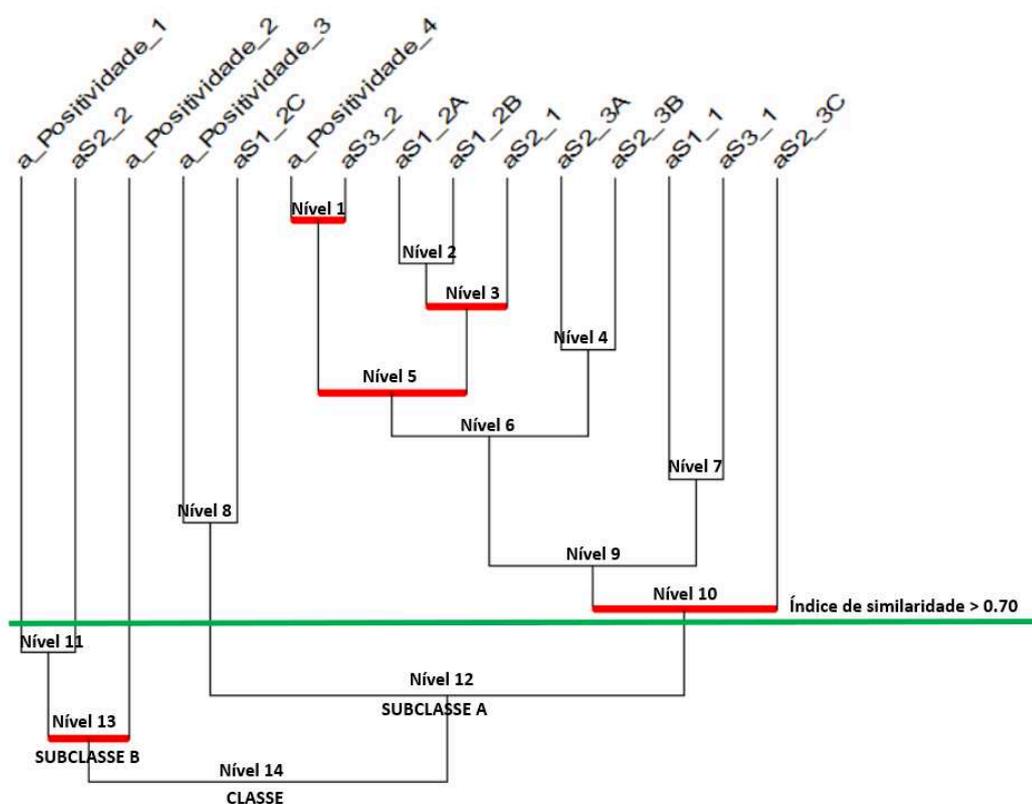
As variáveis foram tratadas seguindo a lógica binária, conforme seja identificada a ocorrência coerente ou não da resposta, representado pelos valores zero ou um. Esses valores foram lançados em uma planilha eletrônica do Excel (com extensão do arquivo.csv), que é importada pelo CHIC para processar os dados (Anexos W e X). O software CHIC construiu árvores de similaridades que foram definidas a partir do cruzamento do conjunto das variáveis (respostas dos estudantes) com o conjunto dos sujeitos (estudantes). Este tipo de análise permite, segundo Gras e Ag Almouloud (2005), estudar e interpretar, em termos de tipologia e de semelhança (e não semelhança) decrescente, classes de variáveis constituídas significativamente a certos níveis da árvore e se opondo a outras nestes mesmos níveis. Analisamos, também, o grafo implicativo construído pelo software por meio de uma

distribuição binomial, com um valor mínimo de implicação igual a 0,70 que permite, segundo Gras e Ag Almouloud (2005), a apresentação de uma estrutura interessante e ainda significativa do ponto de vista estatístico.

2.1. Resultados das situações antes da formação sobre estimação

Começamos nossas análises com a árvore de similaridades que foi formada por uma classe (Nível 14) e dividida em duas subclasses A (Nível 12) e B (Nível 13).

Figura 87 - Árvore de similaridades "Antes"



Fonte: Do autor (2018).

Na subclasse B encontramos apenas um nó de similaridade ao Nível 11 com um índice inferior a 0,70. Essa subclasse é composta pelas partições a_Positividade_1 e a_Positividade_2 que são zonas de comportamento onde os estudantes possuem um índice de positividade baixo em relação a estatística. É possível observar que apenas aS2_2, que representa a variável "amostra", possui uma fraca relação com essas partições. Percebemos, então, que os estudantes com um nível de positividade baixo em relação a estatística, identificaram a amostra usando a representação de um esquema usual para esse conceito, mas não expressaram a palavra "amostra", como revela a Figura 88.

Figura 88 - Resolução das variáveis aS1_1 e aS2_2 entre os estudantes com os menores índices de positividade



Fonte: Do autor (2018).

A subclasse B indicou uma tendência: os estudantes que estão situados nas zonas de comportamento com níveis mais baixos de positividade não obtiveram sucesso expressivo nas demais soluções que as situações-problema propuseram ou mobilizaram implicitamente conceitos-em-ação como o da amostra. Não foi possível perceber indícios de teoremas-em-ação, uma vez que não utilizaram relações matemáticas para justificar suas respostas.

Na subclasse A encontramos 10 nós de similaridades, aos níveis 1, 2, 3, 4, 5, 6, 7, 8, 9 e 10, sendo que os níveis 1, 3, 5 e 10 foram considerados mais significativos. Essa subclasse foi composta pelas partições a_Positividade_3 e a_Positividade_4 que são zonas de comportamento onde os estudantes possuem um índice de positividade maior em relação a estatística. No entanto, apenas a partição a_Positividade_4 contribuiu para formação dos nós mais significativos da árvore (Anexo Y), como mostra a Tabela 14.

Tabela 14 - Variáveis com maiores contribuições aos nós com níveis superiores a 0,70.

<i>Nós com índice superior a 0,70</i>	<i>Variável com maior contribuição</i>
Nível 1 (índice 1)	a_Positividade_4
Nível 2 (índice 0,99)	a_Positividade_4
Nível 3 (índice 0,99)	a_Positividade_4
Nível 4 (índice 0,99)	a_Positividade_4
Nível 5 (índice 0,99)	a_Positividade_4
Nível 6 (índice 0,98)	a_Positividade_4
Nível 7 (índice 0,89)	a_Positividade_4
Nível 8 (índice 0,81)	a_Positividade_3
Nível 9 (índice 0,79)	a_Positividade_4
Nível 10 (índice 0,72)	a_Positividade_4

Fonte: Do autor (2018).

Analizamos os nós mais significativos:

Nível 1 {a_Positividade_4, aS3_2} – Uma forte similaridade entre os estudantes situados nessa zona de comportamento com a justificativa coerente para possíveis erros nos padrões estabelecidos para as situações. Encontramos justificativas como: “Peguei uma amostra, então não posso afirmar que todos os pacotes de maçãs sigam o padrão que elaborei”; “No caso dos carros, depende muito do tamanho da amostra e do tempo que ficamos observando”.

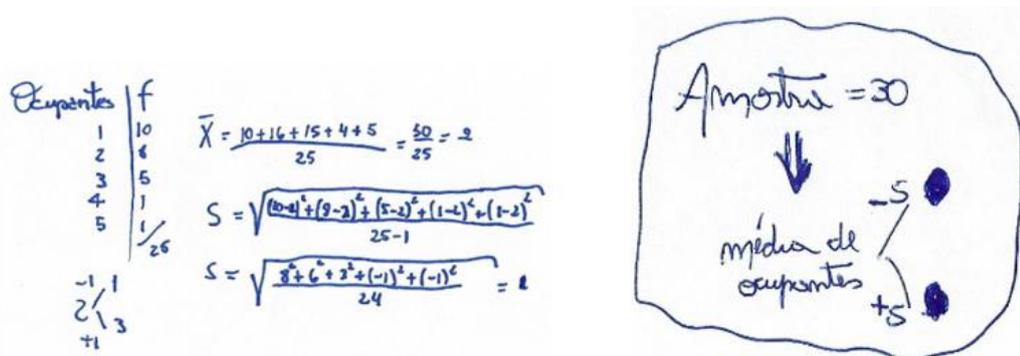
Nível 3 {{aS1_2A, aS1_2B}, aS2_1} – Indica uma tendência: os estudantes que identificaram a média e o desvio-padrão na situação 1, também identificaram a variabilidade do elemento na situação 2.

Nível 5 {{a_Positividade_4 aS3_2}, {{aS1_2A aS1_2B} aS2_1}} – Indica uma tendência: os estudantes que estão nessa zona de comportamento e que justificaram possíveis erros nos padrões estabelecidos, identificaram a média e o desvio-padrão da situação 1 e a variabilidade da situação 2.

Nível 10 {{{{a_Positividade_4 aS3_2} {{aS1_2A aS1_2B} aS2_1}} {aS2_3A aS2_3B}} {aS1_1 aS3_1}} aS2_3C} – Indica uma tendência: os estudantes que estão nessa zona de comportamento concordaram e justificaram os possíveis erros nos padrões estabelecidos, identificaram a amostra, a média e o desvio-padrão da situação 1, identificaram a variabilidade, a média, o desvio-padrão e estabeleceram um padrão para a situação 2.

Destacamos o Nível 10 que mostra as relações de similaridades entre a partição a_Positividade_4 com quase todas as variáveis analisadas. Os estudantes situados nessa zona de comportamento interpretaram as situações em um contexto mais científico, fazendo o uso de elementos conceituais adequados, como mostra a Figura 89.

Figura 89 - Respostas encontradas na primeira situação entre os estudantes com maiores índices de positividade



Fonte: Do autor (2018).

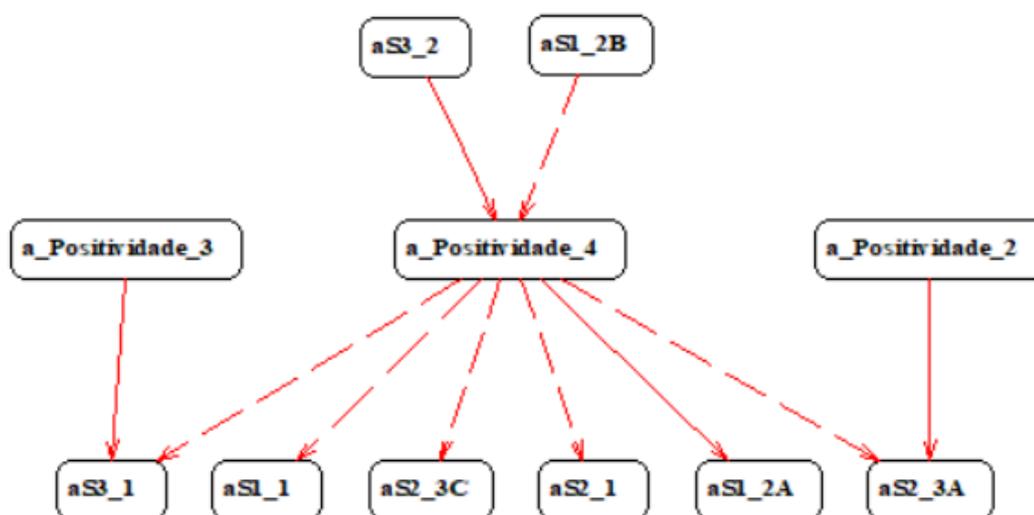
Dessa forma, podemos interpretar que os estudantes que se encontram nessa zona de comportamento possuem uma forte tendência na compreensão da linguagem estatística que as

situações-problema apresentaram, utilizando teoremas e conceitos em ação adequados ao contexto. Em linhas gerais, os dois tipos de invariantes estão imbricados dialeticamente na ação desses estudantes, onde, segundo Vergnaud (1990) os conceitos-em-ação permitiram retirar do meio as informações pertinentes e selecionar os teoremas-em-ação necessários para o desenvolvimento dos cálculos.

Ainda sobre a árvore de similaridades da Figura 87, foi interessante observar que houve uma nítida separação entre os estudantes que possuem um índice de positividade baixo com os que possuem um índice de positividade maior. Esse fato nos leva a deduzir que existe uma influência do comportamento do estudante em relação a estatística com as respostas do questionário, validando nossa segunda hipótese da pesquisa.

Construímos, também, o grafo implicativo observando as implicações centradas nas partições (Figura 90), com índice de implicação 0,70.

Figura 90 - Grafo implicativo "Antes"

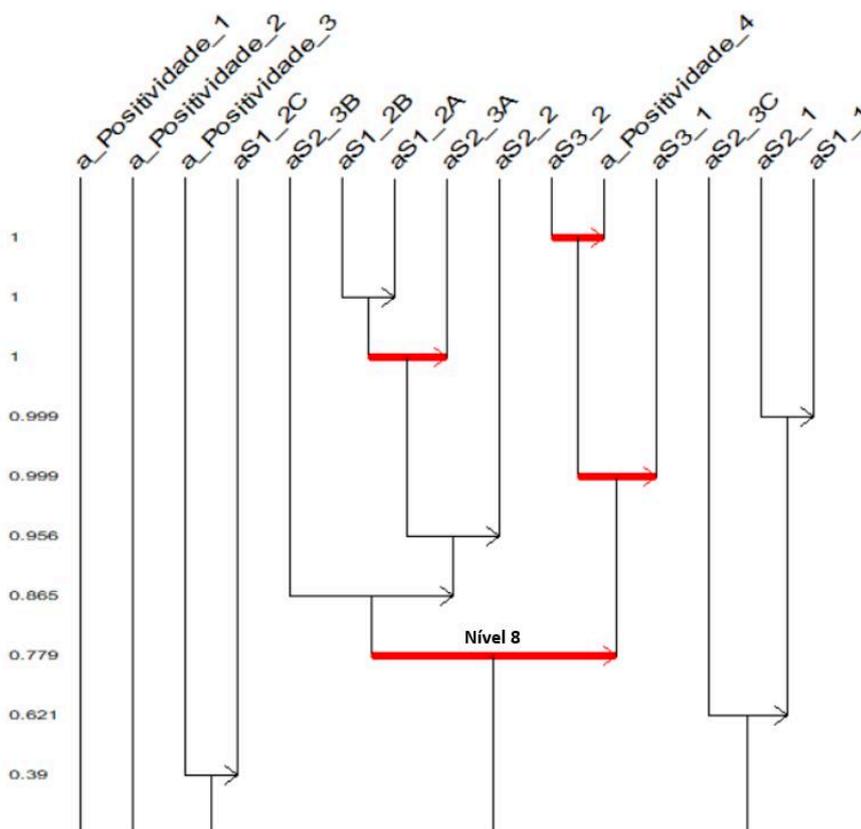


Fonte: Do autor (2018).

É possível perceber que a partição `a_Positividade_1` não foi representada no grafo, isso por que os estudantes que estão nessa zona de comportamento não obtiveram êxito relevante nas soluções das situações-problema. Podemos identificar que a partição `a_Positividade_4` possui implicações com grande parte das variáveis, como foi visualizado na árvore de similaridades. Os estudantes que possuem um índice de positividade mais elevado recorreram aos conceitos-em-ação e teoremas-em-ação adequados para as soluções das situações-problema. Nesse sentido, diante das informações até aqui apuradas, compreendemos que os estudantes com um maior índice de positividade mostraram uma ligação entre os componentes

cognitivos (aS1_1, aS1_2A, aS1_2B, aS2_3C, aS2_1) com os componentes de disposição (aS3_1 e aS3_2), que, segundo Gal (2002), precisam estar correlacionados para o desenvolvimento do letramento estatístico.

Figura 91 - Árvore coesiva "Antes"



Fonte: Do autor (2018).

Além da árvore de similaridades o CHIC nos forneceu a árvore coesiva (Figura 91) que reforçou nossas observações anteriores. De acordo com Couturier *et al.* (2003) a árvore coesiva apresenta uma janela com resultados numéricos contendo índices decrescentes das coesões e a árvore, indicando os níveis de coesão das classes. A árvore apresenta as implicações orientadas dentro de cada classe, que traduz a noção de regras sobre regras.

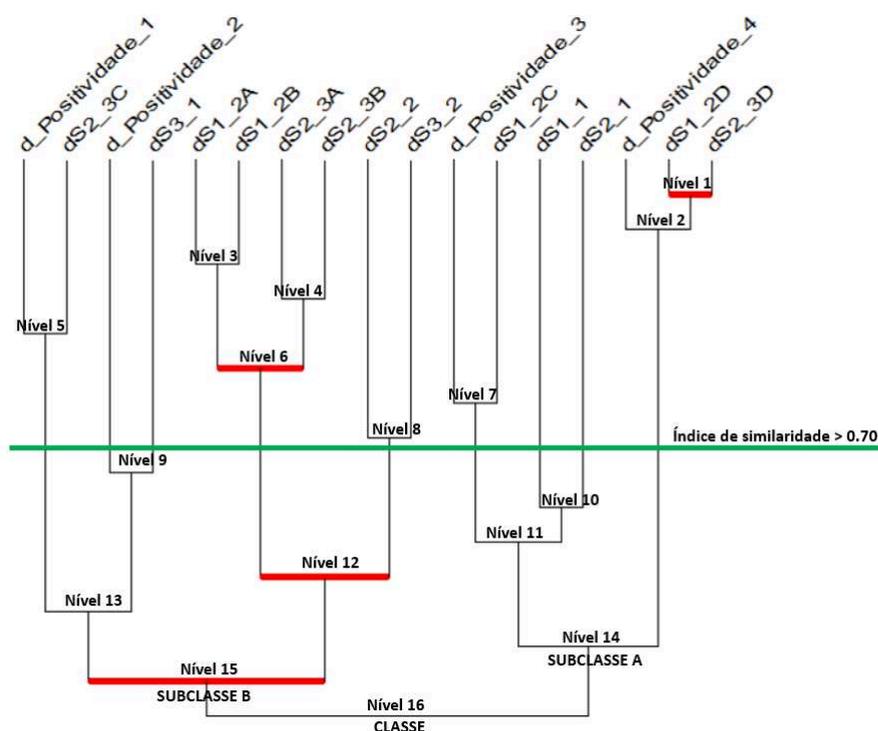
Começamos nossa análise examinando as variáveis que não implicam e não foram implicadas por nenhuma outra variável, como é o caso de a_Positividade_1 e a_Positividade_2, onde não houve uma mobilização expressiva dos invariantes operatórios. Uma relação de coesão significativa é representada pelo nó existente entre aS3_2 e a_Positividade_4, que pode ser interpretada como: uma postura crítica em relação aos resultados das situações-problema é satisfeita pelos estudantes com maiores índices de positividade em relação a estatística. Observamos a relação entre as variáveis aS1_2B e aS1_2A que pode ser interpretada como: os

estudantes que identificaram o desvio padrão, também identificaram a média. Destacamos o Nível 8 da árvore coesiva que revela um nó significativo, com um índice de coesão de 0,779, entre as diversas variáveis observadas nas situações-problema implicadas com a variável *a_Positividade_4*. Assim sendo, os estudantes com os maiores índices de positividade empregaram corretamente os conceitos de amostra, média, variabilidade e desvio-padrão, além de mobilizarem as relações matemáticas para justificar suas respostas. De fato, podemos perceber que o nível de positividade influenciou nas respostas dadas pelos estudantes.

2.2. Resultados das situações depois da formação sobre estimação

Nossa análise começa novamente pela árvore de similaridades (Figura 92), que agora foi formada por uma classe (Nível 16) e dividida em duas subclasses A (Nível 14) e B (Nível 15).

Figura 92 - Árvore de similaridades "Depois"



Fonte: Do autor (2018).

Mais uma vez, foi possível observar uma separação entre os estudantes que possuem um índice de positividade baixo (subclasse B) com os que possuem um índice de positividade maior (subclasse A). No entanto, na Subclasse B, onde encontram-se as zonas *d_Positividade_1* e *d_Positividade_2*, apresenta um maior número de nós de similaridades em relação as variáveis

analisadas anteriormente. Isso mostra que houve uma tendência na melhora quanto ao desenvolvimento do letramento estatístico por parte dos estudantes, uma vez que as variáveis que mais contribuíram na formação dos nós de similaridades (Anexo Z) foram melhor distribuídas como mostra a Tabela 15.

Tabela 15 - Variáveis com maiores contribuições aos nós com níveis superiores a 0,70

<i>Nós com índice superior a 0,70</i>	<i>Variável com maior contribuição</i>
Nível 1 (índice 0,99)	d_Positividade_4
Nível 2 (índice 0,99)	d_Positividade_4
Nível 3 (índice 0,96)	d_Positividade_4
Nível 4 (índice 0,93)	d_Positividade_1
Nível 5 (índice 0,88)	d_Positividade_1
Nível 6 (índice 0,85)	d_Positividade_2
Nível 7 (índice 0,83)	d_Positividade_3
Nível 8 (índice 0,71)	d_Positividade_4

Fonte: Do autor (2018).

Análise dos nós com índice de similaridade superior a 0,70:

Nível 1 {dS1_2D dS2_3D} – Indica uma tendência: os estudantes que utilizaram elementos da estimação na situação 1, também empregaram na situação 2.

Nível 2 {d_Positividade_4 {dS1_2D dS2_3D}} – Indica uma tendência: os estudantes que apresentaram os maiores índices de positividade, identificaram elementos da estimação na situação 1 e na situação 2.

Nível 3 {dS1_2A dS1_2B} – Indica uma tendência: os estudantes que mobilizaram o conceito de média, também mobilizaram o conceito do desvio-padrão na situação 1.

Nível 4 {dS2_3A dS2_3B} – Indica uma tendência: os estudantes que mobilizaram o conceito de média, também mobilizaram o conceito do desvio-padrão na situação 2.

Nível 5 {d_Positividade_1, dS2_3C} – Indica uma tendência: os estudantes que estão nessa zona de comportamento mobilizaram os conceitos-em-ação e os teoremas-em-ação adequados para a situação 2.

Nível 6 {{dS1_2A dS1_2B} {dS2_3A dS2_3B}} – Indica uma tendência: os estudantes que mobilizaram os conceitos de média e desvio-padrão na situação 1, também mobilizaram na situação 2.

Nível 7 {d_Positividade_3, dS1_2C} – Indica uma tendência: os estudantes que estão nessa zona de comportamento mobilizaram os conceitos-em-ação e os teoremas-em-ação adequados para a situação 1.

Nível 8 {dS2_2, dS3_2} – Indica uma tendência: os estudantes que mobilizaram o conceito de amostra, concordaram com possíveis erros nos padrões estabelecidos.

Constatamos que os invariantes operatórios foram melhor mobilizados em todas as zonas de comportamentos, o que antes era característica apenas do grupo de estudantes com os maiores índices de positividade. Os estudantes, de modo geral, empregaram os conceitos-em-ação e os teoremas-em-ação adequados, justificando as respostas por meio da explicitação dos conceitos de média e desvio-padrão, por exemplo, além do uso das relações matemáticas existentes. Entretanto, destacamos, na Subclasse A, o Nível 2 (d_Positividade_4, dS1_2D e dS2_3D), onde mostra que os estudantes com os maiores índices de positividade mobilizaram os invariantes operatórios relativos a estimação, como mostra a Figura 93.

Figura 93 - Invariantes operatórios observados sobre a estimação

$$\begin{aligned}
 \text{amostra} &= 30 \text{ carros} \\
 \text{média} &= 2 \text{ por carro} \\
 s &= 1 \\
 \text{nível de confiança } &95\% \\
 1,96 \cdot \frac{s}{\sqrt{n}} &= 1 \\
 \sqrt{n} &= 30
 \end{aligned}$$

Fonte: Do autor (2018).

Essa constatação sinalizou que os estudantes situados nessa zona de comportamento possuem uma tendência em utilizar novos conhecimentos de forma mais satisfatória que os demais, indicando que a relação do estudante com a estatística pode influenciar no desenvolvimento do raciocínio estatístico.

Construímos o grafo implicativo observando as implicações centradas nas partições (Figura 94), com índice de implicação 0,70.

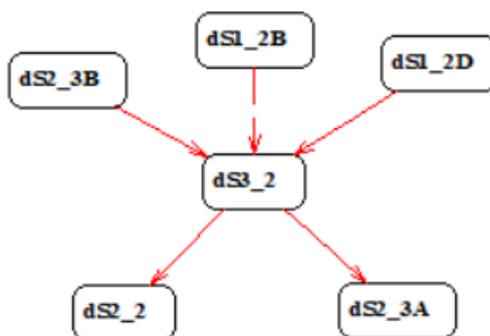
Figura 94 - Grafo implicativo "Depois"



Fonte: Do autor (2018).

Dessa vez, diferente do grafo da Figura 90, todas as partições foram representadas. De modo geral, as situações foram melhor desenvolvidas em todas as zonas de comportamento. Em d_Positividade_4 é possível observar, novamente, que os estudantes situados nessa zona de comportamento obtiveram implicações com os invariantes operatórios advindos da estimação, reforçando a interpretação de que a relação dos estudantes com a estatística pode influenciar no desenvolvimento do letramento estatístico. O grafo implicativo da Figura 95 foi construído observando as implicações centradas na variável dS3_2, que buscou averiguar nas repostas dos estudantes justificativas que remetesse para a conceptualização da estimação.

Figura 95 - Grafo implicativo em relação a variável dS3_2



Fonte: Do autor (2018).

De acordo com as implicações, verificamos que os estudantes que identificaram a média, o desvio-padrão, a amostra e ocuparam em suas respostas elementos da estimação, conseguiram justificar os possíveis erros nas simulações realizadas nas situações-problema, evocando uma conceptualização para a estimação: “Para verificar a quantidade de maçãs nos pacotes, ocupei uma amostra, fiz a média, o desvio padrão e um intervalo de confiança para a média”; “Realizei uma estimativa para a média, calculando o erro amostral para os carros e para os pacotes de maçãs”; “A estimativa pontual pode não ser confiável, então encontrei o erro amostral com um nível de confiança para estabelecer um intervalo de confiança.”.

Considerando os posicionamentos dos estudantes, elaboramos um esquema com os elementos que foram empregados para o desenvolvimento da conceptualização da estimação estatística, como mostra a Figura 96.

Figura 96 - Formação de um possível campo conceitual para a estimação



Fonte: Do autor (2018).

Contudo, o uso do software CHIC nos proporcionou uma visão global sobre a influência das atitudes dos estudantes com a estatística em relação as repostas encontradas no questionário Q2. Com a ajuda da Teoria dos Campos Conceituais observamos que a mobilização dos invariantes operatórios mais adequados para a resolução das situações-problema foram empregados, principalmente, por estudantes que possuem um índice de positividade mais elevado.

3. Considerações finais das hipóteses levantadas e prolongamentos

Diante das análises realizadas vislumbramos elementos que nos ajudaram nas validações das hipóteses levantadas para a pesquisa.

Figura 97 - H1

H1: Existem limitações na formação do raciocínio estatístico, na estimação estatística, ligadas aos procedimentos e conceitos básicos da Estatística Descritiva. ➡ **VALIDADA**

Fonte: Do autor (2018).

O grupo de estudantes investigados nessa pesquisa validou essa hipótese. Os estudantes, de acordo com a variável V03, reconhecem os elementos estatísticos em diversas situações de enfrentamento. Contudo, em sua maioria, indicam dificuldades em usar os conhecimentos estatísticos para o exercício da abstração dessas situações (V04). Ao enfrentarem as situações-problema oferecidas na primeira aplicação de Q2, observamos que apenas um grupo (a_Positividade_4) apresentou uma propensão para o sucesso das soluções, mobilizando invariantes operatórios adequados. Essas soluções abarcaram conceitos básicos da estatística descritiva, como o de amostra, média, desvio padrão, além de seus processos operatórios. Dessa forma, sem antes haver uma retomada desses conceitos e procedimentos operatórios, os estudantes encontrariam obstáculos cognitivos para o avanço do raciocínio estatístico com o novo conteúdo de estimação estatística. Nesse sentido, apesar de uma melhora importante observada na segunda aplicação de Q2, muitos estudantes ainda carregam estigmas decorrentes de uma base estatística enfraquecida advinda, possivelmente, da educação básica.

Figura 98 - H2

H2: A relação afetiva do estudante com a Estatística pode  **VALIDADA** influenciar na formação do raciocínio estatístico.

Fonte: Do autor (2018).

O grupo de estudantes investigados nessa pesquisa validou essa hipótese. Os estudantes, de acordo com V07, possuem uma tendência positiva em atribuir um bom nível de interesse pessoal em relação aos conhecimentos estatísticos. Entendem que essa formação é importante para a sua carreira profissional (V08). Contudo, a variável V05 apresentou indícios de uma relação não tão cordial com a estatística, mostrando que muitos estudantes optam em cursar essa disciplina “mais tarde”. Além disso, nosso instrumento de mensuração das atitudes positivas e negativas em relação a estatística observou, antes da formação sobre estimação, quatro grupos com extremidades discordantes. As atitudes podem auxiliar ou prejudicar a aprendizagem de estatística, podendo afetar o desenvolvimento do raciocínio estatístico. Os grupos com os menores índices de positividade (a_Positividade_1 e a_Positividade_2) estiveram sempre vinculados a manifestações de baixo interesse ou não capazes de solucionar os problemas estatísticos apresentados em Q2. O reforço dos conceitos e procedimentos básicos da estatística descritiva, indicou uma melhora em relação as atitudes e, quanto melhor forem essas atitudes, mais os estudantes estarão propensos a se aproximar da estatística. Depois do conteúdo de estimação, apesar da melhora, ainda assim, os que apresentaram índices de

positividades mais elevadas fizeram o uso do novo conteúdo na resolução das situações-problema, avançando no desenvolvimento do raciocínio estatístico.

Figura 99 - H3

H3: A relação afetiva do estudante com a Matemática pode  **VALIDADA**
influenciar na formação do raciocínio estatístico.

Fonte: Do autor (2018).

O grupo de estudantes investigados nessa pesquisa validou essa hipótese. Os estudantes, de acordo com V10, possuem uma tendência em identificar termos estatísticos. Reconhecem as utilidades da estatística (V11) aplicadas em diversas situações. Contudo, não basta a compreensão e aplicação dos conceitos estatísticos para desenvolver atitudes positivas em relação à estatística, mas também seus processos operatórios. A matemática, ou elementos da matemática, foi ostensiva quando os estudantes foram indagados sobre o que lhes invocavam a estatística (V06) ou sobre as dificuldades que encontravam em uma disciplina de estatística (V12). Os estudantes que entendem a estatística como elemento da matemática e, se suas vivências com a matemática for algo que provoca frustrações, eles penderão a demonstrar esses desencantamentos em relação à estatística, acarretando atitudes negativas. A adesão por parte dos estudantes nas variáveis das situações-problema que solicitavam procedimentos operatórios foi acanhada. A exceção ocorreu nos grupos (antes e depois) com os maiores índices de positividade. Diante disso, constatamos uma acentuada tendência de que as atitudes dos estudantes em relação a estatística podem despontar das atitudes em relação a matemática, influenciando no desenvolvimento do raciocínio estatístico.

3.1. Prolongamentos da pesquisa

Nosso objetivo geral era melhor entender o processo de formação do raciocínio estatístico na aprendizagem da estimação estatística no Ensino Superior. Para tanto, aplicamos questionários em estudantes de graduação que nos ajudaram a levantar informações para observar a formação do raciocínio estatístico. Tratamos nossos dados com uma ferramenta estatística no quadro A.S.I. e utilizamos o software SPAD para as análises descritivas. Observamos um campo conceitual para a estimação estatística e procuramos na Teoria dos Campos Conceituais, subsídios para analisar as respostas dos estudantes.

Verificamos a validação das hipóteses levantadas na pesquisa, avistando limitações na formação do raciocínio estatístico, por meio da estimação, ligadas aos procedimentos e

conceitos básicos da estatística descritiva. Ainda, observamos que o relacionamento dos estudantes com a estatística pode influenciar na formação do raciocínio estatístico, salientando a matemática como elemento principal entre as dificuldades apontadas pelos estudantes.

Muitos foram os dados coletados para a elaboração do trabalho, de onde inúmeras informações poderiam ser extraídas, contudo, apresentamos as que nos pareceram ser mais relevantes para dar início a essa pesquisa que não ficará inerte. Dessa forma, apontamos algumas limitações que pretendemos suprir e alguns seguimentos que farão parte da continuidade da nossa vida científica acadêmica:

- 1) Nossa pesquisa tratou sobre as atitudes dos estudantes em relação a estatística, contudo, se limitou a fazer uma verificação pontual, sem buscar subsídios na literatura que fundamentassem melhor essas ocorrências. Assim sendo, vamos investigar as definições e relações de afetividade, os componentes pedagógicos e antropológicos das atitudes no ensino da estatística.
- 2) Limitamos o uso da A.S.I., por meio do software CHIC, em apenas um questionário, sendo assim, iremos explorar as demais variáveis, procurando novas abordagens e estimulando o desenvolvimento do campo teórico da análise estatística implicativa.
- 3) O software SPAD é uma poderosa ferramenta estatística computacional, apropriado para implementar uma série de métodos na análise de dados. Nosso uso foi modesto, o que nos instiga a realizar uma melhor exploração sobre seus comandos e funções.
- 4) Limitamos a investigação tomando como base o conteúdo de estimação estatística, o que nos motiva a explorar novos conteúdos. Assim sendo, os questionários continuarão sendo aplicados em novas turmas de estatística, levantando novos dados e procurando, em nosso trabalho docente, melhor atender aos alunos por meio da apuração de suas respostas.

Acreditamos que uma pesquisa não tem fim. Nosso interesse é promover novas discussões e contribuir para a melhoria do ensino, em especial, na área da Educação Estatística.

CONCLUSÃO GERAL

O raciocínio estatístico, conectado com o pensamento estatístico, é responsável pela expansão do letramento estatístico. Sobre suas especificidades de desenvolvimento, por se tratar da compreensão de uma ciência, o estudante deve possuir ferramentas básicas da estatística advindas de sua escolarização. A evolução das competências estatísticas depende de um conjunto de processos mentais que articulam as ferramentas estatísticas para executar cálculos e interpretar esses resultados. Nesse sentido, a matemática, ou o raciocínio matemático, visa proporcionar atributos operacionais e deduções lógicas, enquanto o raciocínio estatístico versa sobre as interpretações subjetivas ligadas, frequentemente, por algum conceito matemático preterido por análises e relações abstratas da linguagem estatística. Dessa forma, a relação fronteira entre a matemática e a estatística faz com que os estudantes não as desvinculem, tornando a estatística um assunto dependente da qualidade da relação dos estudantes com a matemática. Por conseguinte, considerando as análises realizadas, muitas das atitudes negativas em relação a estatística, foram promovidas em função das atitudes com a matemática, ocorrendo perturbações no desenvolvimento do raciocínio estatístico.

Nessa perspectiva, entendemos que para haver a promoção do desenvolvimento do raciocínio estatístico, os estudantes devem possuir algumas referências como mostra a Figura 100.

Figura 100 - Referências para o desenvolvimento do raciocínio estatístico



Fonte: Do autor (2018).

Os dados são números que estão contextualizados para uma determinada situação e servirão para orientar as interpretações dos resultados. Diante disso, verifica-se o melhor processo estatístico a ser empregado para a manipulação dos dados. As operações matemáticas estão inseridas na efetivação desses processos. Um bom reforço dessas operações, propiciam um avanço cognitivo em relação a estatística, diminuindo as atitudes negativas em relação a matemática, ainda que dentro do contexto do processo estatístico específico, promovendo, assim, uma aproximação do estudante com a estatística. No caso da estimação de parâmetros,

o reforço ajudou no domínio dos conceitos e procedimentos operacionais da média e desvio padrão, por exemplo, que foram fundamentais para observarmos um avanço em todas as zonas de comportamento construídas na pesquisa.

Os resultados obtidos de uma situação-problema estatística não devem ser confundidos com uma solução matemática. Geralmente, os problemas estatísticos apresentam um questionamento cujo resultado deve ser articulado por meio da manipulação ocorrida nos dados e pela interpretação particular do contexto da situação-problema. Essa subjetividade do resultado, não é tarefa fácil de ser desenvolvida nos estudantes, contudo, a clareza de que as respostas não são absolutas, promovem o conceito chave da estatística que é a variabilidade.

Ainda que muitos pesquisadores como Hand (1998), Snee (1988), Stuart (1995) e outros, apontem na defesa da simplificação da matemática na formação dos usuários da estatística, oferecendo recursos computacionais para o incremento dos cálculos, o que também utilizamos em nossa prática docente, não podemos deixar de reconhecer a presença desses obstáculos operacionais advindos da matemática, gerando incompletos usuários de estatística. Devemos incentivar e colaborar com a atenuação dessas dificuldades, provendo o raciocínio e o desenvolvimento da educação, em especial da Educação Estatística.

REFERÊNCIAS

- AAKER, D. A. **Marketing research**. 7. ed. New York: John Wiley & Sons, 2001.
- ACIOLY-RÉGNIER, N. M. **Culture et cognition**: domaine de recherche, champ conceptuel, cadre d'intelligibilité et objet d'étude fournissant des instruments pour conduire des analyses conceptuelles et méthodologiques en psychologie et en sciences de l'éducation. Note de synthèse HDR - Université Lumière Lyon 2 – septembre 2010.
- ACIOLY-RÉGNIER, N. M.; RÉGNIER, J.-C. Repérage d'obstacles didactiques et socioculturels au travers de l'A.S.I. des données issues d'un questionnaire. *In*: GRAS, R.; SPAGNOLO, F.; DAVID, J. (Coords.). **Proceedings Third International Conference A.S.I.** Palerme: Implicative Statistic Analysis, 2005. p. 63-87.
- AGAMBEN, G. **Che cos'è un dispositivo?**. Milano: Nottetempo, 2006.
- AGUIAR JR., O. G. A. **Modelo de ensino para mudanças cognitivas**: um instrumento para o planejamento do ensino e a avaliação da aprendizagem em ciências. 2001. 377f. Tese (Doutorado em Educação) - Universidade Federal de Minas Gerais, Belo Horizonte, 2001.
- ANDRADE, V. L. V. X. **Os conceitos de medidas de tendência central e de dispersão na formação estatística no ensino médio no Brasil e na França**. Abordagem exploratória no quadro da teoria antropológica do didático e da Teoria dos Campos Conceituais. 2013. 552f. Tese (Doutorado em Ensino de Ciências) - Universidade Federal Rural de Pernambuco, Recife, 2013.
- ANDRADE, V. L. V. X.; RÉGNIER, J.-C. A formação do licenciando em matemática para o ensino de estatística na educação básica: problemas e desafios. *In*: SEMINÁRIO INTERNACIONAL DE EDUCAÇÃO MATEMÁTICA, 2., 2009, São Paulo. **Anais [...]**. São Paulo: UNIBAN, 2009.
- ARDILLY, P. **Echantillonnage et methodes d'enquetes**. 2. ed. Paris: Dunod, 2004.
- ASSIS, L. B. **A formação do usuário de estatística pelo desenvolvimento da literacia estatística, do raciocínio estatístico e do pensamento estatístico através de atividades exploratórias**. 2015. 88f. Dissertação (Mestrado Profissional em Educação Matemática) - Universidade Federal de Juiz de Fora, Juiz de Fora, 2015.
- BACHELARD, G. **A filosofia do não**: o novo espírito científico: a poética do espaço. São Paulo: Abril Cultural, 1978. (Os pensadores).
- _____. **A formação do espírito científico**: contribuição para uma psicanálise do conhecimento. Tradução de Estela dos Santos Abreu. Rio de Janeiro: Contraponto, 1996.
- _____. **La formation de L'esprit scientifique**. Paris: Vrin, 1989.
- BARBETTA, P. A. **Estatística aplicada às ciências sociais**. 6. ed. Florianópolis: UFSC, 2006.

BARBETTA, P. A.; REIS, M. M.; BORNIA, A. C. **Estatística para cursos de engenharia e informática**. 3. ed. São Paulo: Atlas, 2010.

BARCELÓ, S.; CAPILLA, C.; ZÚNICA, L. Estatística e probabilidades no ensino superior: diagnóstico de competências e propostas curriculares globais. *In*: LOUREIRO, C.; OLIVEIRA, F.; BRUNHEIRA, L. (Ed.). **Ensino e aprendizagem da estatística**. Lisboa: SPE e APM, 2000. p. 83-92.

BATANERO, C. **Didáctica de la estadística**. Granada: Grupo de investigación em Educación Estadística do Departamento de Didáctica de la Matemática da Universidad de Granada, 2001.

BATANERO, C. *et al.* **Investigación en educación estadística**: algunas cuestiones prioritarias. **Statistical Education Research Newsletter**, v. 1, n. 2, p. 1-31, 2000.

BELLHOUSE, D. R. De vetula: a medieval manuscript containing probability calculations. **International Statistical Review**, v. 68, n. 2, p. 123-136, 2000.

BERLIKOWSKI, M. E. **Análise das atitudes e imagem em relação à estatística**: um estudo comparativo com alunos da graduação. 2012. 63f. Dissertação (Mestrado em Educação em Ciências e Matemática) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2012.

BERNSTEIN, B. **A estruturação do discurso pedagógico**: classe, códigos e controle. Petrópolis: Vozes, 1996.

_____. **Class, codes and control**: theoretical studies towards Sociology of Language. London: Routledge & Kegan Paul, 1971.

BESSON, J.-L. **A ilusão das estatísticas**. São Paulo: Unesp, 1995.

BOYER, C. B. **História da matemática**. Tradução de Elza F. Gomide. São Paulo: Edgard Blücher, 1974.

_____. **História da matemática**. 2. ed. São Paulo: Edgard Blücher, 2003.

BRITO, M. O papel das competências, das habilidades e das atitudes na aprendizagem e no ensino de Matemática e estatística. *In*: BRITO, Márcia. **Mapeamento de pesquisas em educação matemática**. Brasília: INEP, 1994. v. 1, p. 17-18.

BROUSSEAU, G. Que peut-on enseigner en mathématiques à l'école primaire et pourquoi?. **Repère IREM**, b. 38, p. 7-10, 2000.

_____. **Théorisation des phénomènes d'enseignement de mathématiques**. Bordeaux: Université de Bordeaux I, 1986.

CAI, J. Exploring students' conceptual understanding of the averaging algorithm. **School Science and Mathematics**, v. 98, n. 2, p. 93-98, fev. 1998.

CAMPOS, C. R. **A educação estatística: uma investigação acerca dos aspectos relevantes à didática da estatística em cursos de graduação.** 2007. 242f. Tese (Doutorado em Educação Matemática) - Universidade Estadual Paulista, Rio Claro, 2007.

CARZOLA, I. M. **A relação entre a habilidade viso-pictórica e o domínio de conceitos estatísticos na leitura de gráficos.** 2002. 335f. Tese (Doutorado em Educação) - Universidade Estadual de Campinas, Campinas, 2002.

CHANCE, B. L. Components of statistical thinking and implications for instruction and assessment. **Journal of Statistics Education**, v. 10, n. 3, p. 1-14, nov. 2002.

CHAVALLARD, Y. **La transposition didactique: du savoir savant au savoir enseigné.** Paris: La Pensee Sauvage, 1991.

COSTA, C. A. **Educação matemática nos cursos superiores de tecnologia: revelações sobre a formação estatística.** 2013. 273f. Tese (Doutorado em Educação) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2013.

COSTA, G. D. F. **A metodologia de projetos como uma alternativa para ensinar estatística no ensino superior.** 2012. 261f. Tese (Doutorado em Educação) - Universidade Estadual de Campinas, Campinas, 2012.

COURNOT, A. A. **Exposition de la théorie des chances et des probabilités.** Paris: Librairie de L. Hachette, 1843. Disponível em: <https://gallica.bnf.fr/ark:/12148/bpt6k285042.texteImage>. Acesso em: 12 fev. 2018.

COUTANSON, B. E. **La question de l'éducation statistique et de la formation de l'esprit statistique à l'école primaire en France: étude exploratoire de quelques caractéristiques de situations inductrices d'un enseignement de la statistique au cycle III.** 2010. 267f. Tese (Doutorado em Ciências da Educação) - Université Lumière - Lyon II, Lyon, 2010.

COUTINHO, C. Q. S. Conceitos probabilísticos: quais contextos a história nos aponta?. **Revista Eletrônica de Educação Matemática**, v. 2, n. 3, p. 50-67, 2007.

_____. Educação estatística e os livros didáticos para o ensino médio. **Revista Educação Matemática em Foco**, Campina Grande, v. 2, n. 1, p. 68-86, 2013.

COUTURIER, R. Traitement de l'analyse statistique dans CHIC. *In: COUTURIER, R. Dans caen: actes des journées: la fouille dans les données par la méthode d'analyse statistique implicative. Applications et traitement par CHIC.* Caen: IUFM, 2001. p. 33-50.

DALBÃO, F. C. **As contribuições do raciocínio estocástico para a formação do economista.** 2015. Tese (Doutorado em Ensino de Ciências e Matemática) – Universidade Cruzeiro do Sul, São Paulo, 2015.

DAVID, F. N. **Games, gods and gambling: the origins and history of probability and statistical ideas from the earliest times to the Newtonian era.** New York: Hafner Publishing Company, 1962.

DEHON, C. DROESBEKE, J.-J. ; VERMANDELE, C. **Eléments de statistique**. 5. ed. Bruxelles: Université de Bruxelles, 2008.

DELMAS, B. C. What makes the standard deviation larger or smaller?. In : DELMAS, B. C. **Statistics teaching and resource library**. Minneapolis: University of Minnesota, 2001.

DIAZ, D. **Les facteurs influençant la réussite des activités collaboratives médiées par les TICE dans une situation de formation universitaire à la statistique**. 2016. Tese (Doutorado em Ciências da Educação) - Université Lumière - Lyon II, Lyon, 2016.

DOWNING, D.; CLARK, J. **Estatística aplicada**. 2. ed. São Paulo: Saraiva, 2002.

DROESBEKE, J.-J. *et al.* **Histoire de la statistique**. 2. ed. Paris: Presses Universitaires de France, 1997.

ECHEVESTE, S. *et al.* Educação estatística: perspectiva e desafios. **Acta Scientiae**, v. 7, n. 1, p. 103-109, jan./jun. 2005.

FERREIRA, N. S. A. As pesquisas denominadas “Estado da Arte”. **Educação & Sociedade**, São Paulo, v. 23, n. 79, p. 258, ago. 2002.

FERREIRA, D. H. L.; JACOBINI, O. R. Modelagem matemática e ambiente de trabalho: uma combinação pedagógica voltada para aprendizagem. **Rencima**, v. 1, p. 9-26, 2010.

FOUCAULT, M. O sujeito e o poder. In: DREYFUS, H.; RABINOW, P. **Michel Foucault: uma trajetória filosófica (para além do estruturalismo e da hermenêutica)**. Rio de Janeiro: Forense Universitária, 1995. p. 231-149.

_____. Sobre a história da sexualidade. In: FOUCAULT, M. **Microfísica do poder**. Rio de Janeiro: Graal, 2000. p. 243-270.

_____. **Tecnologias del yo y otros textos afines**. Barcelona: Paidós, 1990.

FREUND, J. E. **Estatística aplicada: economia, administração e contabilidade**. 11. ed. Porto Alegre: Bookman, 2006.

GAL, I. Adult's statistical literacy: meanings, components, responsibilities. **International Statistical Review**, v. 70, n. 1, p. 1-25, 2002.

GAL, I.; GARFIELD. J. **The assessment challenge in statistic education**. Amsterdã: IOS Press, 1997.

GAL, I.; GINSBURG, L. role os beliefs and attitudes in learning statistic : towards an assessment framework. **Journal os Statistics Edcation**, v. 2, p. 64-89, 1994.

GARFIELD, J. The challenge of developing statistical reasoning. **Journal of Statistics Education**, v. 10, n. 3, 2002.

GARFIELD, J.; GAL, I. Teaching and assessing statistical reasoning. *In*: STIFF, L.; CURCIO, F. R. **Developing mathematical reasoning in grades K-12**. Reston: Ed. L. Staff, 1999. p. 207-219.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 5. ed. São Paulo: Atlas, 1999.

GITIRANA, V. *et al.* Média aritmética no ensino fundamental. *In*: LOPES, C.; COUTINHO, C. Q. S.; ALMOULOU, S. A. (Org.). **Estudos e reflexões em educação estatística**. Campinas: Mercado das Letras, 2010.

GNEDENKO, B. V. **A teoria da probabilidade**. Tradução da série de textos clássicos da American Mathematical Society. Rio de Janeiro: Ciência Moderna, 2008.

GODINO, J. D.; BATANERO, C.; CAÑIZARES, J. **Azar y probabilidad**: matemáticas: cultura y aprendizaje. Madrid: Editorial Síntesis, 1996.

GONÇALEZ, N. **Atitudes dos alunos do curso de pedagogia com relação à disciplina de estatística no laboratório de informática**. 2002. 191f. Tese (Doutorado em Educação) - Universidade Estadual de Campinas, Campinas, 2002.

GONSALVES, E. P. **Iniciação à pesquisa científica**. Campinas: Alínea, 2003.

GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. O ensino de estatística na UNESP/Campus de Marília. **Educação Matemática em Revista**, v. 11, n. 17, p. 9-15, 2005.

GRAS, R. L'analyse implicative: ses bases, ses développements. **Educação Matemática Pesquisa**, v. 4, n. 2, p. 11-48, 2004.

GRAS, R.; ALMOULOU, S. A. A implicação estatística usada como ferramenta em um exemplo de análise de dados multidimensionais. **Educação Matemática Pesquisa**, v. 4, n. 2, p. 75-88, 2002.

GRAS, R.; RÉGNIER, J.-C.; GUILLET, F. **Analyse statistique implicative**: une méthode d'analyse de données pour la recherche de causalités. Toulouse: Éditions Cépaduès, 2009.

GRAS, R. *et al.* **Analyse statistique implicative**: méthode exploratoire et confirmatoire à la recherche de causalités. Toulouse: Éditions Cépaduès, 2013.

GRAS, R. *et al.* **Analyse statistique implicative**: des Sciences dures aux sciences humaines et sociales. Toulouse: Éditions Cépaduès, 2017.

GRAY, D. E. **Pesquisa no mundo real**. 2. ed. Porto Alegre: Penso, 2012.

HAND, D. J. Breaking misconceptions – statistics and its relationship to mathematics. **The Statistician**, v. 47, n. 2, p. 245-250, 1998.

JUTAND, M.-A. **Études des phénomènes de transposition didactique de la statistique dans le champ universitaire et ses environnements**: une contribution à la pédagogie universitaire. 2015. 251f. Tese (Doutorado em Ciências da Educação) - Universidade Bordeaux Montaigne, Bordeaux, 2015.

- KATAOKA, V. Y. *et al.* Evidências de validade de uma prova de letramento estatístico: um estudo com estudantes universitários de cursos tecnológicos. **Bolema**, Rio Claro, v. 24, n. 40, p. 873-895, dez. 2011.
- KENDALL, M.; KENDALL, D.; STUART, A. (Ed.). **Statistic theory and practice**: selected papers. Oxford: Oxford University Press, 1997.
- LAPLACE, P.-S. **Essai philosophique sur les probabilités**. 6. ed. Paris: Christian Bourgeois Éditeur, 1840.
- LARANJEIRA, R. Estado da Arte do direito agrário no Brasil. *In*: SEMINÁRIO INTERNACIONAL DO DIREITO AGRÁRIO, 11., 2003, Maranhão. **Anais [...]**. Maranhão, UFMG, 2003.
- LARROSA, J. Tecnologias do eu e educação. *In*: SILVA, Tomaz Tadeu da. **O sujeito da educação**: estudos foucaultianos. Petrópolis: Vozes, 1995. p. 35-86.
- LERMAN, I. C. **Classification et analyse ordinale des données**. Paris: Dunod, 1981.
- LETHIELLEUX, M.; CHEVALIER, C. **Probabilités**: estimation statistique en 24 fiches. 5. ed. Paris: Dunod, 2016.
- LEVINE, D. M. *et al.* **Estatística**: teoria e aplicações usando Microsoft Excel em português. Rio de Janeiro: LTC, 2000.
- LIGHTNER, J. E. Um resumo da história da probabilidade e da estatística. Tradução de Antônio C. Patrocínio. **Mathematics Teacher**, nov. 1991.
- LOPES, C. A. E. **O conhecimento profissional dos professores e suas relações com estatística e probabilidade na educação infantil**. 2003. 290f. Tese (Doutorado em Educação) - Universidade Estadual de Campinas, Campinas, 2003.
- _____. Educação estatística no curso de licenciatura em matemática. **Bolema**, Rio Claro, v. 27, n. 47, p. 901-915, dez. 2013.
- MALLOWS, C. The zeroth problem. **American Statistician**, Washington, n. 52, p. 1-9, 1998.
- MANTOVANI, D. M. N.; VIANA, A. B. N. Ensino de estatística para cursos de graduação em administração de empresas – novas perspectivas. *In*: SEMINÁRIO DE ADMINISTRAÇÃO, 7., 2004, São Paulo. **Anais [...]**. São Paulo: FEA/USP, 2004.
- MARTINS, G. A. **Estatística geral e aplicada**. 3. ed. São Paulo: Atlas, 2008.
- MAYÉN GALICIA, S. A. M. **Comprensión de las medidas de tendencia central en estudiantes mexicanos de educación secundaria y bachillerato**. 2009. 376f. Tese (Doutorado em Matemática) - Universidade de Granada, Granada, 2009.
- MERINO, B. C. **Significados de las medidas de posición central para los estudiantes de secundaria**. 2003. 301f. Tese (Doutorado em Matemática) - Universidade de Granada, Granada, 2003.

MESNARD, J. **Blaise pascal**: aevres complètes. Paris: Desclée de Brouwer, 1991.

MESSINA, G. Estudio sobre el estado da arte de la investigacion acerca de la formación docente en los noventa. Organización de Estados IberoAmericanos para La Educación, La Ciência Y La Cultura. *In*: REÚNION DE CONSULTA TÉCNICA SOBRE INVESTIGACIÓN EN FORMACIÓN DEL PROFESSORADO, 1., 1998, México. **Anais [...]**. México: [S. n.], 1998.

MEYER, P. L. **Probabilidade**: aplicações à estatística. 2. ed. Rio de Janeiro: LTC 2011.

MONTGOMERY, D. C.; RUNGER, G. C. **Estatística aplicada e probabilidade para engenheiros**. 5. ed. Rio de Janeiro: LTC, 2012.

MONTUCLA, J.-F. Histoire des mathématiques. *In*: MONTUCLA, J.-F. **Tome troisième**. Paris: Blanchard, 1968. p. 375-426.

MOORE, D. S.; FARIAS, A. A. de. **A estatística básica e sua prática**. Rio de Janeiro: LTC, 2000.

MORETTIN, P. A.; BUSSAB, W. O. **Estatística básica**. 6. ed. São Paulo: Saraiva 2010.

_____. **Estatística básica**. 7. ed. São Paulo: Saraiva, 2011.

MOUNTASSIR, M. **Probabilités et statistique**. Canada: Modulo, 2016.

ORIOU, J.-C. **Formation à la statistique par la pratique d'enquêtes par questionnaires et la simulation**: étude didactique d'une expérience d'enseignement dans un département d'IUT. 2007. 278f. Tese (Doutorado em Ciências da Educação) - Université Lumière-Lyon II, Lyon, 2007.

PIAGET, J. **Psicologia e pedagogia**. Rio de Janeiro: Forense Universitária, 1985.

PICHARD, J.-F. Points d'histoire des probabilités. Instituts de Recherche sur l'Enseignement des Mathématiques – Commission Inter-IREM, Statistique et Probabilités, 1997.

PONCE, A. **Educação e luta de classes**. Tradução: José Severo de Camargo Pereira. 6. ed. São Paulo: Cortez/Autores Associados, 1986.

RÉGNIER, J.-C. À propos de la formation en statistique. Aprocches praxéologiques et épistémologiques de questions du champ de la didactique de la statistique. **Revue du Centre de Recherche en Education**, v. 22, p. 157-201, 2003.

_____. **Approche méthodologique de l'enquête par questionnaire & Statistique**. 2009. Tesis (Doctorat de Sciences de L'éducation) - Université Lumière Lyon 2, Lyon, 2009. 2008.

_____. Formation de l'esprit statistique et raisonnement statistique. Que peut-on attendre de la didactique de la statistique. *In*: RÉGNIER, J.-C. **Actes du séminaire national de didactique des mathématiques**. Paris: Association pour la Recherche en Didactique des Mathématiques, 2005. p. 13-38.

RÉGNIER, J.-C. **Quelques propôs autour de la statistique**. Département de Sciences de l'éducation. Université Lumière Lyon 2, Décembre, 1996.

_____. **Vocabulaire de statistique**. Lyon: ISPEF - Université Lumière Lyon 2, 2007.

RÉGNIER, J.-C.; BRAGA, E. M. Instrumentos estatísticos para uma leitura do mundo: formação do espírito estatístico e cidadania. **Conjectura: Filosofia e Educação**, Caxias do Sul, v. 13, n. 2, p. 9-41, 2008.

RÉGNIER, J.-C.; ALMOULOU, S. A.; GRAS, R. Analyse statistique implicative: cadre théorique et applicatif pour l'exploration sémantique et non symétrique des données. *In*: COLLOQUE INTERNATIONAL D'ANALYSE STATISTIQUE IMPLICATIVE, 7., 2013, São Paulo. **Anais [...]**. São Paulo: Université Lumière-Lyon II, 2013.

RÉGNIER, J.-C.; GRAS, R.; BAILLEUL, M. L'analyse statistique implicative: de l'exploratoire au confirmatoire. *In*: COLLOQUE INTERNATIONAL D'ANALYSE STATISTIQUE IMPLICATIVE, 6., 2012, Caen. **Anais [...]**. Caen: Université Lumière-Lyon II, 2012.

ROMANOWSKI, J. P.; ENS, R. T. As pesquisas denominadas do tipo "Estado da Arte" em educação. **Revista Diálogo Educacional**, v. 6, n. 19, p. 37-50, set./dez. 2006.

RUMSEY, D. J. Statistical literacy as goal for introductory statistics courses. **Journal of Statistics Education**, v. 10, n. 3, 2002.

SALSBURG, D. **Uma senhora toma chá...**: como a estatística revolucionou a ciência no século XX. Rio de Janeiro: Zahar, 2009.

SANTOS, G. R. **A educação estatística no ensino superior, o trabalho com projetos e o uso de tecnologias**. 2003. (Mestrado em Ensino de Ciências) - Universidade Cruzeiro do Sul, São Paulo, 2003.

SANTOS, R. F. V. S. **Probabilidade circa 1914 e a construção de Pacheco d'Amorim**. 2008. 769f. Tese (Doutorado em Estatística e Investigação Operacional) - Universidade de Lisboa, Lisboa, 2008.

SEDLMEIER, P. **Improving statistical reasoning**: theoretical models and practical implication. Mahwah: Springer Verlag, 1999.

SELZ, M.; MAILLOCHON, F. **Le raisonnement statistique en sociologie**. Paris: Presses Universitaires de France, 2009.

SGOUTAS-EMCH, S. A.; JOHNSON, C. J. Is journal writing an effective method of reducing anxiety towards statistics? **Journal of Instructional Psychology**, v. 25, n. 1, p. 49-57, mar. 1998.

SILVA, C. B. **Pensamento estatístico e raciocínio sobre variação**: um estudo com professores de matemática. 2007. 354f. Tese (Doutorado em Educação Matemática) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2007.

SILVA JR., G. B. **O ensino de estatística na formação inicial do engenheiro de produção**. 2014. 217f. Tese (Doutorado em Ensino de Ciências e Matemática) - Universidade Cruzeiro do Sul, São Paulo, 2014.

SNEE, R. D. Mathematics is only one tool that statistician use. **The College Mathematics Journal**, v. 19, p. 30-32, 1988.

SOARES, M. Letramento e alfabetização: as muitas facetas. **Revista Brasileira de Educação**, n. 25, p. 5-17, 2004. 2003

SOARES, M. B.; MACIEL, F. **Alfabetização/Organização**. Brasília: MEC/Inep/Comped, 2000.

SOUZA JR., A. J.; CAMPOS, S. G. V. B. Trabalho de projetos no processo de ensinar e aprender estatística na universidade. **Bolema**, Rio Claro, v. 24, n. 39, p. 413-426, ago. 2011.

SPEARMAN, C. **Las habilidades del hombre**: su naturaleza y medición. Buenos Aires: Paidós, 1927.

SPIEGEL, M. R.; SCHILLER, J. J.; SRINIVASAN, R. A. **Teoria e problemas de probabilidade e estatística**. 2. ed. Porto Alegre: Bookman, 2004.

STEVENSON, W. J. **Estatística aplicada à administração**. 1. ed. São Paulo: Harbra, 2001.

STUART, M. Changing the teaching of statistics. **The Statistician**, v. 44, n. 1, p. 45-54, 1995.

TUFFÉRY, S. **Data mining et statistique décisionnelle**: l'intelligence des données. Paris: Editions Technip, 2010.

VERE-JONES, D. The coming of age statistical education. **International Statistical Review**, v. 63, n. 1, ago. 1995.

VERGNAUD. G. A comprehensive theory of representation for mathematics education. **Journal of Mathematical Behavior**, v. 17, n. 2, p. 167-181, 1998.

_____. A teoria dos campos conceituais. *In*: BRUN, J. **Didáctica das matemáticas**. Lisboa: Instituto Piaget, 1996. p. 155-191.

_____. Epistemology and psychology of mathematics education. *In*: NECHER, P.; KILPATRICK, J. (Eds.). **Mathematics and cognition**. Cambridge: Cambridge University Press, 1990a. p. 14-30.

_____. Forme opératoire et forme prédicative de la connaissance (operative and predicative forms of knowledge). **Investigações em Ensino de Ciências**, v. 17, n. 2, p. 287-304, 2001.

_____. La théorie des champs conceptuels. **Récherches em Didactique des Mathématiques**, v. 10, n. 23, p. 133-170, 1990b.

VERGNAUD, G. Teoria dos campos conceituais. *In*: SEMINÁRIO INTERNACIONAL DE EDUCAÇÃO MATEMÁTICA DO RIO DE JANEIRO, 1., Rio de Janeiro, 1993. **Anais [...]**. Rio de Janeiro: UFRJ Projeto Fundação, Instituto de Matemática, 1993. p. 1-26.

VERGNAUD, G.; PLAISANCE, É. **Les sciences de l'éducation**. 5. ed. Paris: Éditions la Découverte, 2012.

VIALI, L. Algumas considerações sobre a origem da teoria da probabilidade. **Revista Brasileira de História da Matemática**, v. 8, n. 16, p. 143-153, 2008.

VIANA, G. S. **Atitude e motivação em relação ao desempenho acadêmicos de alunos do curso de graduação em administração em disciplinas de estatística**. 2012. 199f. Dissertação (Mestrado em Administração de Organizações) - Universidade de São Paulo, Ribeirão Preto, 2012.

VON MISES, R. **Probability, statistics and truth**. 2. ed. New York: Dover, 1981.

WADA, R. S. **Estatística e ensino**: um estudo sobre representações de professores de 3º grau. 1996. 119f. Tese (Doutorado em Educação) - Universidade Estadual de Campinas, Campinas, 1996.

WALLMANN, K. K. Enhancing statistical literacy: enriching our society. **Journal of the American Statistical Association**, v. 88, p. 1-8, 1993.

WATSON, J. Assessing statistical thinking using the media. *In*: GAL, I.; GARFIELD, J. (Eds.). **The assessment challenge in statistics education**. Amsterdam: IOS Press and International Statistical Institute, 1997. p. 107-121.

WATSON, J. *et al.* The measurement of school students' understanding of statistical variation. **Journal of Mathematical Education in Science and Technology**, v. 34, p. 1-29, 2003.

WILD, C. J.; PFANNKUCH, M. Statistical thinking in empirical enquiry. **International Statistical Review**, v. 67, n. 3, p. 223-265, 1999.

WODEWOTZKI, M. L. *et al.* Temas contemporâneos nas aulas de estatística: um caminho para combinar aprendizagem e reflexões políticas. *In*: LOPES, C. E.; COUTINHO, C. Q. S.; ALMOULOUD, S. A. (Orgs.). **Estudos e reflexões em educação estatística**. 1. ed. Campinas: Mercado de Letras, 2010. p. 65-83.

WODEWOTZKI, M. L.; JACOBINI, O. R. O ensino de estatística no contexto da educação matemática. *In*: BICUDO, M. A. V.; BORBA, M. C. (Orgs.). **Educação matemática**: pesquisa em movimento. São Paulo: Cortez, 2004. p. 232-249.

YULE, G. U.; KENDALL, M. G. **Introdução à teoria da estatística**. Rio de Janeiro: Conselho Nacional de Estatística, 1948.

ANEXO A - QUESTIONÁRIO Q1



Programa de Pós-Graduação em Ensino
de Ciências e Matemática

UNIVERSITÉ
LUMIÈRE
LYON 2
UNIVERSITÉ DE LYON

1/3



Ano letivo 2016 – Centro Universitário Franciscano

Estudo das dificuldades e facilidades de aprendizagem em estatística

Questionário apresentado aos estudantes que estão cursando a disciplina de Estatística

Data :/...../2016

Você pode optar por permanecer anônimo.

Nome :	Sobrenome :	Matrícula :
--------	-------------	-------------

1 Informações da sua formação

[V1] Masculino (assinale a resposta)	Feminino	[V2] Data de nascimento :	Código do anonimato :
---	----------	---------------------------	-----------------------

[V3] Você já enfrentou situações que exigiram conhecimentos em estatística? (Marque sua resposta)

1_Nunca	2_Raramente	3_Algunas vezes	4_Frequentemente	5_Muito Frequentemente	6_Quase todos os dias
---------	-------------	--------------------	------------------	---------------------------	-----------------------------

[V4] Quando os conhecimentos em Estatística são requisitados, isso lhe causa problemas? (Marque sua resposta)

1_De modo algum	2_Raramente	3_Algunas vezes	4_Frequentemente	5_Muito Frequentemente	6_Quase todos os dias
--------------------	-------------	--------------------	------------------	---------------------------	-----------------------------

[V5] Semestre :

2 Você e a Estatística

[V6] Hoje, o que lhe evoca a Estatística?

(Nomear três palavras)

--	--	--

(Explicitar em uma frase)

--

[V7] Qual o nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação?

0	1	2	3	4	5	6	7	8	9	10
Muito fraco										Muito forte

[V8] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação?

0	1	2	3	4	5	6	7	8	9	10
Muito fraco										Muito forte

[V9] Atualmente o que significa para você os termos a seguir ?

Representativo	
Significativo	
Estimação	
Testes de Hipóteses	
Probabilidade	
Sondagem	

[V10] Atualmente o que significa para você uma sondagem de opinião?

--

[V11] Nos dias de hoje, qual utilidade você atribui para a Estatística?

--

[V12] Para você, quais são as **três maiores dificuldades** encontradas numa disciplina de Estatística?

Dif_1 :	Dif_2 :	Dif_3 :
---------	---------	---------

[V13] Para você, quais são as **três maiores contribuições** encontradas numa disciplina de Estatística?

Cont_1 :	Cont_2 :	Cont_3 :
----------	----------	----------

[V14] Você já obteve conhecimentos em Estatística ? *(Marque sua resposta)*

1_Nunca	2_ Em um nível muito superficial	3_ Em um nível superficial	4_Introdução ao básico	5_Introdução ao avançado	6_Eu tenho uma formação em Estatística ou quase
---------	----------------------------------	----------------------------	------------------------	--------------------------	---

2.1 [V15] Atualmente qual o seu sentimento em relação a Estatística?

Instruções : Cada uma das frases seguintes exprime um sentimento em relação a Estatística. Você deverá comparar o seu sentimento pessoal com o que é expresso nas frases marcando com um X conforme o seu grau de adesão.

1	2	3	4
Absolutamente contra.	Não estou bem de acordo.	Estou de acordo.	Concordo plenamente.

1. Eu não tenho um bom conhecimento em Estatística.	1	2	3	4
2. A Estatística é algo que aprecio muito.	1	2	3	4
3. Eu me sinto inseguro quando preciso utilizar a Estatística.	1	2	3	4
4. Eu tenho uma relação muito boa com a Estatística: eu gosto e aprecio esta disciplina.	1	2	3	4
5. Quando estudo Estatística, não consigo pensar claramente pois minha cabeça fica bloqueada.	1	2	3	4
6. Quando eu escuto a palavra Estatística eu sinto um sentimento de aversão.	1	2	3	4
7. Eu realmente não gosto de matemática.	1	2	3	4
8. Pensar em resolver um problema de Estatística me deixa nervoso.	1	2	3	4
9. A Estatística é fascinante e divertida.	1	2	3	4
10. Eu nunca gostei de Estatística, ela é a disciplina que me causa mais medo.	1	2	3	4
11. O sentimento que experimento em relação a Estatística é bom.	1	2	3	4
12. Eu tenho medo de Estatística pois tenho medo de Matemática.	1	2	3	4
13. Eu acho a Estatística muito interessante e gosto de ir na aula dessa disciplina.	1	2	3	4
14. A Estatística me dá segurança e me estimula.	1	2	3	4
15. A Estatística me dá a impressão que estou perdido em uma sala cheia de números na qual eu não acho saída.	1	2	3	4
16. Eu me sinto bem com a Estatística e gosto muito dessa disciplina.	1	2	3	4
17. Eu estou sempre estressado durante as aulas de Estatística.	1	2	3	4
18. A Estatística é uma das matérias que eu realmente gosto de estudar na universidade ou na escola.	1	2	3	4
19. Eu não gosto de Estatística e nem de ir nas aulas de Estatística.	1	2	3	4
20. A Estatística me provoca um sentimento de indecisão que resulta no meu medo de não ser capaz de ter êxito na disciplina.	1	2	3	4
21. Eu me sinto mais feliz com a Estatística do que com as outras disciplinas.	1	2	3	4
22. A Estatística me deixa inquieto, descontente, irritado e impaciente.	1	2	3	4
23. Eu sou competente em Matemática.	1	2	3	4
24. Eu realmente gosto de Estatística.	1	2	3	4

2.2 [V16] Atualmente, qual a sua definição de Estatística?

ANEXO B - QUESTIONÁRIO Q2



Programa de Pós-Graduação em Ensino
de Ciências e Matemática

UNIVERSITÉ
LUMIÈRE
LYON 2
UNIVERSITÉ DE LYON

1/2



Questionário 2

Identificação : (você pode optar em permanecer anônimo)

Nome :	Sobrenome :	Matrícula :

Situação 1

S1_1) Vamos supor que você esteja interessado em investigar a quantidade de ocupantes dentro dos automóveis de passeio que transitam por dia em uma avenida movimentada de sua cidade. Como você procederia para dar início a essa investigação ?

S1_2) Frente as informações do item anterior, que tipo de medidas estatísticas você utilizaria para estabelecer algum tipo de padrão ? Faça uma simulação.

Situação 2

S2_1) Em um supermercado o funcionário responsável pelo hortifrutigranjeiro recebeu a incumbência de empacotar as maçãs em pacotes de dois quilos. O que o funcionário deverá levar em consideração ?

S2_2) Após empacotar todas as maçãs num total de 150 pacotes, o funcionário ficou curioso em saber sobre a distribuição das frutas nos pacotes. Não possuindo tempo para fazer uma contagem sobre todos os pacotes, que procedimento ele poderia tomar para fazer essa descoberta ?

S2_3) Que tipos de medidas estatísticas podemos proceder para tentarmos estabelecer algum padrão em relação aos pacotes de maçãs ? Faça uma simulação.

S3_1) Em ambas as situações anteriores (Situação 1 e Situação 2), você acredita que possam haver erros em relação aos padrões estabelecidos ? Justifique sua resposta.

OBRIGADO PELA SUA CONTRIBUIÇÃO !

ANEXO C - CURVA NORMAL REDUZIDA

A Curva Normal Reduzida

Curvas normais, com qualquer μ e σ , podem ser transformadas em uma curva normal que tem média igual a 0 ($\mu = 0$) e desvio padrão igual a 1 ($\sigma = 1$). Esta curva normal, com média 0 e desvio padrão 1, é conhecida como curva normal reduzida. Suas probabilidades são apresentadas em tabelas de fácil utilização.

Como a normal é simétrica, a tabela apresenta somente as probabilidades da metade direita da curva. A probabilidade de um intervalo qualquer da metade esquerda é igual à probabilidade do intervalo equivalente na metade direita.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

ANEXO D - ÍNDICE DE POSITIVIDADE

Estudantes	Soma ATT(-)	Soma ATT(+)	Nível de positividade antes	Positividade antes	Nível de positividade antes padronizado	Soma ATT(-)	Soma ATT(+)	Nível de positividade depois	Positividade depois	Nível de positividade depois padronizado	Diferença Depois Antes	
E01_001	-0.269230769	0.272727273	0.513888889	a_Positividade 2	-0.121867514	-0.173076923	0.477272727	0.708333333	d_Positividade 4	0.84645422	0.988321734	Aumento
E01_002	-0.307692308	0.386363636	0.541666667	a_Positividade 2	-0.059256681	-0.192307692	0.477272727	0.694444444	d_Positividade 3	0.529038888	0.469777307	Aumento
E01_003	-0.403846154	0.272727273	0.416666667	a_Positividade 1	-0.755802195	-0.211538462	0.409090909	0.638888889	d_Positividade 1	-0.740647443	0.015154753	Aumento
E01_004	-0.326923077	0.386363636	0.541666667	a_Positividade 2	0.059256681	0.25	0.431818182	0.625	d_Positividade 1	-1.058067775	-1.117324456	Atenuação
E01_005	-0.307692308	0.386363636	0.541666667	a_Positividade 2	0.059256681	-0.173076923	0.409090909	0.666666667	d_Positividade 2	-0.150067778	-0.150067778	Atenuação
E01_006	-0.288461538	0.386363636	0.569444444	a_Positividade 3	0.240380876	-0.192307692	0.5	0.708333333	d_Positividade 4	0.84645422	0.606073344	Aumento
E01_007	-0.307692308	0.386363636	0.555555556	a_Positividade 3	0.149818778	-0.192307692	0.386363636	0.638888889	d_Positividade 1	-0.740647443	-0.390466221	Atenuação
E01_008	-0.307692308	0.295454545	0.5	a_Positividade 2	0.105429752	-0.211538462	0.386363636	0.625	d_Positividade 1	-1.058067775	-0.845316164	Atenuação
E01_009	-0.423076923	0.295454545	0.416666667	a_Positividade 1	-0.755802195	-0.288461538	0.454545455	0.611111111	d_Positividade 1	-1.375488108	-0.619685912	Atenuação
E01_010	-0.403846154	0.295454545	0.430555556	a_Positividade 1	-0.665240098	-0.173076923	0.454545455	0.694444444	d_Positividade 3	0.529038888	1.194273986	Aumento
E01_011	-0.307692308	0.386363636	0.541666667	a_Positividade 2	0.059256681	-0.173076923	0.431818182	0.680555556	d_Positividade 3	0.211613555	-0.255262556	Aumento
E01_012	-0.25	0.545454545	0.944444444	a_Positividade 4	0.059256681	-0.192307692	0.5	0.708333333	d_Positividade 4	0.84645422	-0.208085532	Atenuação
E01_013	-0.365384615	0.318181818	0.472222222	a_Positividade 2	-0.393553806	-0.192307692	0.431818182	0.666666667	d_Positividade 2	-0.105806778	0.287747028	Aumento
E01_014	-0.346153846	0.340909091	0.5	a_Positividade 2	-0.212429611	-0.230769231	0.431818182	0.638888889	d_Positividade 1	-0.740647443	-0.528217831	Atenuação
E01_015	-0.365384615	0.386363636	0.5	a_Positividade 2	-0.212429611	0.25	0.5	0.666666667	d_Positividade 2	-0.105806778	0.106622834	Aumento
E01_016	-0.230769231	0.472727273	0.666666667	a_Positividade 4	0.874315557	-0.173076923	0.477272727	0.708333333	d_Positividade 4	0.84645422	-0.027061187	Atenuação
E01_017	-0.326923077	0.295454545	0.486111111	a_Positividade 2	-0.302991709	0.25	0.431818182	0.625	d_Positividade 1	-1.058067775	-0.755076067	Atenuação
E01_018	-0.307692308	0.318181818	0.513888889	a_Positividade 2	-0.121867514	-0.173076923	0.409090909	0.666666667	d_Positividade 2	-0.105806778	0.061794777	Aumento
E01_019	-0.307692308	0.386363636	0.555555556	a_Positividade 3	0.149818778	-0.211538462	0.477272727	0.680555556	d_Positividade 3	0.211613555	0.061794777	Aumento
E01_020	-0.307692308	0.386363636	0.555555556	a_Positividade 3	0.149818778	-0.211538462	0.477272727	0.666666667	d_Positividade 2	-0.105806778	-0.255262556	Atenuação
E01_021	-0.307692308	0.409090909	0.569444444	a_Positividade 3	0.240380876	-0.230769231	0.477272727	0.666666667	d_Positividade 2	-0.105806778	-0.346187653	Atenuação
E01_022	-0.269230769	0.477272727	0.638888889	a_Positividade 4	0.693191363	-0.211538462	0.522727273	0.708333333	d_Positividade 4	0.84645422	0.153262857	Aumento
E01_023	-0.326923077	0.431818182	0.569444444	a_Positividade 3	0.240380876	-0.173076923	0.454545455	0.694444444	d_Positividade 3	0.529038888	0.288653012	Aumento
E01_024	-0.326923077	0.318181818	0.5	a_Positividade 2	-0.212429611	-0.230769231	0.431818182	0.638888889	d_Positividade 1	-0.740647443	-0.528217831	Atenuação
E01_025	-0.365384615	0.272727273	0.444444444	a_Positividade 1	-0.574678001	-0.25	0.477272727	0.652777778	d_Positividade 2	-0.42322711	0.151450891	Aumento
E01_026	-0.442307692	0.318181818	0.416666667	a_Positividade 1	-0.755802195	-0.192307692	0.477272727	0.694444444	d_Positividade 3	0.529038888	1.284836083	Aumento
E01_027	-0.423076923	0.272727273	0.402777778	a_Positividade 1	-0.846364293	-0.230769231	0.5	0.680555556	d_Positividade 3	0.211613555	1.057977848	Aumento
E01_028	-0.423076923	0.25	0.388888889	a_Positividade 1	-0.93692639	0.25	0.409090909	0.611111111	d_Positividade 1	-1.375488108	-0.438561717	Atenuação
E01_029	-0.25	0.386363636	0.597222222	a_Positividade 2	0.421505071	-0.211538462	0.5	0.694444444	d_Positividade 3	0.529038888	0.107528817	Aumento
E01_030	-0.365384615	0.340909091	0.486111111	a_Positividade 2	-0.302991709	-0.192307692	0.431818182	0.666666667	d_Positividade 2	-0.105806778	0.197184891	Aumento
E01_031	-0.423076923	0.340909091	0.444444444	a_Positividade 1	-0.574678001	-0.192307692	0.454545455	0.680555556	d_Positividade 3	0.211613555	0.786291556	Aumento
E01_032	-0.288461538	0.318181818	0.527777778	a_Positividade 2	-0.031305416	-0.173076923	0.431818182	0.680555556	d_Positividade 3	0.211613555	0.242918971	Aumento
E01_001	-0.25	0.386363636	0.597222222	a_Positividade 3	0.421505071	-0.211538462	0.431818182	0.652777778	d_Positividade 2	-0.42322711	-0.584731281	Atenuação
E02_002	-0.346153846	0.340909091	0.5	a_Positividade 2	-0.212429611	-0.288461538	0.431818182	0.597222222	d_Positividade 1	-1.058067775	-1.409478209	Atenuação
E02_003	-0.288461538	0.431818182	0.597222222	a_Positividade 3	0.421505071	-0.192307692	0.5	0.708333333	d_Positividade 4	0.84645422	0.424949155	Aumento
E02_004	-0.192307692	0.431818182	0.666666667	a_Positividade 4	0.874315557	-0.230769231	0.522727273	0.694444444	d_Positividade 3	0.529038888	-0.345281277	Atenuação
E02_005	-0.019230769	0.613636364	0.902777778	a_Positividade 4	2.413871213	-0.173076923	0.431818182	0.680555556	d_Positividade 3	0.211613555	-2.202257658	Atenuação
E02_001	-0.25	0.5	0.666666667	a_Positividade 4	0.874315557	0.25	0.5	0.666666667	d_Positividade 2	-0.105806778	-0.980122335	Atenuação
E02_002	-0.269230769	0.318181818	0.541666667	a_Positividade 3	0.240380876	-0.230769231	0.386363636	0.597222222	d_Positividade 1	-1.058067775	-1.752165121	Atenuação
E02_003	-0.230769231	0.409090909	0.625	a_Positividade 4	0.602629265	-0.173076923	0.477272727	0.708333333	d_Positividade 4	0.84645422	0.243824955	Aumento
E02_004	-0.5	0.272727273	0.319444444	a_Positividade 1	-1.389738877	-0.211538462	0.5	0.694444444	d_Positividade 3	0.529038888	1.918770765	Aumento
E02_005	-0.096153846	0.5	0.777777778	a_Positividade 4	1.598812346	0.25	0.454545455	0.638888889	d_Positividade 1	-0.740647443	-2.339499779	Atenuação
E02_006	-0.346153846	0.386363636	0.388888889	a_Positividade 2	-0.518939354	-0.153846154	0.5	0.736111111	d_Positividade 1	-1.058067775	2.148212175	Aumento
E02_007	-0.038461538	0.522727273	0.833333333	a_Positividade 4	1.961067026	-0.230769231	0.477272727	0.666666667	d_Positividade 2	-0.105806778	-2.066676703	Atenuação
E02_008	-0.25	0.295454545	0.541666667	a_Positividade 2	0.059256681	-0.192307692	0.5	0.708333333	d_Positividade 4	0.84645422	0.787139739	Aumento
E02_009	-0.115384615	0.318181818	0.652777778	a_Positividade 2	0.78735346	-0.173076923	0.431818182	0.680555556	d_Positividade 3	0.211613555	-0.572139905	Atenuação
E02_010	-0.557692308	0.272727273	0.277777778	a_Positividade 1	-1.661423169	-0.25	0.545454545	0.694444444	d_Positividade 3	0.529038888	2.190457057	Aumento
E02_011	-0.192307692	0.25	0.555555556	a_Positividade 2	-0.149818778	-0.211538462	0.386363636	0.652777778	d_Positividade 1	-1.058067775	-1.752165121	Atenuação
E02_012	-0.115384615	0.522727273	0.777777778	a_Positividade 2	1.598812346	-0.211538462	0.431818182	0.652777778	d_Positividade 2	-0.42322711	-2.022019447	Atenuação
E02_013	-0.269230769	0.318181818	0.541666667	a_Positividade 2	-0.269230769	0.386363636	0.569444444	0.611111111	d_Positividade 1	-1.375488108	-2.387005786	Atenuação
E02_014	-0.288461538	0.272727273	0.472222222	a_Positividade 2	-0.393553806	0.25	0.409090909	0.611111111	d_Positividade 1	-1.375488108	-0.981934002	Atenuação
E02_015	-0.173076923	0.409090909	0.625	a_Positividade 4	0.602629265	-0.230769231	0.477272727	0.666666667	d_Positividade 2	-0.105806778	-0.708430463	Atenuação
E02_016	-0.096153846	0.522727273	0.791666667	a_Positividade 4	1.689374434	0.25	0.409090909	0.611111111	d_Positividade 1	-1.375488108	-3.064862541	Atenuação
E02_017	-0.288461538	0.386363636	0.416666667	a_Positividade 1	-0.755802195	-0.192307692	0.431818182	0.666666667	d_Positividade 2	-0.105806778	0.649995418	Aumento
E02_018	-0.134615385	0.659090909	0.847222222	a_Positividade 4	2.051622823	-0.230769231	0.477272727	0.666666667	d_Positividade 2	-0.105806778	-2.157429601	Atenuação
E02_019	-0.192307692	0.522727273	0.722222222	a_Positividade 4	1.236562947	-0.326923077	0.454545455	0.583333333	d_Positividade 1	-2.010328773	-3.246892772	Atenuação
E02_020	-0.38615385	0.295454545	0.444444444	a_Positividade 1	-0.574678001	0.25	0.5	0.666666667	d_Positividade 2	-0.105806778	0.468871223	Aumento
E02_021	-0.596153846	0.022727273	0.125	a_Positividade 1	-2.65760624	-0.134615385	0.545454545	0.777777778	d_Positividade 4	2.43555883	5.091162123	Aumento
E02_022	-0.538461538	0.090909091	0.208333333	a_Positividade 1	-2.11423656	-0.134615385	0.590909091	0.805555556	d_Positividade 4	3.068396548	5.182630204	Aumento
E02_023	-0.557692308	0.136363636	0.222222222	a_Positividade 1	-2.023671559	-0.25	0.522727273	0.680555556	d_Positividade 3	0.211613555	2.235285114	Aumento
E02_024	-0.596153846	0.081818182	0.152777778	a_Positividade 1	-2.47482046	-0.134615385	0.590909091	0.805555556	d_Positividade 4	3.068396548	5.544878993	Aumento
E02_025	-0.442307692	0.386363636	0.444444444	a_Positividade 1	-0.574678001	-0.288461538	0.545454545	0.666666667	d_Positividade 2	-0.105806778	0.468871223	Aumento
E02_026	-0.326923077	0.25	0.458333333	a_Positividade 2	-0.484115903	0.25	0.454545455	0.638888889	d_Positividade 1	-0.740647443	-0.256515339	Atenuação
E02_027	-0.269230769	0.295454545	0.527777778	a_Positividade 2	-0.031305416	-0.211538462	0.477272727	0.680555556	d_Positividade 3	0.211613555	0.242918971	Aumento
E02_028	-0.365384615	0.272727273	0.444444444	a_Positividade 1	-0.574678001	-0.211538462	0.5	0.694444444	d_Positividade 3	0.529038888	1.103711888	Aumento
E02_029	-0.480769231	0.318181818	0.305555556	a_Positividade 1	-1.40298974	0.25	0.477272727	0.652777778	d_Positividade 2	-0.42322711	1.05701864	Aumento
E04_001	-0.192307692	0.204545455	0.527777778	a_Positividade 2	-0.031305416	-0.211538462	0.386363636	0.625	d_Positividade 1	-1.05806777		

ANEXO E - V03 X POSITIVIDADE

a_POSITIVIDADE / [aV03] Você já enfrentou situações que exigiram conhecimentos em estatística ?

Nombre de cases d'effectif théorique < 5	19
Nombre de degrés de liberté	15
Khi-2	19,841
P-value	0,178
T de Tschuprow	0,251
V de Cramer	0,286

Effectif/poids		1_Nunca	2_Raramente	3_Algunas vezes	4_Frequentemente	5_Muito Frequentemente	6_Quase todos os dias	Ensemble
% en ligne	% en colonne							
a_Positividade 1	0,0	0	4	11	5	2	0	22
	0,0	18,2	50,0	22,7	9,1	0,0	100,0	27,2
a_Positividade 2	2	3	8	4	4	5	26	
	7,7	11,5	30,8	15,4	15,4	19,2	100,0	32,1
a_Positividade 3	0	0	6	3	5	0	14	
	0,0	0,0	42,9	21,4	35,7	0,0	100,0	17,3
a_Positividade 4	0	2	7	5	4	1	19	
	0,0	10,5	36,8	26,3	21,1	5,3	100,0	23,5
Ensemble	2,5	9	32	17	15	6	81	
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

d_POSITIVIDADE / [dV03] Você já enfrentou situações que exigiram conhecimentos em estatística ?

Nombre de cases d'effectif théorique < 5	12
Nombre de degrés de liberté	12
Khi-2	9,333
P-value	0,674
T de Tschuprow	0,182
V de Cramer	0,196

Effectif/poids		2_Raramente	3_Algunas vezes	4_Frequentemente	5_Muito Frequentemente	6_Quase todos os dias	Ensemble
% en ligne	% en colonne						
d_Positividade 1	1	6	9	2	4	22	
	4,5	27,3	40,9	9,1	18,2	100,0	27,2
d_Positividade 2	2	4	8	3	3	20	
	10,0	20,0	40,0	15,0	15,0	100,0	24,7
d_Positividade 3	0	7	7	4	3	21	
	0,0	33,3	33,3	19,0	14,3	100,0	25,9
d_Positividade 4	0	7	7	4	0	18	
	0,0	38,9	38,9	22,2	0,0	100,0	22,2
Ensemble	3	24	31	13	10	81	
	3,7	29,6	38,3	16,0	12,3	100,0	100,0

ANEXO G - SEMESTRE X NÍVEL DE POSITIVIDADE

[aV05] Semestre / a_POSITIVIDADE

Nombre de cases d'effectif théorique < 5	19
Nombre de degrés de liberté	15
Khi-2	11,299
P-value	0,731
T de Tschuprow	0,190
V de Cramer	0,216

Effectif/poids	a_Positividade 1	a_Positividade 2	a_Positividade 3
% en ligne			
% en colonne			
Semestre2	0	0	0
	0,0	0,0	0,0
	0,0	0,0	0,0
Semestre3	8	11	5
	28,6	39,3	17,9
	36,4	42,3	35,7
Semestre4	3	8	3
	14,3	38,1	14,3
	13,6	30,8	21,4
Semestre5	2	2	2
	25,0	25,0	25,0
	9,1	7,7	14,3
Semestre6	5	1	2
	50,0	10,0	20,0
	22,7	3,8	14,3
Semestre7	4	4	2
	30,8	30,8	15,4
	18,2	15,4	14,3
Ensemble	22	26	14
	27,2	32,1	17,3
	100,0	100,0	100,0

[aV05] Semestre / d_POSITIVIDADE

Nombre de cases d'effectif théorique < 5	17
Nombre de degrés de liberté	15
Khi-2	10,327
P-value	0,799
T de Tschuprow	0,181
V de Cramer	0,206

Effectif/poids	d_Positividade 1	d_Positividade 2	d_Positividade 3
% en ligne			
% en colonne			
	1	0	0

Semestre2	100,0	0,0	0,0
	4,5	0,0	0,0
Semestre3	7	8	9
	25,0	28,6	32,1
	31,8	40,0	42,9
Semestre4	6	7	4
	28,6	33,3	19,0
	27,3	35,0	19,0
Semestre5	1	1	2
	12,5	12,5	25,0
	4,5	5,0	9,5
Semestre6	3	2	3
	30,0	20,0	30,0
	13,6	10,0	14,3
Semestre7	4	2	3
	30,8	15,4	23,1
	18,2	10,0	14,3
Ensemble	22	20	21
	27,2	24,7	25,9
	100,0	100,0	100,0

a_Positividade 4	Ensemble
1	1
100,0	100,0
5,3	1,2
4	28
14,3	100,0
21,1	34,6
7	21
33,3	100,0
36,8	25,9
2	8
25,0	100,0
10,5	9,9
2	10
20,0	100,0
10,5	12,3
3	13
23,1	100,0
15,8	16,0
19	81
23,5	100,0
100,0	100,0

d_Positividade 4	Ensemble
0	1

0,0	100,0
0,0	1,2
4	28
14,3	100,0
22,2	34,6
4	21
19,0	100,0
22,2	25,9
4	8
50,0	100,0
22,2	9,9
2	10
20,0	100,0
11,1	12,3
4	13
30,8	100,0
22,2	16,0
18	81
22,2	100,0
100,0	100,0

ANEXO H - FRASES X POSITIVIDADE

Phrases longues caractéristiques pour "a_POSITIVIDADE"

a_Positividade 1

Critère de classement	Index	Libellé de la réponse	Identifiant
0,830	1	É importante para conhecer gráficos e tabelas.	Indiv. 10
0,848	2	Uma coleta de dados para discutir seus resultados.	Indiv. 66
0,852	3	É uma disciplina necessária para a formação acadêmica.	Indiv. 68
0,853	4	Não sei matemática e não sei ler gráficos.	Indiv. 31
0,854	5	Como são criados os gráficos e tabelas.	Indiv. 74
0,858	6	Ciência que estuda gráficos e tabelas.	Indiv. 28
0,878	7	Vejo um grande problema para resolver.	Indiv. 41
0,888	8	Não entendo por que tem matemática.	Indiv. 9
0,894	9	Interpretações dos gráficos e matemática.	Indiv. 60
0,899	10	Pesquisa para o trabalho final.	Indiv. 25
0,906	11	Interessante porém tenho muita dificuldade.	Indiv. 43
0,908	12	Fazer média e gráficos.	Indiv. 26
0,912	13	Não consigo entender matemática.	Indiv. 62
0,939	14	Muitas interpretações com números.	Indiv. 3
0,952	15	Matemática com muitos cálculos e fórmulas.	Indiv. 27
0,979	16	Cálculo.	Indiv. 65
0,995	17	Analisar dados.	Indiv. 47

a_Positividade 2

Critère de classement	Index	Libellé de la réponse	Identifiant
0,758	1	Não conheço muito bem para falar, mas tirando os números eu gosto da disciplina.	Indiv. 24
0,804	2	Eu sei que é importante mas não conheço muito ainda.	Indiv. 18
0,818	3	Ajuda a fazer e interpretar os gráficos que precisamos durante a vida.	Indiv. 34
0,819	4	Cálculos fazem parte da nossa vida e dos estudos.	Indiv. 67
0,892	5	É uma matemática escondida, mas muito útil.	Indiv. 8
0,896	6	Por não ter conhecimento causa insegurança.	Indiv. 39
0,896	7	Importante para a vida profissional.	Indiv. 14
0,901	8	Eleições sempre tem estatística dos candidatos.	Indiv. 32
0,906	9	Matemática com exercícios mais práticos	Indiv. 17
0,908	10	Informações com muitos números e matemática.	Indiv. 1
0,911	11	Interpretações dos gráficos e matemática.	Indiv. 4
0,915	12	São cálculos com porcentagens e gráficos.	Indiv. 30
0,932	13	Avaliar as melhores decisões.	Indiv. 50
0,935	14	Matemática com aplicações.	Indiv. 11
0,936	15	Não sei nunca vi.	Indiv. 15
0,947	16	Estatística é probabilidade.	Indiv. 63
0,948	17	Ferramenta para conhecer acontecimentos.	Indiv. 75
0,953	18	Analisar os dados estatísticos.	Indiv. 73
0,956	19	Contas de matemática.	Indiv. 5
0,978	20	Muitas fórmulas.	Indiv. 2
0,991	21	Muito cálculo.	Indiv. 13

a_Positividade 3

Critère de classement	Index	Libellé de la réponse	Identifiant
0,809	1	Analisar muitos dados e fazer muita conta.	Indiv. 7
0,829	2	Informação e representação através de dados coletados e tratados.	Indiv. 80
0,832	3	Ela é importante para fazer pesquisas e entender problemas.	Indiv. 33
0,836	4	Não gosto por que tem matemática e muita conta.	Indiv. 6
0,838	5	Muito cálculo mas fundamental para a área que escolhi.	Indiv. 19
0,851	6	Usar o Excel para fazer cálculos e gráficos.	Indiv. 48
0,874	7	Envolve pesquisa e opiniões com cálculo.	Indiv. 35
0,887	8	Quando envolve números fica muito complicado.	Indiv. 20
0,892	9	Recolhe e analisa os dados.	Indiv. 72
0,897	10	Ocupa em todo curso quase.	Indiv. 23
0,900	11	Ajuda na elaboração de gráficos e tabelas.	Indiv. 76
0,901	12	Concentração para efetuar os cálculos.	Indiv. 21
0,901	13	Coletar muitas informações e tabelar.	Indiv. 29

a_Positividade 4

Critère de classement	Index	Libellé de la réponse	Identifiant
0,776	1	É uma forma de obter possíveis conclusões através de dados coletados na realidade.	Indiv. 42
0,786	2	A estatística serve para a gente aprender a ler gráficos e fazer previsões com cálculos.	Indiv. 22
0,806	3	Na maioria das vezes os dados que encontramos em tabelas e gráficos são números.	Indiv. 71
0,815	4	Acho muito interessante essas pesquisas através do recolhimento de dados.	Indiv. 38
0,860	5	Análise de dados.	Indiv. 81
0,867	6	Análise e interpretação de dados.	Indiv. 52
0,870	7	Projeção de cálculos visando um resultado satisfatório.	Indiv. 36
0,870	8	A probabilidade de eu me formar.	Indiv. 53
0,879	9	Muito útil para compreender muitas informações que lemos.	Indiv. 12
0,898	10	Organização e análise das informações.	Indiv. 44
0,908	11	A estatística engloba pesquisa.	Indiv. 40
0,924	12	Muita dificuldade mas muito útil.	Indiv. 16
0,929	13	Conjunto de números.	Indiv. 55
0,933	14	Faço cálculos no Excel.	Indiv. 49
0,935	15	Estimativa de acontecimentos.	Indiv. 46
0,949	16	Para fazer trabalhos.	Indiv. 56
1,028	17	Não sei.	Indiv. 37

ANEXO I - FRASES X POSITIVIDADE DEPOIS

Phrases longues caractéristiques pour "d_POSITIVIDADE"

d_Positividade 1

Critère de classement	Index	Libellé de la réponse	Identifiant
0,748	1	Usada para a realização de pesquisas por meio de uma amostra.	Indiv. 70
0,830	2	Tomar decisões por meio de dados coletados, amostra.	Indiv. 50
0,856	3	Ciência que analisa os dados de uma amostra para uma pesquisa.	Indiv. 28
0,876	4	Tira conclusões com base em uma amostra.	Indiv. 24
0,882	5	Responsável no tratamentos de dados de uma pesquisa.	Indiv. 34
0,886	6	Coletar dados para realizar uma pesquisa.	Indiv. 48
0,891	7	Ciência que estuda uma amostra.	Indiv. 75
0,895	8	Coletar dados para verificar alguma pesquisa.	Indiv. 4
0,901	9	É uma disciplina que aprende muito.	Indiv. 8
0,919	10	Cuida dos dados de uma pesquisa.	Indiv. 14
0,923	11	Analisa uma amostra coletada.	Indiv. 67
0,928	12	Usada em pesquisas.	Indiv. 56
0,928	13	Pesquisa sobre determinado assunto.	Indiv. 7
0,935	14	Estuda e interpreta dados numéricos.	Indiv. 51
0,939	15	Analisa uma pesquisa de opinião.	Indiv. 3
0,944	16	Estuda fenômenos coletivos.	Indiv. 63
0,952	17	Construção de tabelas e gráficos.	Indiv. 39
0,954	18	Pesquisa numérica.	Indiv. 17
0,961	19	Pode fazer projeções.	Indiv. 77
0,983	20	Pesquisas.	Indiv. 9
1,021	21	Coleta e interpreta dados para uma pesquisa.	Indiv. 42

d_Positividade 2

Critère de classement	Index	Libellé de la réponse	Identifiant
0,822	1	Coleta e interpreta dados para um trabalho de pesquisa.	Indiv. 66
0,877	2	É uma ciência capaz de coletar e interpretar dados.	Indiv. 54
0,880	3	Informações dos dados de uma pesquisa.	Indiv. 5
0,888	4	Estuda como coleta e trata dados de uma pesquisa.	Indiv. 21
0,890	5	Coleta e interpreta dados que servem para realizar pesquisas.	Indiv. 52
0,894	6	Interpreta os números de uma pesquisa.	Indiv. 15
0,907	7	Faz um tratamentos da amostra.	Indiv. 30
0,911	8	Organização e análise de dados.	Indiv. 44
0,911	9	Intepretações dos cálculos de matemática.	Indiv. 13
0,916	10	São medidas para analisar dados.	Indiv. 62
0,922	11	Desenvolvimento de pesquisas em diversas áreas.	Indiv. 38
0,922	12	Cálculos que interpretam uma pesquisa.	Indiv. 25
0,922	13	Tratar dados de uma amostra.	Indiv. 33
0,930	14	Desvio padrão, média, mediana.	Indiv. 55
0,934	15	Analisa os dados estatísticos.	Indiv. 73
0,943	16	Trabalhos de pesquisas.	Indiv. 57
0,952	17	Interpretações de resultados.	Indiv. 78
0,963	18	Muito útil para pesquisa.	Indiv. 49
0,969	19	Interpretar dados numéricos.	Indiv. 20
0,979	20	Analisa os dados coletados de uma pesquisa.	Indiv. 18

d_Positividade 3

Critère de classement	Index	Libellé de la réponse	Identifiant
0,832	1	Matemática que é aplicada para cuidar de uma pesquisa.	Indiv. 32
0,870	2	Aplica a matemática para desenvolver uma pesquisa.	Indiv. 65
0,872	3	Ciência que domina a área de pesquisas.	Indiv. 68
0,873	4	Informação e representação através de dados coletados e tratados.	Indiv. 80
0,876	5	Usa a matemática para fazer pesquisa.	Indiv. 31
0,896	6	Estudo de como proceder para uma pesquisa.	Indiv. 11
0,898	7	Coleta e interpretações de dados.	Indiv. 37
0,901	8	Matemática com interpretações.	Indiv. 2
0,905	9	Interpretar dados para uma pesquisa de opinião.	Indiv. 27
0,907	10	Como são criados os gráficos e tabelas.	Indiv. 74
0,913	11	Dar opiniões com fundamento matemático.	Indiv. 10
0,926	12	Opiniões para pesquisas em diversas áreas.	Indiv. 64
0,929	13	Importante para realizar uma pesquisa.	Indiv. 23
0,933	14	Base para uma pesquisa.	Indiv. 19
0,934	15	Estimar alguma ação futura.	Indiv. 26
0,943	16	Ciência que estuda pesquisas.	Indiv. 41
0,950	17	Pode fazer previsões.	Indiv. 46
0,959	18	Interpretar dados..	Indiv. 36
1,019	19	Coleta e interpreta dados.	Indiv. 60
1,026	20	Analisa uma amostra.	Indiv. 29
1,037	21	Análise de dados de uma amostra.	Indiv. 47

d_Positividade 4

Critère de classement	Index	Libellé de la réponse	Identifiant
0,766	1	Serve para coletar e interpretar dados.	Indiv. 43
0,808	2	Análise de dados coletados para uma pesquisa.	Indiv. 1
0,809	3	Analisa os dados para fazer uma pesquisa usando a matemática.	Indiv. 35
0,825	4	Examina os dados de uma pesquisa e apresenta uma conclusão.	Indiv. 22
0,828	5	Muito útil para coletar e interpretar dados.	Indiv. 58
0,848	6	Interpretações de dados para uma pesquisa sobre algum assunto.	Indiv. 59
0,851	7	Avaliar os dados coletados para uma pesquisa.	Indiv. 61
0,858	8	Trata de dados e como tirar informações deles.	Indiv. 12
0,862	9	Interpretação de dados coletados para uma pesquisa.	Indiv. 45
0,877	10	Ajuda na elaboração de gráficos e tabelas.	Indiv. 76
0,898	11	Construir gráficos e interpretar.	Indiv. 71
0,903	12	Ciência que coleta dados e transforma em opiniões.	Indiv. 16
0,907	13	Interpretações usando matemática.	Indiv. 6
0,909	14	Análise de dados.	Indiv. 81
0,922	15	Descreve e analisa dados.	Indiv. 79
0,923	16	Serve para diversas pesquisas.	Indiv. 40
0,941	17	Analisa dados.	Indiv. 72
0,990	18	Coleta dados.	Indiv. 69

ANEXO J - aV07 X dV07

Statistiques sur variables continues

Tableau

Variable	Effectif	Moyenne	Écart-type (N)
[aV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	81	7,210	1,817
[dV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	81	8,247	1,233

T test pour deux échantillons

T test pour deux échantillons -[aV07], [dV07]

Paramètres de fonctionnement

Variables	[aV07] , [dV07]
Données appariées	Non
Seuil de confiance	0,950
Type de test	Bilatéral
Différence de moyennes supposée	0,000
Variances supposées égales	Oui

Statistiques descriptives

	[aV07]	[dV07]
Moyenne observée	7,210	8,247
Variance observée	3,343	1,538
Nombre d'individus	81	81
Minimum	2,000	6,000
Maximum	10,000	10,000

Hypothèse testée

H0	moyenne([aV07]) - moyenne([dV07]) = 0
H1	moyenne([aV07]) - moyenne([dV07]) <> 0

Résultats

Valeur de la statistique F	2,173
p-valeur associée à la statistique F	0,001
T test	-4,224
p-valeur du test	0,000
Degrés de liberté	160
Intervalle de confiance de la différence des moyennes]-1,522; -0,552[
Intervalle de confiance de la variance commune]1,983; 3,078[
Intervalle de confiance du ratio des variances]1,398; 3,379[

Conclusion du test

On rejette l'hypothèse nulle H0

ANEXO K - aV08 X dV08

Statistiques sur variables continues

Tableau

Variable	Effectif	Moyenne	Écart-type (N)
[aV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?	81	7,852	1,957
[dV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?	81	8,951	1,121

T test pour deux échantillons

T test pour deux échantillons - ([aV08] , [dV08])

Paramètres de fonctionnement

Variables	[aV08] , [dV08]
Données appariées	Non
Seuil de confiance	0,950
Type de test	Bilatéral
Différence de moyennes supposée	0,000
Variances supposées égales	Oui

Statistiques descriptives

	[aV08]	[dV08]
Moyenne observée	7,852	8,951
Variance observée	3,878	1,273
Nombre d'individus	81	81
Minimum	1,000	6,000
Maximum	10,000	10,000

Hypothèse testée

H0	moyenne([aV08]) - moyenne([dV08]) = 0
H1	moyenne([aV08]) - moyenne([dV08]) <> 0

Résultats

Valeur de la statistique F	3,047
p-valeur associée à la statistique F	0,000
T test	-4,357
p-valeur du test	0,000
Degrés de liberté	160
Intervalle de confiance de la différence des moyennes] -1,597; -0,601[
Intervalle de confiance de la variance commune] 2,092; 3,248[
Intervalle de confiance du ratio des variances] 1,960; 4,738[

Conclusion du test

On rejette l'hypothèse nulle H0

ANEXO L - CORRELAÇÃO ENTRE V07 E V08

Matrice des corrélations

Matrice des corrélations de Pearson

Variables	[aV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	[aV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?	[dV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	[dV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?
[aV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	1,000			
[aV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?	0,675	1,000		
[dV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	0,787	0,640	1,000	
[dV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?	0,502	0,683	0,509	1,000

Matrice des p-values

Variables	[aV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	[aV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?	[dV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	[dV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?
[aV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	0,000			
[aV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?	0,000	0,000		
[dV07] Qual nível de interesse pessoal você atribui para uma formação em Estatística na sua graduação ?	0,000	0,000	0,000	
[dV08] Qual o nível de utilidade você atribui para uma formação em Estatística na sua graduação ?	0,000	0,000	0,000	0,000

ANEXO M - V09 FRASES

REPRESENTATIVO	
Libellé de la réponse	Identifiant
Uma amostra grande, mais forte.	Indiv. 45
Que define alguma coisa.	Indiv. 61
O que representa a maioria dos dados.	Indiv. 22
O que tem representação significativa.	Indiv. 7
O mais expressivo.	Indiv. 75
Representação forte.	Indiv. 24
Representação do todo.	Indiv. 48
É um valor expressivo de opiniões.	Indiv. 50
Tudo o que tem representação forte.	Indiv. 24
Representa algo como um todo.	Indiv. 42
Representa o todo no geral.	Indiv. 61
Representação dos dados.	Indiv. 35
Elementos principais de uma pesquisa.	Indiv. 8
O que tem mais importância.	Indiv. 67
Representação da amostra.	Indiv. 14
Amostra para representação de uma pesquisa.	Indiv. 17
Que vai representar.	Indiv. 79

PROBABILIDADE	
Libellé de la réponse	Identifiant
Chance ou não de ocorrer provado com cálculo.	Indiv. 67
Porcentagens.	Indiv. 34
Cálculos que provam a estatística.	Indiv. 7
A tendência de algo.	Indiv. 43
Provável ou não de acontecer.	Indiv. 22
Jogos de azar.	Indiv. 1
Matemática.	Indiv. 16
Relações prováveis de acontecer.	Indiv. 4
Matemática na estatística.	Indiv. 7
Viabilidade de respostas.	Indiv. 70
Fazer uma previsão usando porcentagem.	Indiv. 10
O que pode ou não acontecer.	Indiv. 46
Chances de um sorteio, por exemplo.	Indiv. 71

SIGNIFICATIVO	
Libellé de la réponse	Identifiant
É destacado na pesquisa.	Indiv. 47
Nível de importância.	Indiv. 26
Tem relevância.	Indiv. 60
Demonstra importância alta.	Indiv. 42
A parte mais significativa.	Indiv. 23
O mais forte da pesquisa.	Indiv. 29
Valores importantes para a estatística.	Indiv. 22
Tudo que for relevante.	Indiv. 58
Sinaliza o que é importante.	Indiv. 2
Que tem o significado mais importante.	Indiv. 18

TESTES DE HIPÓTESES	
Libellé de la réponse	Identifiant
Teste da estatística.	Indiv. 24
Estatística não descritiva.	Indiv. 34
Testes estatísticos para fazer conclusões.	Indiv. 42
teste-t-Student, teste-qui-quadrado.	Indiv. 67
Testes da estatística para provar pesquisas.	Indiv. 8
Problemas resolvidos com hipóteses.	Indiv. 48
Ver se certa hipótese tem chance de ter sucesso	Indiv. 71
Estatística com inferências.	Indiv. 21
Testes para chegar em alguma conclusão.	Indiv. 17

SONDAGEM	
Libellé de la réponse	Identifiant
Aplicar questionário e coletar informações.	Indiv. 58
É uma espécie de enquete, de entrevista.	Indiv. 6
São entrevistas para fazer pesquisas.	Indiv. 16
Entrevistas.	Indiv. 35
Levantamento de dados.	Indiv. 76
Entrevista de coleta de dados.	Indiv. 6
Pesquisas para os testes estatísticos.	Indiv. 12
Coleta de dados.	Indiv. 61
Entrevistar a população e fazer uma amostra.	Indiv. 60
Levantar informações para pesquisa.	Indiv. 20
Retirada de amostra.	Indiv. 62
Informações para uma pesquisa.	Indiv. 33
Buscar dados para uma pesquisa.	Indiv. 54

ANEXO N - FRASES SOBRE ESTIMAÇÃO ANTES

Phrases longues caractéristiques pour "a_POSITIVIDADE"

a_Positividade 1

Critère de classement	Index	Libellé de la réponse	Identifiant
0,715	1	Não sei explicar.	Indiv. 27
0,715	2	Não sei explicar.	Indiv. 9
0,728	3	Não sei.	Indiv. 31
0,728	4	Não sei.	Indiv. 3
0,843	5	Algo que tenho muito carinho.	Indiv. 43
0,854	6	Não faço ideia.	Indiv. 59
0,874	7	Tem relação com previsões.	Indiv. 26
0,888	8	Estimar o resultado.	Indiv. 74
0,891	9	Um recorte de uma população.	Indiv. 66
0,916	10	Quando estima.	Indiv. 68
0,916	11	Tempo estimado.	Indiv. 65
0,928	12	Animais de estimação.	Indiv. 61
0,956	13	Para estimar, estatística para estimar.	Indiv. 10
0,956	14	Estimar.	Indiv. 28
0,957	15	Previsões.	Indiv. 25
0,957	16	Próximo.	Indiv. 47
0,957	17	Suposição.	Indiv. 58

a_Positividade 2

Critère de classement	Index	Libellé de la réponse	Identifiant
0,738	1	Teste da estatística.	Indiv. 14
0,738	2	Teste da estatística.	Indiv. 8
0,761	3	Usamos na estatística para previsões futuras.	Indiv. 11
0,790	4	Não sei explicar mas acho que é especular algo.	Indiv. 18
0,802	5	Estatística para chegar em alguma conclusão.	Indiv. 17
0,807	6	Tudo que faz previsões na estatística.	Indiv. 34
0,808	7	São avaliações baseadas em dados.	Indiv. 50
0,818	8	Onde usamos os dados para interpretar.	Indiv. 32
0,829	9	Uma estatística para fazer pesquisa.	Indiv. 13
0,863	10	Conteúdo da estatística	Indiv. 4
0,895	11	São conclusões que chegamos.	Indiv. 1
0,912	12	Estímulos de estatística.	Indiv. 67
0,913	13	Teste estatístico.	Indiv. 24
0,924	14	Um valor superficial.	Indiv. 79
0,937	15	Não sei explicar.	Indiv. 5
0,939	16	Resultados que deram.	Indiv. 30
0,946	17	Algo que criamos.	Indiv. 2
0,969	18	Proximidade.	Indiv. 73
0,969	19	Supor.	Indiv. 75
0,975	20	Previsão.	Indiv. 45
0,980	21	Estima	Indiv. 63
1,004	22	Estimar.	Indiv. 51
1,009	23	De estimar.	Indiv. 15

a_Positividade 3

Critère de classement	Index	Libellé de la réponse	Identifiant
0,733	1	Fazer previsões com a estatística.	Indiv. 33
0,738	2	Fazer estimação com os dados.	Indiv. 19
0,790	3	A estatística ocupa para realizar previsões.	Indiv. 29
0,804	4	É uma previsão que podemos fazer.	Indiv. 6
0,824	5	Fazer uma suposição.	Indiv. 69
0,844	6	Explorar e fazer conclusões.	Indiv. 35
0,875	7	Visa prever fatos.	Indiv. 23
0,910	8	Resultado da pesquisa.	Indiv. 20
0,918	9	Estima dados.	Indiv. 72
0,923	10	Estatística de inferência.	Indiv. 7
0,941	11	Valor aproximado.	Indiv. 48
0,957	12	Aproximar.	Indiv. 80
0,970	13	Estimativa.	Indiv. 76
0,973	14	Para estimar.	Indiv. 21

a_Positividade 4

Critère de classement	Index	Libellé de la réponse	Identifiant
0,724	1	Opinião sobre algo que não se tem certeza.	Indiv. 40
0,783	2	Serve para estimar e saber o que vai acontecer.	Indiv. 12
0,827	3	Faz previsões do que não conhecemos.	Indiv. 16
0,830	4	O que se espera da pesquisa.	Indiv. 81
0,853	5	Representa uma previsão, tendência significativa.	Indiv. 37
0,862	6	Um dado mais próximo de algo.	Indiv. 77
0,880	7	Estimativa não exata.	Indiv. 52
0,895	8	Apego e amor.	Indiv. 56
0,899	9	É um valor aproximado.	Indiv. 38
0,903	10	Que pode ser analisado.	Indiv. 36
0,911	11	Previsão da realidade.	Indiv. 42
0,915	12	Valor de algo.	Indiv. 46
0,932	13	Estimar uma população.	Indiv. 71
0,960	14	Teste estatístico.	Indiv. 22
0,967	15	Tendência	Indiv. 44
0,967	16	Probabilidade.	Indiv. 53
0,969	17	Estimar.	Indiv. 55
1,086	18	Não sei.	Indiv. 49

ANEXO O - FRASES SOBRE ESTIMAÇÃO DEPOIS

Phrases longues caractéristiques pour "d_POSITIVIDADE"

d_Positividade 1

Critère de classement	Index	Libellé de la réponse	Identifiant
0,787	1	Realizar um cálculo estatístico para prever alguma coisa.	Indiv. 9
0,809	2	inferência da estatística.	Indiv. 14
0,839	3	Estatística de inferência.	Indiv. 7
0,839	4	Estatística de inferência.	Indiv. 63
0,844	5	Cálculo com dados para inferir o geral.	Indiv. 50
0,857	6	Realizar uma previsão com cálculos da amostra.	Indiv. 75
0,864	7	Estatística para tirar conclusões de parâmetros.	Indiv. 51
0,879	8	O que espera acontecer no cálculo.	Indiv. 48
0,880	9	Ferramentas para fazer inferência.	Indiv. 34
0,881	10	Vai estimar o valor de uma pesquisa.	Indiv. 39
0,883	11	Para o total da população.	Indiv. 53
0,909	12	Cálculo estatístico.	Indiv. 17
0,932	13	Fazer uma estatística inferencial.	Indiv. 70
0,944	14	Interpretação de uma amostra.	Indiv. 28
0,951	15	Cálculo de estimação.	Indiv. 67
0,958	16	Concluir algo.	Indiv. 8
0,962	17	Mais próximo da realidade.	Indiv. 77
0,966	18	Inferência estatística.	Indiv. 4
0,985	19	Previsões de parâmetros.	Indiv. 3
0,987	20	Cálculo com a amostra.	Indiv. 56
0,991	21	Previsão.	Indiv. 42

d_Positividade 2

Critère de classement	Index	Libellé de la réponse	Identifiant
0,814	1	Fazer estudo com dados para prever alguma coisa.	Indiv. 15
0,831	2	É uma tendência calculada pela estatística.	Indiv. 44
0,833	3	A estatística usa para prever um acontecimento.	Indiv. 13
0,835	4	Cálculo com uma amostra para estimar uma população.	Indiv. 5
0,852	5	Uma tendência de valor de uma população.	Indiv. 73
0,890	6	Previsões que podem ter erros.	Indiv. 33
0,902	7	Amostra para a população.	Indiv. 57
0,903	8	Prever, estimar.	Indiv. 30
0,926	9	Fazer uma previsão com a amostra.	Indiv. 55
0,937	10	O mais real do verdadeiro.	Indiv. 78
0,939	11	Inferência dos dados, dos parâmetros.	Indiv. 52
0,948	12	Prever.	Indiv. 18
0,954	13	Pressupor algo.	Indiv. 20
0,956	14	Tirar conclusões.	Indiv. 62
0,956	15	Previsões.	Indiv. 25
0,963	16	Estimar, calcular.	Indiv. 38
0,981	17	Mais próximo dos parâmetros.	Indiv. 49
0,984	18	Estimar.	Indiv. 54
1,027	19	Cálculo da estatística para fazer previsões.	Indiv. 21

d_Positividade 3

Critère de classement	Index	Libellé de la réponse	Identifiant
0,790	1	É uma previsão da pesquisa de população por meio de uma amostra.	Indiv. 36
0,833	2	Usar uma amostra para fazer previsões.	Indiv. 19
0,841	3	Usamos na estatística para realizar previsões futuras.	Indiv. 11
0,854	4	Analisa uma amostra e tira conclusões.	Indiv. 27
0,855	5	É uma previsão com intervalo de confiança.	Indiv. 26
0,862	6	Coletar dados e fazer previsões.	Indiv. 31
0,863	7	A estatística ocupa para previsões pontuais ou intervalares.	Indiv. 29
0,887	8	Resultado perto do que se espera.	Indiv. 68
0,903	9	Mais perto do real.	Indiv. 47
0,914	10	Cálculo de uma pesquisa.	Indiv. 10
0,916	11	Amostra fala para o geral.	Indiv. 41
0,920	12	Nível de confiança de um intervalo.	Indiv. 37
0,930	13	O mais real possível.	Indiv. 60
0,931	14	O que prevê fatos.	Indiv. 23
0,934	15	Valor próximo da realidade.	Indiv. 46
0,943	16	Fazer uma previsão.	Indiv. 2
0,953	17	Ponto ou intervalo.	Indiv. 32
0,967	18	Previsão, intervalos, pontual.	Indiv. 64
0,992	19	Cálculo da amostra para a população.	Indiv. 80
1,017	20	Estimar.	Indiv. 74

d Positividade 4

Critère de classement	Index	Libellé de la réponse	Identifiant
0,717	1	É uma previsão (amostra) da população dentro de um intervalo ou ponto.	Indiv. 6
0,774	2	Conhecer um possível resultado dentro de um intervalo ou por uma média.	Indiv. 16
0,805	3	Estimar por ponto ou intervalo os parâmetros.	Indiv. 43
0,806	4	Cálculo da amostra usando regra pontual ou intervalar.	Indiv. 61
0,818	5	Fazer uma projeção pontual ou intervalar da população.	Indiv. 45
0,832	6	Estimar, calcular valores perto dos reais, intervalo, pontual.	Indiv. 81
0,847	7	Cálculo do parâmetro por meio da amostra.	Indiv. 69
0,854	8	Um previsão de alguma coisa (população) num intervalo numérico.	Indiv. 12
0,863	9	São previsões através de cálculos, pontuais, intervalos de confiança.	Indiv. 71
0,895	10	Estimação pontual ou intervalar.	Indiv. 35
0,928	11	Previsão, com nível de confiança.	Indiv. 22
0,934	12	Avaliar, prever, intervalo.	Indiv. 1
0,939	13	Estimar parâmetros.	Indiv. 79
0,941	14	Cálculo com a amostra para fazer previsões pontuais ou intervalares.	Indiv. 40
0,951	15	Intervalos de confiança.	Indiv. 58
0,965	16	Estimar dados.	Indiv. 72
0,968	17	Para uma amostra.	Indiv. 76
0,978	18	Cálculo mais próximo do real.	Indiv. 59

ANEXO P - V011 FRASES SOBRE UTILIDADE DE ESTATÍSTICA ANTES E DEPOIS

Phrases longues caractéristiques pour "a_POSITIVIDADE"

a_Positividade 1

Critère de classement	Index	Libellé de la réponse	Identifiant
0,804	1	Na coleta de dados e nos cálculos para saber o resultado de um trabalho.	Indiv. 61
0,846	2	Saber interpretar tabelas e gráficos e cálculos.	Indiv. 43
0,869	3	Para fazer pesquisas.	Indiv. 27
0,878	4	Análise de dados e interpretações.	Indiv. 66
0,881	5	Ler tabelas e gráficos e fazer os cálculos.	Indiv. 27
0,886	6	Para ler o mundo numérico.	Indiv. 31
0,897	7	Útil em diversas áreas.	Indiv. 65
0,898	8	Uma grande utilidade, pois é importante para a análise de dados.	Indiv. 66
0,908	9	Fazer gráficos e tabelas.	Indiv. 9
0,912	10	Estudar os números de uma pesquisa.	Indiv. 25
0,914	11	É fundamental para uma boa atuação em todas as áreas.	Indiv. 65
0,918	12	Saber fazer questionários, tabelas e gráficos.	Indiv. 31
0,924	13	Construí gráficos e tabelas no Excel.	Indiv. 68
0,924	14	Muito útil nas pesquisas para validar as coisas.	Indiv. 61
0,924	15	Olhar os números e interpretar os problemas.	Indiv. 47
0,928	16	Coletar dados e realizar pesquisas.	Indiv. 28
0,931	17	Muito útil nas pesquisas.	Indiv. 60
0,933	18	Embora eu não goste muito, é muito importante.	Indiv. 43
0,935	19	Entender uma pesquisa, por exemplo.	Indiv. 62
0,935	20	Leituras de números.	Indiv. 26
0,936	21	Nas interpretações dos dados.	Indiv. 60
0,941	22	Saber calcular gráficos.	Indiv. 3
0,944	23	Útil para interpretar dados de uma amostra.	Indiv. 9
0,944	24	Descritiva e Inferência.	Indiv. 41
0,945	25	Interpretar dados numéricos.	Indiv. 3
0,946	26	Para saber sobre pesquisas.	Indiv. 59
0,948	27	Coletar, organizar e calcular para concluir.	Indiv. 54
0,956	28	Usar em pesquisas.	Indiv. 58
0,957	29	Conteúdo importante na formação.	Indiv. 57
0,958	30	Passar informações numéricas e gráficas.	Indiv. 25

a_Positividade 2

Critère de classement	Index	Libellé de la réponse	Identifiant
0,800	1	Usar para fazer qualquer pesquisa de opinião.	Indiv. 11
0,807	2	Para fazer uma pesquisa.	Indiv. 17
0,813	3	A estatística é necessária para interpretar uma amostra e fazer pesquisas.	Indiv. 34
0,828	4	É utilizada em várias áreas para a realização de pesquisas.	Indiv. 63
0,842	5	Fazer uma pesquisa.	Indiv. 17
0,852	6	Fazer pesquisas.	Indiv. 30
0,856	7	Entender e fazer pesquisas.	Indiv. 30
0,857	8	Saber usar para tirar conclusões em uma pesquisa.	Indiv. 5
0,860	9	Fazer pesquisas.	Indiv. 8
0,869	10	Para fazer pesquisas.	Indiv. 2
0,873	11	Fazer pesquisa e saber usar as informações.	Indiv. 5
0,878	12	Coleta, interpreta dados.	Indiv. 39
0,885	13	Muito importante para qualquer pesquisa.	Indiv. 32
0,885	14	Fundamental para pesquisas em geral.	Indiv. 1
0,887	15	Para qualquer coisa existe um pouco dela.	Indiv. 45
0,900	16	Para sabermos do funcionamento, opiniões e projeções por lugares que não estamos.	Indiv. 39
0,910	17	É válido o conhecimento, nem que seja básico em estatística.	Indiv. 63
0,917	18	Essencial em uma pesquisa.	Indiv. 15
0,919	19	Para realizar qualquer pesquisa.	Indiv. 1
0,919	20	Aplicar em pesquisas.	Indiv. 14
0,922	21	Interpretações dos cálculos da matemática.	Indiv. 67
0,923	22	Usar os cálculos para pesquisar.	Indiv. 8
0,924	23	Fazer gráficos.	Indiv. 15
0,932	24	Organizar os dados de uma pesquisa para retirar informações.	Indiv. 51
0,932	25	Interpretações para uma pesquisa.	Indiv. 4
0,934	26	Conhecer as informações que são passadas.	Indiv. 24
0,934	27	Pesquisas.	Indiv. 78
0,934	28	Pesquisas.	Indiv. 78
0,935	29	Saber o que fazer com dados de uma pesquisa.	Indiv. 50
0,936	30	Elaborar e aplicar uma pesquisa.	Indiv. 14

a_Positividade 3

Critère de classement	Index	Libellé de la réponse	Identifiant
0,796	1	Ler gráficos e interpretar os resultados de uma pesquisa.	Indiv. 19
0,799	2	Tirar conclusões com os dados de uma pesquisa.	Indiv. 29
0,812	3	Levantar dados e fazer uma pesquisa.	Indiv. 33
0,848	4	Ferramenta para obter informações de dados.	Indiv. 72
0,848	5	Prever com números e ler os resultados.	Indiv. 48
0,858	6	Ferramenta fundamental para obter informações de dados.	Indiv. 72
0,864	7	Levantar dados e interpretar para concluir uma pesquisa.	Indiv. 29
0,875	8	Ler informações numéricas que são passadas.	Indiv. 35
0,879	9	Nas pesquisas.	Indiv. 69
0,892	10	Realizar interpretações para uma pesquisa.	Indiv. 7
0,898	11	Clarear as informações.	Indiv. 76
0,898	12	Importante para o dia a dia mas tenho dificuldade.	Indiv. 69
0,906	13	Trazer informações mais explicadas e com justificativa.	Indiv. 33
0,913	14	Para pesquisa em geral.	Indiv. 20
0,917	15	Coletar, interpretar dados.	Indiv. 6
0,931	16	Construí gráficos, tabelas, e estimar.	Indiv. 21
0,931	17	Interpretar dados.	Indiv. 23
0,931	18	Interpretar dados.	Indiv. 23
0,935	19	Entender informações de gráficos e tabelas.	Indiv. 19
0,936	20	Saber usar os dados encontrados em uma pesquisa.	Indiv. 6
0,941	21	Tabular dados e interpretar.	Indiv. 7
0,942	22	Realizar uma pesquisa sozinho.	Indiv. 20
0,942	23	Estudo das diversas informações.	Indiv. 35
0,946	24	É útil para o entendimento dos dados.	Indiv. 76
0,952	25	Muito útil nas empresas.	Indiv. 48
0,960	26	Pesquisas gerais.	Indiv. 80
0,964	27	Criar gráficos e tabelas.	Indiv. 21
1,045	28	Muito útil em pesquisas.	Indiv. 80

a_Positividade 4

Critère de classement	Index	Libellé de la réponse	Identifiant
0,862	1	Para coletar dados e fazer pesquisas.	Indiv. 71
0,879	2	Saber usar para os resultados de uma pesquisa.	Indiv. 37
0,885	3	Útil para todas as pesquisas.	Indiv. 42
0,891	4	Fazer inferências com dados numéricos.	Indiv. 22
0,893	5	Saber estimar e fazer previsões.	Indiv. 38
0,901	6	Coletar, interpretações de dados e de opiniões para pesquisa.	Indiv. 36
0,902	7	As pesquisas, coleta de dados para se chegar a mais dados.	Indiv. 71
0,903	8	Organização dos dados para uma pesquisa.	Indiv. 40
0,908	9	Usar os dados para o desenvolvimento de uma pesquisa.	Indiv. 55
0,912	10	Usar o Excel e construir os gráficos para meu trabalho.	Indiv. 46
0,917	11	Trabalhar com dados.	Indiv. 49
0,918	12	Buscar dados de uma amostra e prever para a população.	Indiv. 81
0,918	13	Hoje é útil em todas as formas de pesquisas.	Indiv. 42
0,927	14	Está em todo lugar, muito útil.	Indiv. 56
0,932	15	Serve para as pesquisas.	Indiv. 70
0,933	16	Resumir dados e tirar o maior proveito.	Indiv. 22
0,933	17	Conhecer as conclusões de uma pesquisa.	Indiv. 16
0,933	18	Interpretação de forma mais geral e efetiva de problemas ou estudos.	Indiv. 37
0,935	19	Acho importante principalmente na parte de estimativas.	Indiv. 38
0,941	20	Fazer cálculos e investigar.	Indiv. 52
0,946	21	Ajuda a entender coisas do cotidiano.	Indiv. 40
0,946	22	Saber interpretar os números dentro de um contexto.	Indiv. 36
0,947	23	Para tudo na vida profissional.	Indiv. 49
0,950	24	Amplamente usada.	Indiv. 77
0,953	25	Para interpretar tabelas, dados.	Indiv. 44
0,958	26	Interpretar dados.	Indiv. 53
0,966	27	Muito importante para analisar.	Indiv. 52
0,971	28	Interpretações de uma pesquisa de opinião, por exemplo.	Indiv. 16
0,971	29	Tem ampla utilidade.	Indiv. 77
0,976	30	Muito útil.	Indiv. 56

Phrases longues caractéristiques pour "d_POSITIVIDADE"

d_Positividade 1

Critère de classement	Index	Libellé de la réponse	Identifiant
0,803	1	A estatística é necessária para interpretar uma amostra e fazer pesquisas.	Indiv. 34
0,820	2	É utilizada em várias áreas para a realização de pesquisas.	Indiv. 63
0,856	3	Útil para todas as pesquisas.	Indiv. 42
0,873	4	Coleta, interpreta dados.	Indiv. 39
0,889	5	Útil para interpretar dados de uma amostra.	Indiv. 9
0,894	6	Hoje é útil em todas as formas de pesquisas.	Indiv. 42
0,898	7	Para sabermos do funcionamento, opiniões e projeções por lugares que não estamos.	Indiv. 39
0,899	8	É válido o conhecimento, nem que seja básico em estatística.	Indiv. 63
0,911	9	Fazer pesquisas.	Indiv. 8
0,915	10	Realizar interpretações para uma pesquisa.	Indiv. 7
0,916	11	Está em todo lugar, muito útil.	Indiv. 56
0,918	12	Interpretações dos cálculos da matemática.	Indiv. 67
0,928	13	Aplicar em pesquisas.	Indiv. 14
0,928	14	Conhecer as informações que são passadas.	Indiv. 24
0,929	15	Serve para as pesquisas.	Indiv. 70
0,938	16	Prever com números e ler os resultados.	Indiv. 48
0,940	17	Coletar dados e realizar pesquisas.	Indiv. 28
0,941	18	Usar os cálculos para pesquisar.	Indiv. 8
0,942	19	Organizar os dados de uma pesquisa para retirar informações.	Indiv. 51
0,942	20	Elaborar e aplicar uma pesquisa.	Indiv. 14
0,943	21	Usar seus recursos para pesquisar.	Indiv. 34
0,949	22	Muito útil.	Indiv. 56
0,951	23	Interpretações para uma pesquisa.	Indiv. 4
0,952	24	Amplamente usada.	Indiv. 77
0,958	25	Aprender a fazer escalas.	Indiv. 67
0,960	26	Muito útil nas empresas.	Indiv. 48
0,960	27	Pesquisa para o trabalho final.	Indiv. 24
0,960	28	Interpretar dados.	Indiv. 53
0,964	29	Pesquisas.	Indiv. 75
0,965	30	Tabular dados e interpretar.	Indiv. 7

d_Positividade 2

Critère de classement	Index	Libellé de la réponse	Identifiant
0,847	1	Saber estimar e fazer previsões.	Indiv. 38
0,855	2	Análise de dados e interpretações.	Indiv. 66
0,860	3	Levantar dados e fazer uma pesquisa.	Indiv. 33
0,879	4	Construí gráficos, tabelas, e estimar.	Indiv. 21
0,889	5	Uma grande utilidade, pois é importante para a análise de dados.	Indiv. 66
0,891	6	Trabalhar com dados.	Indiv. 49
0,892	7	Entender uma pesquisa, por exemplo.	Indiv. 62
0,896	8	Estudar os números de uma pesquisa.	Indiv. 25
0,899	9	Fazer cálculos e investigar.	Indiv. 52
0,905	10	Analisar dados.	Indiv. 73
0,917	11	Fazer gráficos, tabelas e interpretar.	Indiv. 13
0,921	12	Acho importante principalmente na parte de estimativas.	Indiv. 38
0,921	13	Tratar informações.	Indiv. 18
0,922	14	Fazer pesquisa e saber usar as informações.	Indiv. 5
0,923	15	Saber usar para tirar conclusões em uma pesquisa.	Indiv. 5
0,924	16	Para tudo na vida profissional.	Indiv. 49
0,925	17	Usar os dados para o desenvolvimento de uma pesquisa.	Indiv. 55
0,925	18	Essencial em uma pesquisa.	Indiv. 15
0,927	19	Trazer informações mais explicadas e com justificativa.	Indiv. 33
0,936	20	Muito importante para analisar.	Indiv. 52
0,938	21	Análise de dados.	Indiv. 73
0,938	22	Para pesquisa em geral.	Indiv. 20
0,940	23	Saber calcular e interpretar informações.	Indiv. 18
0,941	24	Passar informações numéricas e gráficas.	Indiv. 25
0,941	25	Para interpretar tabelas, dados.	Indiv. 44
0,942	26	Conteúdo importante na formação.	Indiv. 57
0,945	27	Utilizamos em quase tudo.	Indiv. 62
0,947	28	Entender e fazer pesquisas.	Indiv. 30
0,947	29	Criar gráficos e tabelas.	Indiv. 21
0,950	30	Fazer gráficos.	Indiv. 15

d_Positividade 3

Critère de classement	Index	Libellé de la réponse	Identifiant
0,780	1	Ler gráficos e interpretar os resultados de uma pesquisa.	Indiv. 19
0,821	2	Muito importante para qualquer pesquisa.	Indiv. 32
0,838	3	Saber usar para os resultados de uma pesquisa.	Indiv. 37
0,846	4	Para fazer pesquisas.	Indiv. 27
0,851	5	Ler tabelas e gráficos e fazer os cálculos.	Indiv. 27
0,860	6	Coletar, interpretações de dados e de opiniões para pesquisa.	Indiv. 36
0,872	7	Usar o Excel e construir os gráficos para meu trabalho.	Indiv. 46
0,877	8	Olhar os números e interpretar os problemas.	Indiv. 47
0,882	9	Diversas utilidades, entre elas, interpretar gráficos e tabelas.	Indiv. 32
0,882	10	Para ler o mundo numérico.	Indiv. 31
0,894	11	Útil em diversas áreas.	Indiv. 65
0,904	12	Construí gráficos e tabelas no Excel.	Indiv. 68
0,909	13	Tirar conclusões com os dados de uma pesquisa.	Indiv. 29
0,911	14	Entender os processos.	Indiv. 64
0,915	15	Saber interpretar os números dentro de um contexto.	Indiv. 36
0,921	16	Realização de uma pesquisa.	Indiv. 10
0,923	17	Interpretação de forma mais geral e efetiva de problemas ou estudos.	Indiv. 37
0,925	18	É fundamental para uma boa atuação em todas as áreas.	Indiv. 65
0,927	19	Saber fazer questionários, tabelas e gráficos.	Indiv. 31
0,930	20	Usar para fazer qualquer pesquisa de opinião.	Indiv. 11
0,934	21	Leituras de números.	Indiv. 26
0,944	22	Pesquisas gerais.	Indiv. 80
0,944	23	Levantar dados e interpretar para concluir uma pesquisa.	Indiv. 29
0,950	24	Entender informações de gráficos e tabelas.	Indiv. 19
0,950	25	Inúmeras utilidades.	Indiv. 74
0,950	26	Inúmeras utilidades.	Indiv. 74
0,951	27	Descritiva e Inferência.	Indiv. 41
0,951	28	Nas interpretações dos dados.	Indiv. 60
0,952	29	Para fazer pesquisas.	Indiv. 2
0,954	30	Muito útil nas pesquisas.	Indiv. 60

d_Positividade 4

Critère de classement	Index	Libellé de la réponse	Identifiant
0,796	1	Na coleta de dados e nos cálculos para saber o resultado de um trabalho.	Indiv. 61
0,839	2	Para coletar dados e fazer pesquisas.	Indiv. 71
0,862	3	Para qualquer coisa existe um pouco dela.	Indiv. 45
0,868	4	Nas pesquisas.	Indiv. 69
0,871	5	Ferramenta para obter informações de dados.	Indiv. 72
0,872	6	Ferramenta fundamental para obter informações de dados.	Indiv. 72
0,884	7	Organização dos dados para uma pesquisa.	Indiv. 40
0,891	8	Importante para o dia a dia mas tenho dificuldade.	Indiv. 69
0,894	9	As pesquisas, coleta de dados para se chegar a mais dados.	Indiv. 71
0,895	10	Fazer inferências com dados numéricos.	Indiv. 22
0,898	11	Clarear as informações.	Indiv. 76
0,899	12	Muito útil nas pesquisas para validar as coisas.	Indiv. 61
0,912	13	Saber interpretar tabelas e gráficos e cálculos.	Indiv. 43
0,912	14	Buscar dados de uma amostra e prever para a população.	Indiv. 81
0,925	15	Ler informações numéricas que são passadas.	Indiv. 35
0,926	16	Fundamental para pesquisas em geral.	Indiv. 1
0,929	17	Coletar, interpretar dados.	Indiv. 6
0,931	18	Resumir dados e tirar o maior proveito.	Indiv. 22
0,932	19	Serve para coletar dados e calcular.	Indiv. 79
0,933	20	Embora eu não goste muito, é muito importante.	Indiv. 43
0,933	21	Ajuda a entender coisas do cotidiano.	Indiv. 40
0,936	22	Importante mas não sei definir.	Indiv. 45
0,938	23	É útil para o entendimento dos dados.	Indiv. 76
0,948	24	Elaborar, calcular e interpretar informações para uma pesquisa.	Indiv. 12
0,948	25	Para saber sobre pesquisas.	Indiv. 59
0,956	26	Usar em pesquisas.	Indiv. 58
0,960	27	Estudo das diversas informações.	Indiv. 35
0,961	28	Saber usar os dados encontrados em uma pesquisa.	Indiv. 6
0,965	29	Conhecer as conclusões de uma pesquisa.	Indiv. 16
0,965	30	Interpretar uma pesquisa.	Indiv. 12

ANEXO Q - aV14 E dV14 X POSITIVIDADE

[aV14]Você já obteve conhecimentos em Estatística ? / a_POSITIVIDADE

Nombre de cases d'effectif théorique < 5	7
Nombre de degrés de liberté	9
Khi-2	31,752
P-value	0,000
T de Tschuprow	0,361
V de Cramer	0,361

Effectif / poids					Ensemble
	a_Positividade 1	a_Positividade 2	a_Positividade 3	a_Positividade 4	
% en ligne					
% en colonne					
1_Nunca	4	0	1	1	6
	66,7	0,0	16,7	16,7	100,0
	18,2	0,0	7,1	5,3	7,4
2_ Em um nível muito superficial	11	9	2	2	24
	45,8	37,5	8,3	8,3	100,0
	50,0	34,6	14,3	10,5	29,6
3_ Em um nível superficial	5	10	4	1	20
	25,0	50,0	20,0	5,0	100,0
	22,7	38,5	28,6	5,3	24,7
4_Introdução ao básico	2	7	7	15	31
	6,5	22,6	22,6	48,4	100,0
	9,1	26,9	50,0	78,9	38,3
Ensemble	22	26	14	19	81
	27,2	32,1	17,3	23,5	100,0
	100,0	100,0	100,0	100,0	100,0

[dV14]Você já obteve conhecimentos em Estatística ? / d_POSITIVIDADE

Nombre de cases d'effectif théorique < 5	1
Nombre de degrés de liberté	3
Khi-2	1,083
P-value	0,781
T de Tschuprow	0,088
V de Cramer	0,116

Effectif / poids					Ensemble
	d_Positividade 1	d_Positividade 2	d_Positividade 3	d_Positividade 4	
% en ligne					
% en colonne					
3_ Em um nível superficial	5	7	6	4	22
	22,7	31,8	27,3	18,2	100,0
	22,7	35,0	28,6	22,2	27,2
4_Introdução ao básico	17	13	15	14	59
	28,8	22,0	25,4	23,7	100,0
	77,3	65,0	71,4	77,8	72,8
Ensemble	22	20	21	18	81
	27,2	24,7	25,9	22,2	100,0
	100,0	100,0	100,0	100,0	100,0

ANEXO R - TODAS AS ATITUDES

ATT(-)01_[aV15a]1. Eu não tenho um bom conhecimento em Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	3	3,7 %
2_Não estou bem de acordo.	21	25,9 %
3_Estou de acordo.	44	54,3 %
4_Concordo plenamente.	13	16,0 %
Ensemble	81	100,0 %

ATT(+)02_[aV15b]2. A Estatística é algo que aprecio muito.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	3	3,7 %
2_Não estou bem de acordo.	49	60,5 %
3_Estou de acordo.	24	29,6 %
4_Concordo plenamente.	5	6,2 %
Ensemble	81	100,0 %

ATT(-)03_[aV15c]3. Eu me sinto inseguro quando preciso utilizar a Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	6	7,4 %
2_Não estou bem de acordo.	13	16,0 %
3_Estou de acordo.	42	51,9 %
4_Concordo plenamente.	20	24,7 %
Ensemble	81	100,0 %

ATT(+)04_[aV15d]4. Eu tenho uma relação muito boa com a Estatística : eu gosto e aprecio esta disciplina.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	5	6,2 %
2_Não estou bem de acordo.	42	51,9 %
3_Estou de acordo.	31	38,3 %
4_Concordo plenamente.	3	3,7 %
Ensemble	81	100,0 %

ATT(-)05_[aV15e]5. Quando estudo Estatística, não consigo pensar claramente pois minha cabeça fica bloqueada.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	11	13,6 %
2_Não estou bem de acordo.	23	28,4 %
3_Estou de acordo.	31	38,3 %
4_Concordo plenamente.	16	19,8 %
Ensemble	81	100,0 %

ATT(-)06_[aV15f]6. Quando eu escuto a palavra Estatística eu sinto um sentimento de aversão.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	22	27,2 %
2_Não estou bem de acordo.	45	55,6 %
3_Estou de acordo.	12	14,8 %
4_Concordo plenamente.	2	2,5 %
Ensemble	81	100,0 %

ATT(-)07_[aV15g]7. Eu realmente não gosto de matemática.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	23	28,4 %
2_Não estou bem de acordo.	23	28,4 %
3_Estou de acordo.	25	30,9 %
4_Concordo plenamente.	10	12,3 %
Ensemble	81	100,0 %

ATT(-)08_[aV15h]8. Pensar em resolver um problema de Estatística me deixa nervoso.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	10	12,3 %
2_Não estou bem de acordo.	30	37,0 %

3_Estou de acordo.	29	35,8 %
4_Concordo plenamente.	12	14,8 %
Ensemble	81	100,0 %

ATT(+)⁰⁹ [aV15i]9. A Estatística é fascinante e divertida.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	10	12,3 %
2_Não estou bem de acordo.	39	48,1 %
3_Estou de acordo.	28	34,6 %
4_Concordo plenamente.	4	4,9 %
Ensemble	81	100,0 %

ATT(-)¹⁰ [aV15j]10. Eu nunca gostei de Estatística, ela é a disciplina que me causa mais medo.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	21	25,9 %
2_Não estou bem de acordo.	46	56,8 %
3_Estou de acordo.	13	16,0 %
4_Concordo plenamente.	1	1,2 %
Ensemble	81	100,0 %

ATT(+)¹¹ [aV15k]11. O sentimento que experimento em relação a Estatística é bom.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	3	3,7 %
2_Não estou bem de acordo.	33	40,7 %
3_Estou de acordo.	35	43,2 %
4_Concordo plenamente.	10	12,3 %
Ensemble	81	100,0 %

ATT(-)¹² [aV15l]12. Eu tenho medo de Estatística pois tenho medo de Matemática.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	20	24,7 %
2_Não estou bem de acordo.	29	35,8 %
3_Estou de acordo.	26	32,1 %
4_Concordo plenamente.	6	7,4 %
Ensemble	81	100,0 %

ATT(+)¹³ [aV15m]13. Eu acho a Estatística muito interessante e gosto de ir na aula dessa disciplina.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	5	6,2 %
2_Não estou bem de acordo.	22	27,2 %
3_Estou de acordo.	38	46,9 %
4_Concordo plenamente.	16	19,8 %
Ensemble	81	100,0 %

ATT(+)¹⁴ [aV15n]14. A Estatística me dá segurança e me estimula.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	13	16,0 %
2_Não estou bem de acordo.	48	59,3 %
3_Estou de acordo.	15	18,5 %
4_Concordo plenamente.	5	6,2 %
Ensemble	81	100,0 %

ATT(-)¹⁵ [aV15o]15. A Estatística me dá a impressão que estou perdido em uma sala cheia de números na qual eu não acho saída.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	15	18,5 %
2_Não estou bem de acordo.	47	58,0 %
3_Estou de acordo.	11	13,6 %
4_Concordo plenamente.	8	9,9 %
Ensemble	81	100,0 %

ATT(+)₁₆ [aV15p]16. Eu me sinto bem com a Estatística e gosto muito dessa disciplina.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	8	9,9 %
2_Não estou bem de acordo.	41	50,6 %
3_Estou de acordo.	29	35,8 %
4_Concordo plenamente.	3	3,7 %
Ensemble	81	100,0 %

ATT(-)₁₇ [aV15q]17. Eu estou sempre estressado durante as aulas de Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	50	61,7 %
2_Não estou bem de acordo.	27	33,3 %
3_Estou de acordo.	3	3,7 %
4_Concordo plenamente.	1	1,2 %
Ensemble	81	100,0 %

ATT(+)₁₈ [aV15r]18. A Estatística é uma das matérias que eu realmente gosto de estudar na universidade ou na escola.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	19	23,5 %
2_Não estou bem de acordo.	40	49,4 %
3_Estou de acordo.	21	25,9 %
4_Concordo plenamente.	1	1,2 %
Ensemble	81	100,0 %

ATT(-)₁₉ [aV15s]19. Eu não gosto de Estatística e nem de ir nas aulas de Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	45	55,6 %
2_Não estou bem de acordo.	31	38,3 %
3_Estou de acordo.	4	4,9 %
4_Concordo plenamente.	1	1,2 %
Ensemble	81	100,0 %

ATT(-)₂₀ [aV15t]20. A Estatística me provoca um sentimento de indecisão que resulta no meu medo de não ser capaz de ter êxito na disciplina.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	15	18,5 %
2_Não estou bem de acordo.	32	39,5 %
3_Estou de acordo.	27	33,3 %
4_Concordo plenamente.	7	8,6 %
Ensemble	81	100,0 %

ATT(+)₂₁ [aV15u]21. Eu me sinto mais feliz com a Estatística do que com as outras disciplinas.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	18	22,2 %
2_Não estou bem de acordo.	41	50,6 %
3_Estou de acordo.	17	21,0 %
4_Concordo plenamente.	5	6,2 %
Ensemble	81	100,0 %

ATT(-)₂₂ [aV15v]22. A Estatística me deixa inquieto, descontente, irritado e impaciente.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	40	49,4 %
2_Não estou bem de acordo.	31	38,3 %
3_Estou de acordo.	6	7,4 %
4_Concordo plenamente.	4	4,9 %
Ensemble	81	100,0 %

ATT(+)₂₃ [aV15w]23. Eu sou competente em Matemática.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	22	27,2 %
2_Não estou bem de acordo.	43	53,1 %

3_Estou de acordo.	14	17,3 %
4_Concordo plenamente.	2	2,5 %
Ensemble	81	100,0 %

ATT(+)-24 [aV15x]24. Eu realmente gosto de Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	4	4,9 %
2_Não estou bem de acordo.	36	44,4 %
3_Estou de acordo.	29	35,8 %
4_Concordo plenamente.	12	14,8 %
Ensemble	81	100,0 %

ATT(-)01 [dV15a]1. Eu não tenho um bom conhecimento em Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	6	7,4 %
2_Não estou bem de acordo.	67	82,7 %
3_Estou de acordo.	8	9,9 %
Ensemble	81	100,0 %

ATT(+)-02 [dV15b]2. A Estatística é algo que aprecio muito.

Modalités	Effectifs	% sur exprimés
2_Não estou bem de acordo.	7	8,6 %
3_Estou de acordo.	65	80,2 %
4_Concordo plenamente.	9	11,1 %
Ensemble	81	100,0 %

ATT(-)03 [dV15c]3. Eu me sinto inseguro quando preciso utilizar a Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	4	4,9 %
2_Não estou bem de acordo.	65	80,2 %
3_Estou de acordo.	12	14,8 %
Ensemble	81	100,0 %

ATT(+)-04 [dV15d]4. Eu tenho uma relação muito boa com a Estatística : eu gosto e aprecio esta disciplina.

Modalités	Effectifs	% sur exprimés
2_Não estou bem de acordo.	2	2,5 %
3_Estou de acordo.	59	72,8 %
4_Concordo plenamente.	20	24,7 %
Ensemble	81	100,0 %

ATT(-)05 [dV15e]5. Quando estudo Estatística, não consigo pensar claramente pois minha cabeça fica bloqueada.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	3	3,7 %
2_Não estou bem de acordo.	69	85,2 %
3_Estou de acordo.	9	11,1 %
Ensemble	81	100,0 %

ATT(-)06 [dV15f]6. Quando eu escuto a palavra Estatística eu sinto um sentimento de aversão.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	31	38,3 %
2_Não estou bem de acordo.	50	61,7 %
Ensemble	81	100,0 %

ATT(-)07 [dV15g]7. Eu realmente não gosto de matemática.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	10	12,3 %
2_Não estou bem de acordo.	29	35,8 %
3_Estou de acordo.	40	49,4 %
4_Concordo plenamente.	2	2,5 %

Ensemble	81	100,0 %
----------	----	---------

ATT(-)08 [dV15h]8. Pensar em resolver um problema de Estatística me deixa nervoso.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	22	27,2 %
2_Não estou bem de acordo.	41	50,6 %
3_Estou de acordo.	17	21,0 %
4_Concordo plenamente.	1	1,2 %
Ensemble	81	100,0 %

ATT(+)09 [dV15i]9. A Estatística é fascinante e divertida.

Modalités	Effectifs	% sur exprimés
2_Não estou bem de acordo.	13	16,0 %
3_Estou de acordo.	56	69,1 %
4_Concordo plenamente.	12	14,8 %
Ensemble	81	100,0 %

ATT(-)10 [dV15j]10. Eu nunca gostei de Estatística, ela é a disciplina que me causa mais medo.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	19	23,5 %
2_Não estou bem de acordo.	56	69,1 %
3_Estou de acordo.	6	7,4 %
Ensemble	81	100,0 %

ATT(+)11 [dV15k]11. O sentimento que experimento em relação a Estatística é bom.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	7	8,6 %
2_Não estou bem de acordo.	10	12,3 %
3_Estou de acordo.	59	72,8 %
4_Concordo plenamente.	5	6,2 %
Ensemble	81	100,0 %

ATT(-)12 [dV15l]12. Eu tenho medo de Estatística pois tenho medo de Matemática.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	3	3,7 %
2_Não estou bem de acordo.	33	40,7 %
3_Estou de acordo.	36	44,4 %
4_Concordo plenamente.	9	11,1 %
Ensemble	81	100,0 %

ATT(+)13 [dV15m]13. Eu acho a Estatística muito interessante e gosto de ir na aula dessa disciplina.

Modalités	Effectifs	% sur exprimés
2_Não estou bem de acordo.	3	3,7 %
3_Estou de acordo.	57	70,4 %
4_Concordo plenamente.	21	25,9 %
Ensemble	81	100,0 %

ATT(+)14 [dV15n]14. A Estatística me dá segurança e me estimula.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	1	1,2 %
2_Não estou bem de acordo.	18	22,2 %
3_Estou de acordo.	48	59,3 %
4_Concordo plenamente.	14	17,3 %
Ensemble	81	100,0 %

ATT(-)15 [dV15o]15. A Estatística me dá a impressão que estou perdido em uma sala cheia de números na qual eu não acho saída.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	32	39,5 %
2_Não estou bem de acordo.	46	56,8 %

3_Estou de acordo.	3	3,7 %
Ensemble	81	100,0 %

ATT(+)₁₆ [dV15p]16. Eu me sinto bem com a Estatística e gosto muito dessa disciplina.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	1	1,2 %
2_Não estou bem de acordo.	7	8,6 %
3_Estou de acordo.	52	64,2 %
4_Concordo plenamente.	21	25,9 %
Ensemble	81	100,0 %

ATT(-)₁₇ [dV15q]17. Eu estou sempre estressado durante as aulas de Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	47	58,0 %
2_Não estou bem de acordo.	32	39,5 %
3_Estou de acordo.	2	2,5 %
Ensemble	81	100,0 %

ATT(+)₁₈ [dV15r]18. A Estatística é uma das matérias que eu realmente gosto de estudar na universidade ou na escola.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	3	3,7 %
2_Não estou bem de acordo.	33	40,7 %
3_Estou de acordo.	42	51,9 %
4_Concordo plenamente.	3	3,7 %
Ensemble	81	100,0 %

ATT(-)₁₉ [dV15s]19. Eu não gosto de Estatística e nem de ir nas aulas de Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	52	64,2 %
2_Não estou bem de acordo.	29	35,8 %
Ensemble	81	100,0 %

ATT(-)₂₀ [dV15t]20. A Estatística me provoca um sentimento de indecisão que resulta no meu medo de não ser capaz de ter êxito na disciplina.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	28	34,6 %
2_Não estou bem de acordo.	53	65,4 %
Ensemble	81	100,0 %

ATT(+)₂₁ [dV15u]21. Eu me sinto mais feliz com a Estatística do que com as outras disciplinas.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	4	4,9 %
2_Não estou bem de acordo.	28	34,6 %
3_Estou de acordo.	40	49,4 %
4_Concordo plenamente.	9	11,1 %
Ensemble	81	100,0 %

ATT(-)₂₂ [dV15v]22. A Estatística me deixa inquieto, descontente, irritado e impaciente.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	50	61,7 %
2_Não estou bem de acordo.	31	38,3 %
Ensemble	81	100,0 %

ATT(+)₂₃ [dV15w]23. Eu sou competente em Matemática.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	20	24,7 %
2_Não estou bem de acordo.	47	58,0 %
3_Estou de acordo.	10	12,3 %
4_Concordo plenamente.	4	4,9 %

Ensemble	81	100,0 %
----------	----	---------

ATT(+)²⁴. [dV15x]²⁴. Eu realmente gosto de Estatística.

Modalités	Effectifs	% sur exprimés
1_Absolutamente contra.	2	2,5 %
2_Não estou bem de acordo.	7	8,6 %
3_Estou de acordo.	62	76,5 %
4_Concordo plenamente.	10	12,3 %
Ensemble	81	100,0 %

ANEXO S - TESTE T - aV15 E dV15

T test pour deux échantillons

T test pour deux échantillons - (ANTES Nivel de positividade, DEPOIS Nivel de positividade)

Paramètres de fonctionnement

Variables	ANTES Nivel de positividade, DEPOIS Nivel de positividade	
Données appariées	Non	
Seuil de confiance	0,950	
Type de test	Bilatéral	
Différence de moyennes supposée	0,000	
Variances supposées égales	Oui	

Statistiques descriptives

	ANTES Nivel de positividad	DEPOIS Nivel de positividade
Moyenne observée	0,533	0,671
Variance observée	0,024	0,002
Nombre d'individus	81	81
Minimum	0,125	0,569
Maximum	0,944	0,806

Hypothèse testée

H0	moyenne(ANTES Nivel de positividade) - moyenne(DEPOIS Nivel de positividade) = 0
H1	moyenne(ANTES Nivel de positividade) - moyenne(DEPOIS Nivel de positividade) <> 0

Résultats

Valeur de la statistique F	12,285
p-valeur associée à la statistique F	0,000
T test	-7,828
p-valeur du test	0,000
Degrés de liberté	160
Intervalle de confiance de la différence des moyennes] -0,174; -0,104[
Intervalle de confiance de la variance commune] 0,010; 0,016[
Intervalle de confiance du ratio des variances] 7,901; 19,102[

Conclusion du test

On rejette l'hypothèse nulle H0

ANEXO T - POSITIVIDADE X SEXO

[aV01] Sexo / a_POSITIVIDADE

Nombre de cases d'effectif théorique < 5	0
Nombre de degrés de liberté	3
Khi-2	6,673
P-value	0,083
T de Tschuprow	0,218
V de Cramer	0,287

Effectif/poids		a_Positividade 1	a_Positividade 2	a_Positividade 3	a_Positividade 4	Ensemble
% en ligne	% en colonne					
Feminino		12	11	11	7	41
		29,3	26,8	26,8	17,1	100,0
		54,5	42,3	78,6	36,8	50,6
Masculino		10	15	3	12	40
		25,0	37,5	7,5	30,0	100,0
		45,5	57,7	21,4	63,2	49,4
Ensemble		22	26	14	19	81
		27,2	32,1	17,3	23,5	100,0
		100,0	100,0	100,0	100,0	100,0

[aV01] Sexo / d_POSITIVIDADE

Nombre de cases d'effectif théorique < 5	0
Nombre de degrés de liberté	3
Khi-2	5,119
P-value	0,163
T de Tschuprow	0,191
V de Cramer	0,251

Effectif/poids		d_Positividade 1	d_Positividade 2	d_Positividade 3	d_Positividade 4	Ensemble
% en ligne	% en colonne					
Feminino		9	8	11	13	41
		22,0	19,5	26,8	31,7	100,0
		40,9	40,0	52,4	72,2	50,6
Masculino		13	12	10	5	40
		32,5	30,0	25,0	12,5	100,0
		59,1	60,0	47,6	27,8	49,4
Ensemble		22	20	21	18	81
		27,2	24,7	25,9	22,2	100,0
		100,0	100,0	100,0	100,0	100,0

ANEXO U - V16 DEFINIÇÃO PARA ESTATÍSTICA

Phrases longues caractéristiques pour "a_POSITIVIDADE"

a_Positividade 1

Critère de classement	Index	Libellé de la réponse	Identifiant
0,722	1	Uma forma simples para quem sabe resolver e muito complicado para quem não domina.	Indiv. 41
0,844	2	Ciência que coleta, trata e calcula seus resultados para uma amostra.	Indiv. 66
0,882	3	Não gosto muito, possuo dificuldades.	Indiv. 62
0,883	4	Estatística é uma disciplina bem difícil.	Indiv. 65
0,888	5	Necessária e importante para o mundo.	Indiv. 43
0,893	6	Trata uma amostra de uma população.	Indiv. 31
0,898	7	Trabalhos importantes como o TFG.	Indiv. 58
0,904	8	Tenho interesse mas tenho dificuldade.	Indiv. 60
0,904	9	Associo com quantificação e porcentagem.	Indiv. 61
0,906	10	Ajuda nas pesquisas do TFG.	Indiv. 3
0,911	11	Interpreta dados e constrói tabelas e gráficos.	Indiv. 9
0,922	12	Coleta dados e faz a interpretação.	Indiv. 25
0,927	13	Estatística estuda fenômenos.	Indiv. 28
0,928	14	Estuda as informações de amostras para pesquisas.	Indiv. 10
0,935	15	Medo e insegurança.	Indiv. 47
0,947	16	Pode prever acontecimentos.	Indiv. 26
0,952	17	Descritiva e de inferência.	Indiv. 27
0,960	18	Conhecimentos gerais.	Indiv. 59

a_Positividade 2

Critère de classement	Index	Libellé de la réponse	Identifiant
0,748	1	Trabalha com números de dados.	Indiv. 1
0,758	2	É uma ciência que trabalha com pesquisas.	Indiv. 4
0,784	3	Trabalha com informações que são tratadas para verificar alguma pesquisa.	Indiv. 24
0,822	4	Trabalha com pesquisas.	Indiv. 8
0,836	5	É importante mas não me adapto bem com ela.	Indiv. 67
0,844	6	Importante para verificar opiniões.	Indiv. 75
0,845	7	Conteúdo para fazer tabelas e gráficos e prever.	Indiv. 2
0,848	8	Ciência que estuda os dados para realizar uma pesquisa.	Indiv. 17
0,855	9	Criar gráficos e tabelas a partir de números pesquisados.	Indiv. 34
0,861	10	Uma das cadeiras mais difíceis e mais importantes do curso.	Indiv. 79
0,864	11	Pegar vários dados e ajeitar para um entendimento.	Indiv. 78
0,866	12	Descreve opiniões com o tratamentos de dados.	Indiv. 32
0,894	13	Muitas contas para fazer no Excel.	Indiv. 64
0,900	14	Trabalha com probabilidades.	Indiv. 18
0,902	15	Ciência onde estudamos os dados.	Indiv. 50
0,902	16	Estuda as opiniões das pessoas.	Indiv. 5
0,905	17	Analisa dados e gera informações.	Indiv. 73
0,917	18	Coleta, trata e interpreta dados de uma amostra.	Indiv. 30
0,925	19	Uma disciplina importante.	Indiv. 45
0,931	20	Organizar os números.	Indiv. 51
0,942	21	Trata de informações.	Indiv. 11
0,944	22	Pode fazer previsões.	Indiv. 14
0,960	23	Descritiva e inferência.	Indiv. 15
0,982	24	Probabilidade.	Indiv. 63

a_Positividade 3

Critère de classement	Index	Libellé de la réponse	Identifiant
0,785	1	Ajuda nas coleta de dados para fazer uma pesquisa. (Intenção de voto)	Indiv. 33
0,815	2	Coleta informações e calcula para saber os resultados.	Indiv. 6
0,815	3	Estuda os dados de uma amostra para tirar conclusões.	Indiv. 21
0,815	4	Um elemento essencial no nosso dia a dia.	Indiv. 72
0,862	5	Matéria que estuda como fazer pesquisas.	Indiv. 23
0,876	6	Resume grandes informações de uma pesquisa.	Indiv. 20
0,878	7	Métodos de calcular e apresentar resultados.	Indiv. 80
0,890	8	Ciência que estuda, interpreta e elabora.	Indiv. 76
0,891	9	Faz cálculos para interpretar pesquisas.	Indiv. 19
0,902	10	Estuda as opiniões de entrevistas.	Indiv. 35
0,948	11	Busca estimativas.	Indiv. 48
0,996	12	É o estudo de tabelas e gráficos.	Indiv. 69
1,046	13	Trabalha com dados.	Indiv. 7

a_Positividade 4

Critère de classement	Index	Libellé de la réponse	Identifiant
0,814	1	É uma forma de interpretar os dados levantados em algum tipo de pesquisa.	Indiv. 42
0,834	2	É uma pesquisa que nos possibilita obter um resultado estimado.	Indiv. 38
0,850	3	Ciência da avaliação de dados que explicam resultados de modo geral.	Indiv. 37
0,857	4	É o estudo aproximado, hipóteses das coisas.	Indiv. 49
0,864	5	É o estudo dos dados.	Indiv. 71
0,865	6	São números estudados para se calcular dados.	Indiv. 77
0,872	7	É a ciência que concilia a exatidão matemática.	Indiv. 44
0,881	8	Tudo que tratar de coleta e análise de dados.	Indiv. 40
0,884	9	Problemas foram feitos para serem resolvidos.	Indiv. 36
0,893	10	São medidas de média, moda, mediana.	Indiv. 16
0,901	11	Fórmulas para calcularmos dados fornecidos.	Indiv. 56
0,911	12	Análise de dados.	Indiv. 53
0,917	13	Conteúdo que mistura números e interpretações.	Indiv. 12
0,930	14	Descritiva, descreve dados numéricos.	Indiv. 22
0,960	15	Analisar dados.	Indiv. 52

Phrases longues caractéristiques pour "d_POSITIVIDADE"

d_Positividade 1

Critère de classement	Index	Libellé de la réponse	Identifiant
0,838	1	Estuda medidas como média, desvio padrão e utiliza para fazer inferências.	Indiv. 14
0,854	2	Utiliza opiniões ou outros dados para criar gráficos, tabelas e interpretá-los.	Indiv. 9
0,877	3	Coleta e organiza os dados para uma pesquisa.	Indiv. 51
0,879	4	Ciência que trabalha com variadas pesquisas podendo fazer previsões importantes.	Indiv. 4
0,896	5	Faz a coleta de informações, chamada de amostra, para realizar inferências.	Indiv. 24
0,903	6	Coleta dados e interpreta-os.	Indiv. 48
0,907	7	Ciência onde se estuda os dados.	Indiv. 50
0,907	8	Está preocupada com pesquisas em gerais.	Indiv. 63
0,918	9	Serve para coletar, interpretar dados para uma pesquisa.	Indiv. 39
0,919	10	Faz cálculos e dados através de uma amostra para pesquisas.	Indiv. 70
0,920	11	Usa a matemática para fazer análises.	Indiv. 56
0,920	12	Ciência que está presente em diversos fenômenos.	Indiv. 28
0,921	13	Trabalha com pesquisas, fazendo previsões.	Indiv. 8
0,926	14	Ajuda na coleta e tratamento de informações para uma pesquisa.	Indiv. 7
0,943	15	Disciplina que forma melhor o raciocínio.	Indiv. 53
0,943	16	Importante para verificar opiniões.	Indiv. 75
0,945	17	Ocupada principalmente em pesquisas.	Indiv. 34
0,948	18	Responsável em desenvolver pesquisas.	Indiv. 77
0,961	19	Analisar dados e projetar.	Indiv. 67
0,974	20	Tratamento de dados coletados.	Indiv. 3
1,067	21	Coleta, calcula e interpreta dados de uma pesquisa.	Indiv. 42
1,073	22	Ciência que trata de dados de uma amostra para tirar conclusões.	Indiv. 17

d_Positividade 2

Critère de classement	Index	Libellé de la réponse	Identifiant
0,781	1	Trata dados de uma pesquisa.	Indiv. 30
0,812	2	Coleta dados, trata e transforma em informações importantes que respondem alguma pesquisa.	Indiv. 21
0,815	3	Coleta e interpreta dados.	Indiv. 54
0,817	4	Coleta e interpreta dados das pesquisas.	Indiv. 52
0,845	5	Coleta e trata dados provenientes de uma amostra para tirar conclusões.	Indiv. 33
0,858	6	Ciência que coleta, trata e calcula dados.	Indiv. 66
0,865	7	Ciência usada para coletar, tratar e estimar dados de uma pesquisa.	Indiv. 15
0,867	8	Responsável por resumir e interpretar dados, podendo fazer estimativas.	Indiv. 20
0,870	9	Interpreta resultados das amostras em uma pesquisa.	Indiv. 78
0,874	10	Ciência que coleta dados, interpreta e conclui.	Indiv. 13
0,882	11	Ciência fundamental para o tratamento de dados numéricos de uma pesquisa.	Indiv. 25
0,900	12	Levanta dados para uma pesquisa.	Indiv. 44
0,911	13	Gera informações através de dados de uma amostra.	Indiv. 73
0,931	14	É uma disciplina que usamos para pesquisas.	Indiv. 62
0,945	15	Interpreta informações numéricas.	Indiv. 49
0,946	16	Trabalha com entrevistas e opiniões.	Indiv. 18
0,955	17	Pode ser descritiva ou inferência.	Indiv. 5
0,956	18	Estuda as pesquisas nas diversas áreas.	Indiv. 38

d_Positividade 3

Critère de classement	Index	Libellé de la réponse	Identifiant
0,783	1	Trabalha com amostra para tirar conclusões na realização de uma pesquisa.	Indiv. 2
0,812	2	É uma ciência responsável em tratar dados e tirar conclusões através deles.	Indiv. 31
0,822	3	Pode ser definida com a ciência que descreve e tira conclusões de dados.	Indiv. 29
0,868	4	Ciência que estuda a relação dos dados com o todo.	Indiv. 68
0,889	5	A estatística está dividida em descritiva e inferência.	Indiv. 27
0,895	6	Pode fazer previsões com dados coletados para uma pesquisa.	Indiv. 32
0,895	7	É uma ciência que interpreta dados numéricos.	Indiv. 37
0,903	8	Recolhe amostra de uma população para tirar inferências.	Indiv. 10
0,904	9	Utilizada principalmente na realização de pesquisas.	Indiv. 23
0,929	10	É ciência das análises.	Indiv. 46
0,930	11	Fornecer resultados para pesquisas.	Indiv. 64
0,932	12	Usa dados para criar tabelas, gráficos e estimar.	Indiv. 19
0,933	13	Organiza e elabora resultados em diversas áreas.	Indiv. 65
0,935	14	Disciplina para estudar dados.	Indiv. 36
0,939	15	Tratamento de informações para prever acontecimentos.	Indiv. 11
0,946	16	Dá informações importantes para pesquisas.	Indiv. 80
0,953	17	Ciência que estuda pesquisas.	Indiv. 41
0,954	18	Uma disciplina importante.	Indiv. 74
0,962	19	Analisa dados de uma amostra.	Indiv. 47
0,969	20	Tratamento de dados.	Indiv. 26
1,161	21	Coleta e interpreta dados.	Indiv. 60

d_Positividade 4

Critère de classement	Index	Libellé de la réponse	Identifiant
0,827	1	Coleta dados e faz cálculos para diversos trabalhos em diversas áreas.	Indiv. 45
0,829	2	Coleta, organiza e calcula informações para trabalhos.	Indiv. 79
0,857	3	Faz o uso de informações para descrever ou prever algum fato.	Indiv. 16
0,859	4	Coleta dados e trabalha com essas informações para responder alguma pesquisa.	Indiv. 6
0,885	5	Interpreta dados de uma amostra para tirar conclusões de um determinado assunto.	Indiv. 12
0,905	6	Tudo que trata de coleta e interpretação de dados.	Indiv. 40
0,906	7	Estuda diversos dados de pesquisas que envolvem números.	Indiv. 71
0,918	8	Ciência que descreve e tira conclusões de amostras de uma população.	Indiv. 22
0,920	9	Cria tabelas, gráficos e calcula dados.	Indiv. 69
0,922	10	Usada nas pesquisas científicas ou não.	Indiv. 43
0,923	11	Coleta e trata as informações para uma pesquisa.	Indiv. 1
0,926	12	Avaliação e conclusão de amostras.	Indiv. 61
0,932	13	Usada para fins de pesquisa.	Indiv. 59
0,936	14	Interpreta e elabora informações.	Indiv. 76
0,952	15	Ciência que estuda informações para realizar inferências.	Indiv. 35
0,968	16	Coleta e interpreta dados numéricos.	Indiv. 58
0,969	17	Analisa dados.	Indiv. 72
0,994	18	Analisa dados.	Indiv. 81

ANEXO V - GRUPOS X POSITIVIDADE

Grupos / a_POSITIVIDADE

Nombre de cases d'effectif théorique < 5	16
Nombre de degrés de liberté	15
Khi-2	20,397
P-value	0,157
T de Tschuprow	0,255
V de Cramer	0,290

Effectif/poids		a_Positividade 1	a_Positividade 2	a_Positividade 3	a_Positividade 4	Ensemble
% en ligne	% en colonne					
1_Administração		8	14	7	3	32
		25,0	43,8	21,9	9,4	100,0
		36,4	53,8	50,0	15,8	39,5
2_Matemática		0	1	2	2	5
		0,0	20,0	40,0	40,0	100,0
		0,0	3,8	14,3	10,5	6,2
3_Ciências Contábeis		12	6	1	10	29
		41,4	20,7	3,4	34,5	100,0
		54,5	23,1	7,1	52,6	35,8
4_Ciências Econômicas		1	1	2	2	6
		16,7	16,7	33,3	33,3	100,0
		4,5	3,8	14,3	10,5	7,4
5_Engenharia Química		1	3	1	1	6
		16,7	50,0	16,7	16,7	100,0
		4,5	11,5	7,1	5,3	7,4
6_Engenharia Ambiental		0	1	1	1	3
		0,0	33,3	33,3	33,3	100,0
		0,0	3,8	7,1	5,3	3,7
Ensemble		22	26	14	19	81
		27,2	32,1	17,3	23,5	100,0
		100,0	100,0	100,0	100,0	100,0

Grupos / d_POSITIVIDADE

Nombre de cases d'effectif théorique < 5	16
Nombre de degrés de liberté	15
Khi-2	11,013
P-value	0,752
T de Tschuprow	0,187
V de Cramer	0,213

Effectif/poids		d_Positividade 1	d_Positividade 2	d_Positividade 3	d_Positividade 4	Ensemble
% en ligne	% en colonne					
1_Administração		9	8	10	5	32
		28,1	25,0	31,3	15,6	100,0
		40,9	40,0	47,6	27,8	39,5
2_Matemática		1	1	2	1	5
		20,0	20,0	40,0	20,0	100,0
		4,5	5,0	9,5	5,6	6,2
3_Ciências Contábeis		8	9	6	6	29
		27,6	31,0	20,7	20,7	100,0
		36,4	45,0	28,6	33,3	35,8
4_Ciências Econômicas		2	0	1	3	6
		33,3	0,0	16,7	50,0	100,0
		9,1	0,0	4,8	16,7	7,4
5_Engenharia Química		2	2	1	1	6
		33,3	33,3	16,7	16,7	100,0
		9,1	10,0	4,8	5,6	7,4
6_Engenharia Ambiental		0	0	1	2	3
		0,0	0,0	33,3	66,7	100,0
		0,0	0,0	4,8	11,1	3,7
Ensemble		22	20	21	18	81
		27,2	24,7	25,9	22,2	100,0
		100,0	100,0	100,0	100,0	100,0

ANEXO W - Q2 ANTES - TABELA BINÁRIA

	a_Positivid	a_Positivid	a_Positivid	a_Positivid	a_Positivid	aS1_1	aS1_2A	aS1_2B	aS1_2C	aS2_1	aS2_2	aS2_3A	aS2_3B	aS2_3C	aS3_1	aS3_2
E01_001	0	1	0	0	2	0	1	0	0	0	1	1	0	0	1	0
E01_002	0	1	0	0	2	1	0	0	1	0	0	1	0	1	0	0
E01_003	1	0	0	0	1	1	0	0	1	1	1	0	0	0	0	0
E01_004	0	1	0	0	2	1	1	0	0	1	1	0	0	1	0	0
E01_005	0	1	0	0	2	0	1	1	0	0	1	1	0	1	0	0
E01_006	0	0	1	0	3	0	0	0	1	0	0	0	0	0	0	0
E01_007	0	0	1	0	3	1	0	0	0	0	1	0	0	0	1	0
E01_008	0	1	0	0	2	1	0	0	0	0	1	1	1	1	1	0
E01_009	1	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0
E01_010	1	0	0	0	1	1	0	0	1	0	1	0	0	1	1	0
E01_011	0	1	0	0	2	1	1	1	0	1	1	1	0	0	1	0
E01_012	0	0	0	1	4	1	1	1	1	1	1	1	1	1	1	1
E01_013	0	1	0	0	2	1	0	0	0	0	1	1	0	1	1	0
E01_014	0	1	0	0	2	1	0	0	1	0	0	0	0	0	1	0
E01_015	0	1	0	0	2	1	0	0	0	1	0	1	0	0	1	0
E01_016	0	0	0	1	4	1	0	1	1	0	1	0	0	0	1	1
E01_017	0	1	0	0	2	0	0	0	0	0	1	1	1	0	0	0
E01_018	0	1	0	0	2	1	0	0	0	0	1	0	0	1	1	0
E01_019	0	0	1	0	3	1	1	0	0	1	1	1	0	1	1	1
E01_020	0	0	1	0	3	0	0	0	1	0	0	0	0	1	0	0
E01_021	0	0	1	0	3	1	0	0	0	0	1	0	0	1	1	0
E01_022	0	0	0	1	4	1	1	1	0	1	1	1	0	0	1	1
E01_023	0	0	1	0	3	1	1	0	0	0	1	1	1	0	1	1
E01_024	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0
E01_025	1	0	0	0	1	1	1	0	0	1	0	0	0	0	1	0
E01_026	1	0	0	0	1	1	0	0	1	0	0	1	1	1	1	0
E01_027	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
E01_028	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0
E01_029	0	0	1	0	3	1	1	0	0	1	0	0	0	1	0	0
E01_030	0	1	0	0	2	0	1	0	0	0	1	1	0	1	0	0
E01_031	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
E01_032	0	1	0	0	2	1	0	0	0	0	0	0	0	0	1	0
E02_001	0	0	1	0	3	1	0	0	0	0	1	0	0	1	1	0
E02_002	0	1	0	0	2	1	0	0	0	0	1	0	0	0	0	0
E02_003	0	0	1	0	3	1	0	0	1	0	1	0	0	1	1	0
E02_004	0	0	0	1	4	1	1	1	0	1	1	1	0	0	1	1
E02_005	0	0	0	1	4	0	0	0	0	0	1	0	0	0	1	1
E03_001	0	0	0	1	4	1	0	0	0	0	1	1	0	1	1	1
E03_002	0	1	0	0	2	1	0	0	1	0	0	0	0	0	1	0
E03_003	0	0	0	1	4	1	0	0	0	1	0	1	0	1	0	0
E03_004	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0
E03_005	0	0	0	1	4	1	1	1	1	1	1	1	1	1	1	1
E03_006	1	0	0	0	1	1	0	0	0	0	1	0	0	1	1	0
E03_007	0	0	0	1	4	1	1	0	0	1	1	1	0	1	1	1
E03_008	0	1	0	0	2	0	0	0	1	0	0	0	0	1	0	0
E03_009	0	0	0	1	4	1	0	0	0	0	1	0	0	1	1	1
E03_010	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
E03_011	0	0	1	0	3	1	0	0	0	1	0	0	0	0	0	0
E03_012	0	0	0	1	4	1	1	0	0	1	0	0	0	1	0	0
E03_013	0	1	0	0	2	0	1	0	0	0	1	1	0	1	0	0
E03_014	0	1	0	0	2	0	0	0	1	0	0	0	0	0	0	0
E03_015	0	0	0	1	4	1	0	0	0	0	0	0	0	0	1	0
E03_016	0	0	0	1	4	1	0	0	0	0	1	0	0	1	1	1
E03_017	1	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0
E03_018	0	0	0	1	4	1	1	1	1	1	1	1	1	1	1	1
E03_019	0	0	0	1	4	1	0	0	0	0	0	0	0	0	1	0
E03_020	1	0	0	0	1	1	0	0	0	0	1	0	0	1	1	0
E03_021	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
E03_022	1	0	0	0	1	0	1	0	0	0	1	1	0	0	1	0
E03_023	1	0	0	0	1	1	0	0	1	0	0	1	0	1	0	0
E03_024	1	0	0	0	1	1	0	0	1	1	1	0	0	0	0	0
E03_025	1	0	0	0	1	1	1	0	0	1	1	0	0	1	0	0
E03_026	0	1	0	0	2	0	1	1	0	0	1	1	0	1	0	0
E03_027	0	1	0	0	2	0	0	0	1	0	0	0	0	0	0	0
E03_028	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0
E03_029	1	0	0	0	1	1	0	0	0	0	1	1	1	1	1	0
E04_001	0	1	0	0	2	1	0	0	0	0	1	0	0	0	0	0
E04_002	1	0	0	0	1	1	0	0	1	0	1	0	0	1	1	0
E04_003	0	0	1	0	3	1	1	1	0	1	1	1	0	0	1	0
E04_004	0	0	0	1	4	1	1	1	1	1	1	1	1	1	1	1
E04_005	0	0	0	1	4	1	0	0	0	0	1	1	0	1	1	0
E04_006	0	0	1	0	3	1	0	0	1	0	0	0	0	0	1	0
E05_001	0	1	0	0	2	1	0	0	0	1	0	1	0	1	0	0
E05_002	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0
E05_003	0	1	0	0	2	1	0	0	0	0	1	0	0	0	0	0
E05_004	0	0	1	0	3	1	0	0	1	0	1	0	0	1	1	0
E05_005	0	0	0	1	4	1	1	1	0	1	1	1	0	0	1	0
E05_006	0	1	0	0	2	0	0	0	0	0	1	0	0	0	0	0
E06_001	0	1	0	0	2	1	0	0	0	0	1	1	0	1	1	0
E06_002	0	0	1	0	3	1	0	0	1	0	0	0	0	0	1	0
E06_003	0	0	0	1	4	1	0	0	0	1	0	1	0	1	1	1

ANEXO Y - DADOS ÁRVORE ANTES

nb col : 12, nb lig : 81

a_Positividade
W 2.67234
W 0
W 0

Partitions optimales :
a_Positividade1 de 1 à 1
a_Positividade2 de 2 à 2
a_Positividade3 de 3 à 3
a_Positividade4 de 4 à 4

nb col : 15, nb lig : 81

	Occurrence	Moyenne	Ecart type
a_Positividade1	: 22.00	0.27	0.44
a_Positividade2	: 26.00	0.32	0.47
a_Positividade3	: 14.00	0.17	0.38
a_Positividade4	: 19.00	0.23	0.42
aS1_1	: 63.00	0.78	0.42
aS1_2A	: 23.00	0.28	0.45
aS1_2B	: 12.00	0.15	0.36
aS1_2C	: 24.00	0.30	0.46
aS2_1	: 24.00	0.30	0.46
aS2_2	: 54.00	0.67	0.47
aS2_3A	: 32.00	0.40	0.49
aS2_3B	: 9.00	0.11	0.31
aS2_3C	: 40.00	0.49	0.50
aS3_1	: 47.00	0.58	0.49
aS3_2	: 15.00	0.19	0.39
a_Positividade1 s	: 22.00	0.27	0.44
a_Positividade2 s	: 26.00	0.32	0.47
a_Positividade3 s	: 14.00	0.17	0.38
a_Positividade4 s	: 19.00	0.23	0.42

Fréquence des couples de variables :

	a_Positividade1	a_Positividade2	a_Positividade3	a_Positividade4	aS1_1	aS1_2A	aS1_2B	aS1_2C	aS2_1	aS2_2	aS2_3A	aS2_3B	aS2_3C	aS3_1	aS3_2
a_Positividade1	0.00	0.00	0.00	0.22	0.04	0.00	0.09	0.06	0.20	0.05	0.02	0.10	0.14	0.00	
a_Positividade2		0.00	0.00	0.19	0.09	0.04	0.07	0.05	0.20	0.16	0.02	0.16	0.11	0.00	
a_Positividade3			0.00	0.15	0.05	0.01	0.07	0.05	0.10	0.04	0.01	0.09	0.12	0.02	
a_Positividade4				0.22	0.11	0.10	0.06	0.14	0.17	0.15	0.05	0.15	0.21	0.16	
aS1_1				0.21	0.12	0.22	0.30	0.53	0.31	0.10	0.42	0.54	0.17		
aS1_2A					0.14	0.05	0.20	0.25	0.22	0.06	0.17	0.19	0.11		
aS1_2B						0.06	0.11	0.15	0.14	0.05	0.07	0.12	0.09		
aS1_2C							0.07	0.14	0.09	0.06	0.16	0.17	0.06		
aS2_1								0.19	0.19	0.05	0.17	0.16	0.11		
aS2_2									0.31	0.10	0.36	0.46	0.17		
aS2_3A										0.11	0.28	0.27	0.14		
aS2_3B											0.09	0.10	0.06		
aS2_3C												0.31	0.12		
aS3_1														0.19	
aS3_2															0.19

Coefficient de corrélation :

	a_Positividade1	a_Positividade2	a_Positividade3	a_Positividade4	aS1_1	aS1_2A	aS1_2B	aS1_2C	aS2_1	aS2_2	aS2_3A	aS2_3B	aS2_3C	aS3_1	aS3_2
a_Positividade1	-0.42	-0.28	-0.34	0.06	-0.20	-0.25	0.03	-0.09	0.08	-0.27	-0.04	-0.16	-0.10	-0.29	
a_Positividade2		-0.31	-0.38	-0.33	-0.02	-0.06	-0.10	-0.21	-0.07	0.15	-0.07	0.01	-0.33	-0.33	
a_Positividade3			-0.25	0.09	0.00	-0.10	0.13	-0.01	-0.09	-0.17	-0.06	0.01	0.12	-0.05	
a_Positividade4				0.23	0.23	0.43	-0.04	0.34	0.08	0.27	0.18	0.15	0.35	0.71	
aS1_1					-0.06	0.06	-0.04	0.35	0.06	0.01	0.09	0.17	0.45	0.18	
aS1_2A						0.59	-0.17	0.55	0.27	0.50	0.21	0.14	0.09	0.33	
aS1_2B							0.11	0.41	0.29	0.44	0.29	0.01	0.21	0.43	
aS1_2C								-0.07	-0.29	-0.14	0.20	0.06	0.00	0.04	
aS2_1									-0.06	0.31	0.11	0.12	-0.05	0.32	
aS2_2										0.20	0.17	0.12	0.30	0.27	
aS2_3A											0.44	0.36	0.18	0.33	
aS2_3B												0.20	0.22	0.34	
aS2_3C													0.09	0.16	
aS3_1															0.41
aS3_2															

Indices de similarité :

	a_Positividade1	a_Positividade2	a_Positividade3	a_Positividade4	aS1_1	aS1_2A	aS1_2B	aS1_2C	aS2_1	aS2_2	aS2_3A	aS2_3B	aS2_3C	aS3_1	aS3_2
a_Positividade1	1.00	0.00	0.03	0.01	0.59	0.10	0.04	0.57	0.28	0.64	0.06	0.39	0.19	0.31	0.02
a_Positividade2		1.00	0.02	0.01	0.12	0.44	0.33	0.27	0.09	0.37	0.80	0.30	0.52	0.06	0.01
a_Positividade3			1.00	0.03	0.63	0.50	0.23	0.82	0.47	0.33	0.14	0.33	0.51	0.74	0.36
a_Positividade4				1.00	0.80	0.94	1.00	0.40	0.99	0.65	0.95	0.90	0.80	0.96	1.00
aS1_1					1.00	0.42	0.59	0.44	0.89	0.56	0.51	0.65	0.70	0.89	0.75
aS1_2A						1.00	0.42	1.00	0.14	1.00	0.88	1.00	0.94	0.78	0.67
aS1_2B							1.00	1.00	0.78	1.00	0.92	1.00	0.99	0.51	0.88
aS1_2C								1.00	1.00	0.34	0.11	0.21	0.92	0.63	0.51
aS2_1									1.00	0.34	1.00	0.40	0.96	0.79	0.40
aS2_2										1.00	0.79	0.79	0.67	0.84	0.90
aS2_3A											1.00	1.00	0.96	0.79	0.98
aS2_3B												1.00	0.89	0.89	1.00
aS2_3C													1.00	0.64	0.83
aS3_1														1.00	0.98
aS3_2															1.00

Classification au niveau : 1 : (a_Positividade4 aS3_2) similarité : 1

Classification au niveau : 2 : (aS1_2A aS1_2B) similarité : 0.99998

Classification au niveau : 3 : ((aS1_2A aS1_2B) aS2_1) similarité : 0.999566

Classification au niveau : 4 : (a52_3A a52_3B) similarité : 0.998057

Classification au niveau : 5 : ((a_Positividade4 a53_2) ((a51_2A a51_2B) a52_1)) similarité : 0.995955

Classification au niveau : 6 : (((a_Positividade4 a53_2) ((a51_2A a51_2B) a52_1)) (a52_3A a52_3B)) similarité : 0.984576

Classification au niveau : 7 : (a51_1 a53_1) similarité : 0.89089

Classification au niveau : 8 : (a_Positividade3 a51_2C) similarité : 0.818389

Classification au niveau : 9 : (((a_Positividade4 a53_2) ((a51_2A a51_2B) a52_1)) (a52_3A a52_3B)) (a51_1 a53_1) similarité : 0.793187

Classification au niveau : 10 : (((((a_Positividade4 a53_2) ((a51_2A a51_2B) a52_1)) (a52_3A a52_3B)) (a51_1 a53_1)) a52_3C) similarité : 0.724992

Classification au niveau : 11 : (a_Positividade1 a52_2) similarité : 0.636138

Classification au niveau : 12 : (((a_Positividade3 a51_2C) (((a_Positividade4 a53_2) ((a51_2A a51_2B) a52_1)) (a52_3A a52_3B)) (a51_1 a53_1)) a52_3C) similarité : 0.203489

Classification au niveau : 13 : ((a_Positividade1 a52_2) a_Positividade2) similarité : 0.140166

Classification au niveau : 14 : (((a_Positividade1 a52_2) a_Positividade2) ((a_Positividade3 a51_2C) (((a_Positividade4 a53_2) ((a51_2A a51_2B) a52_1)) (a52_3A a52_3B)) (a51_1 a53_1)) a52_3C)) similarité : 0.0523954

Le noeud le plus significatif est au niveau : 1

Noeuds significatifs

au niveau : 1
 au niveau : 3
 au niveau : 5
 au niveau : 10
 au niveau : 13

Contribution à la classe : a_Positividade4 a53_2 (1)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.979
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.989
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.914
 La variable a_Positividade4 contribue à cette classe avec un risque de : 3.89e-08

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 3.89e-08

Contribution à la classe : a51_2A,a51_2B (2)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.96
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.478
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.585
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.00215

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.00215

Contribution à la classe : a51_2A,a51_2B,a52_1 (2,3)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.925
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.801
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.471
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.000564

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.000564

Contribution à la classe : a52_3A,a52_3B (4)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.449
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.564
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.471
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.052

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.052

Contribution à la classe : a_Positividade4,a53_2,a51_2A,a51_2B,a52_1 (1,2,3,5)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.816
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.865
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.66
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.000277

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.000277

Contribution à la classe : a_Positividade4,a53_2,a51_2A,a51_2B,a52_1,a52_3A,a52_3B (1,2,3,4,5,6)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.816
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.865
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.66
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.000277

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.000277

Contribution à la classe : a51_1,a53_1 (7)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.733
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.987
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.0572
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.00127

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.00127

Contribution à la classe : a_Positividade3,a51_2C (8)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.816
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.865
 La variable a_Positividade3 contribue à cette classe avec un risque de : 2.63e-05
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.768

La variable qui contribue le plus à cette classe est a_Positividade3 avec un risque de : 2.63e-05

Contribution à la classe : a_Positividade4,a53_2,a51_2A,a51_2B,a52_1,a52_3A,a52_3B,a51_1,a53_1,1,2,3,4,5,6,7,9,1

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.816
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.865
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.66
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.000277

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.000277

Contribution à la classe : a_Positividade4,a53_2,a51_2A,a51_2B,a52_1,a52_3A,a52_3B,a51_1,a53_1,a52_3C (1,2,3,4,5,6,7,9,10)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.816
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.865
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.66
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.000277

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.000277

Contribution à la classe : a_Positividade1,a52_2 (11)

La variable a_Positividade1 contribue à cette classe avec un risque de : 9.97e-09
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.997
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.954
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.985

La variable qui contribue le plus à cette classe est a_Positividade1 avec un risque de : 9.97e-09

Contribution à la classe : a_Positividade3,a51_2C,a_Positividade4,a53_2,a51_2A,a51_2B,a52_1,a52_3A,a52_3B,a51_1,a53_1,a52_3C (1,2,3,4,5,6,7,8,9,10,12)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.863
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.905
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.718
 La variable a_Positividade4 contribue à cette classe avec un risque de : 9.81e-05

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 9.81e-05

Contribution à la classe : a_Positividade1,a52_2,a_Positividade2 (11,13)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.000359
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.0068
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.999
 La variable a_Positividade4 contribue à cette classe avec un risque de : 1

La variable qui contribue le plus à cette classe est a_Positividade1 avec un risque de : 0.000359

Contribution à la classe : a_Positividade1,a52_2,a_Positividade2,a_Positividade3,a51_2C,a_Positividade4,a53_2,a51_2A,a51_2B,a52_1,a52_3A,a52_3B,a51_1,a53_1,a52_3C (1,2,3,4,5,6,7,8,9,10,11,12,13,14)

La variable a_Positividade1 contribue à cette classe avec un risque de : 0.816
 La variable a_Positividade2 contribue à cette classe avec un risque de : 0.865
 La variable a_Positividade3 contribue à cette classe avec un risque de : 0.66
 La variable a_Positividade4 contribue à cette classe avec un risque de : 0.000277

La variable qui contribue le plus à cette classe est a_Positividade4 avec un risque de : 0.000277

ANEXO Z - DADOS ÁRVORE DEPOIS

nb.col = 14 nb.lie = 81

d Positividade

W0 77747

W0

W0

Partições estimadas :

d Positividade1 de 1 à 1

d Positividade2 de 2 à 2

d Positividade3 de 3 à 3

d Positividade4 de 4 à 4

nb.col = 17 nb.lie = 81

Corumana Miramaa Frontana

d Positividade1 : 22.00 0.77 0.44

d Positividade2 : 20.00 0.75 0.43

d Positividade3 : 21.00 0.76 0.44

d Positividade4 : 18.00 0.72 0.42

d51 1 : 75.00 0.93 0.26

d51 2A : 61.00 0.75 0.43

d51 2B : 41.00 0.51 0.50

d51 2C : 45.00 0.56 0.50

d51 2D : 73.00 0.28 0.45

d52 1 : 71.00 0.88 0.33

d52 2 : 75.00 0.93 0.26

d52 2A : 65.00 0.80 0.40

d52 2B : 48.00 0.59 0.49

d52 2C : 63.00 0.65 0.48

d52 2D : 77.00 0.33 0.47

d53 1 : 76.00 0.94 0.24

d53 2 : 55.00 0.68 0.47

d Positividade1 s : 77.00 0.77 0.44

d Positividade2 s : 20.00 0.75 0.43

d Positividade3 s : 71.00 0.76 0.44

d Positividade4 s : 18.00 0.72 0.42

Frequência das contagens de variáveis :

	d Positividade1	d Positividade2	d Positividade3	d Positividade4	d51 1	d51 2A	d51 2B	d51 2C	d51 2D	d52 1	d52 2	d52 2A	d52 2B	d52 2C	d52 2D	d53 1	d53 2	
d Positividade1	0.00	0.00	0.00	0.27	0.27	0.15	0.12	0.01	0.26	0.25	0.23	0.20	0.23	0.06	0.23	0.19		
d Positividade2		0.00	0.00	0.20	0.21	0.14	0.14	0.01	0.20	0.25	0.19	0.15	0.16	0.04	0.23	0.17		
d Positividade3			0.00	0.25	0.19	0.10	0.19	0.09	0.23	0.25	0.20	0.14	0.16	0.09	0.25	0.16		
d Positividade4				0.21	0.14	0.17	0.11	0.17	0.19	0.19	0.19	0.11	0.10	0.15	0.22	0.16		
d51 1				0.68	0.43	0.56	0.27	0.81	0.85	0.73	0.52	0.64	0.31	0.86	0.60			
d51 2A					0.51	0.17	0.65	0.73	0.62	0.53	0.53	0.73	0.74	0.53				
d51 2B					0.21	0.14	0.44	0.47	0.51	0.41	0.27	0.17	0.49	0.38				
d51 2C						0.19	0.47	0.53	0.49	0.38	0.35	0.72	0.49	0.40				
d51 2D							0.23	0.26	0.22	0.16	0.16	0.25	0.27	0.22				
d52 1								0.80	0.74	0.54	0.58	0.77	0.81	0.54				
d52 2									0.73	0.56	0.62	0.30	0.86	0.68				
d52 2A										0.69	0.69	0.77	0.77	0.58				
d52 2B											0.40	0.20	0.57	0.47				
d52 2C												0.71	0.67	0.43				
d52 2D													0.33	0.25				
d53 1															0.64			
d53 2																		

Coefficient de corrélation :

	d Positividade1	d Positividade2	d Positividade3	d Positividade4	d51 1	d51 2A	d51 2B	d51 2C	d51 2D	d52 1	d52 2	d52 2A	d52 2B	d52 2C	d52 2D	d53 1	d53 2	
d Positividade1	-0.35	-0.36	-0.33	0.17	0.09	0.05	-0.12	-0.32	0.14	-0.04	0.09	0.17	0.27	-0.14	-0.19	0.00		
d Positividade2		-0.34	-0.21	-0.28	0.13	0.05	-0.01	-0.30	-0.13	0.16	-0.08	0.01	-0.01	-0.22	0.03	0.03		
d Positividade3			-0.32	0.06	-0.05	-0.15	0.19	0.05	0.05	-0.06	-0.08	-0.04	0.00	0.03	-0.08			
d Positividade4				0.04	-0.18	0.05	-0.06	0.50	-0.07	-0.19	0.04	-0.10	-0.24	0.38	0.14	0.05		
d51 1				-0.16	-0.28	0.32	0.07	0.04	-0.08	-0.14	-0.23	0.29	0.00	-0.07	-0.19			
d51 2A					0.58	-0.24	-0.21	-0.04	0.17	0.44	0.34	0.13	-0.08	0.33	0.10			
d51 2B						-0.29	-0.04	0.00	0.00	0.50	0.44	-0.25	0.02	0.16	0.17			
d51 2C							0.12	-0.11	0.03	-0.07	-0.19	-0.08	0.16	-0.23	0.08			
d51 2D								-0.10	-0.03	-0.03	-0.04	-0.12	0.72	0.05	0.14			
d52 1									-0.11	0.29	0.22	0.04	-0.13	-0.10	-0.26			
d52 2										-0.14	0.05	0.09	-0.10	-0.07	0.41			
d52 2A											0.60	-0.17	0.07	0.13	0.19			
d52 2B												0.03	0.00	0.10	0.29			
d52 2C													-0.04	0.03	-0.05			
d52 2D														0.18	0.09			
d53 1															0.04			
d53 2																		

Indices de similarité :

	d Positividade1	d Positividade2	d Positividade3	d Positividade4	d51 1	d51 2A	d51 2B	d51 2C	d51 2D	d52 1	d52 2	d52 2A	d52 2B	d52 2C	d52 2D	d53 1	d53 2		
d Positividade1	1.00	0.01	0.01	0.01	0.64	0.64	0.60	0.26	0.02	0.65	0.47	0.63	0.79	0.89	0.19	0.36	0.51		
d Positividade2		1.00	0.01	0.02	0.28	0.69	0.61	0.49	0.02	0.36	0.63	0.40	0.52	0.49	0.08	0.52	0.55		
d Positividade3			1.00	0.02	0.55	0.42	0.21	0.84	0.66	0.55	0.55	0.42	0.34	0.42	0.50	0.53	0.37		
d Positividade4				1.00	0.43	0.24	0.62	0.28	0.42	1.00	0.42	0.34	0.46	0.30	0.14	0.99	0.61	0.50	
d51 1					0.64	0.28	0.55	0.53	1.00	0.42	0.32	0.20	0.56	0.51	0.48	0.44	0.36	0.66	0.39
d51 2A						0.64	0.49	0.63	0.74	0.42	0.32	0.16	0.71	0.47	0.58	0.81	0.82	0.63	0.28
d51 2B							0.60	0.61	0.71	0.62	0.32	0.11	0.43	0.50	0.50	0.92	0.96	0.18	0.54
d51 2C								0.76	0.41	0.73	0.41	0.52	0.43	0.34	0.40	0.78	0.37	0.60	0.73
d51 2D									0.02	0.62	0.43	0.43	0.40	0.47	0.46	0.43	0.30	1.00	0.54
d52 1										0.65	0.36	0.55	0.47	0.65	0.67	0.53	0.27	0.47	0.32
d52 2											0.47	0.63	0.55	0.34	0.48	0.58	0.50	0.52	0.47
d52 2A												0.46	0.66	0.44	1.00	0.64	0.25	0.53	0.55
d52 2B													0.63	0.54	0.54	1.00	0.54	0.50	0.56
d52 2C														0.56	0.54	1.00	0.44	0.52	0.43
d52 2D															0.19	0.08	1.00	0.99	0.63
d53 1																0.36	0.52	1.00	0.52
d53 2																	0.55	0.37	1.00

Classification au niveau : 1 - (d51 2D d52 2D) similarité : 0.999996

Classification au niveau : 2 - (d Positividade4 d51 2D d52 2D) similarité : 0.999916

Classification au niveau : 3 - (d51 2A d51 2B) similarité : 0.985762

Classification au niveau : 4 - (d52 2A d52 2B) similarité : 0.936708

Classification au niveau : 5 : I_d Positivité1 d52 3C) similarité : 0.887577

Classification au niveau : 6 : $I_d(I_{d51} 2A d51 2B) d52 3A d52 3B)$ similarité : 0.853895

Classification au niveau : 7 : I_d Positivité3 d51 2C) similarité : 0.835483

Classification au niveau : 8 : $I_d(I_{d52} 2 d53 2)$ similarité : 0.715965

Classification au niveau : 9 : I_d Positivité2 d53 1) similarité : 0.521592

Classification au niveau : 10 : $I_d(I_{d51} 1 d52 1)$ similarité : 0.512754

Classification au niveau : 11 : I_d Positivité3 d51 2C) $I_d(I_{d51} 1 d52 1)$ similarité : 0.258297

Classification au niveau : 12 : $I_d(I_{d51} 2A d51 2B) d52 3A d52 3B)$ $I_d(I_{d52} 2 d53 2)$ similarité : 0.221404

Classification au niveau : 13 : I_d Positivité1 d52 3C) I_d Positivité2 d53 1) similarité : 0.0705421

Classification au niveau : 14 : I_d Positivité3 d51 2C) $I_d(I_{d51} 1 d52 1)$ I_d Positivité4 d51 2D d52 3D)) similarité : 0.0512724

Classification au niveau : 15 : $I_d(I_{d52} 2 d53 2)$ I_d Positivité2 d53 1) $I_d(I_{d51} 2A d51 2B) d52 3A d52 3B)$ $I_d(I_{d52} 2 d53 2)$ $I_d(I_{d51} 2A d51 2B) d52 3A d52 3B)$ similarité : 0.00994963

Classification au niveau : 16 : $I_d(I_{d51} 2A d51 2B) d52 3A d52 3B)$ $I_d(I_{d51} 2A d51 2B) d52 3A d52 3B)$ similarité : 1.92629e-10

Le nœud le plus significatif est au niveau : 1

Nœuds significatifs

au niveau : 1
 au niveau : 6
 au niveau : 12
 au niveau : 15

Contribution à la classe : d51 2D d52 3D) 1)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 0.984
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 0.978
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 0.285
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 2.97e-05
 la variable qui contribue le plus à cette classe est I_d Positivité4 avec un risque de : 2.97e-05

Contribution à la classe : d Positivité4 d51 2D d52 3D) 1 2)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 0.971
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 0.96
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 0.962
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 6.78e-08
 la variable qui contribue le plus à cette classe est I_d Positivité4 avec un risque de : 6.78e-08

Contribution à la classe : d51 2A d51 2B) 3)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 0.281
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 0.77
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 0.824
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 0.257
 la variable qui contribue le plus à cette classe est I_d Positivité3 avec un risque de : 0.824

Contribution à la classe : d52 3A d52 3B) 4)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 0.951
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 0.39
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 0.656
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 0.715
 la variable qui contribue le plus à cette classe est I_d Positivité1 avec un risque de : 0.951

Contribution à la classe : d Positivité1 d52 3C) 5)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 2.54e-11
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 0.995
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 0.995
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 0.997
 la variable qui contribue le plus à cette classe est I_d Positivité1 avec un risque de : 2.54e-11

Contribution à la classe : d51 2A d51 2B) d52 3A d52 3B) 3 4 6)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 0.422
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 0.142
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 0.426
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 0.452
 la variable qui contribue le plus à cette classe est I_d Positivité2 avec un risque de : 0.142

Contribution à la classe : d Positivité3 d51 2C) 7)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 0.988
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 0.982
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 1.5e-08
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 0.975
 la variable qui contribue le plus à cette classe est I_d Positivité3 avec un risque de : 1.5e-08

Contribution à la classe : d52 2 d53 2) 8)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 0.426
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 0.336
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 0.647
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 0.265
 la variable qui contribue le plus à cette classe est I_d Positivité3 avec un risque de : 0.647

Contribution à la classe : d Positivité2 d53 1) 9)

la variable I_d Positivité1 contribue à cette classe avec un risque de : 0.997
 la variable I_d Positivité2 contribue à cette classe avec un risque de : 2.54e-12
 la variable I_d Positivité3 contribue à cette classe avec un risque de : 0.995
 la variable I_d Positivité4 contribue à cette classe avec un risque de : 0.997
 la variable qui contribue le plus à cette classe est I_d Positivité2 avec un risque de : 2.54e-12

Contributions à la classe : dC1 1 dC2 1 (101)

la variable d1 Contributable1 contribue à cette classe avec un risque de : 0.011
 la variable d1 Contributable2 contribue à cette classe avec un risque de : 0.030
 la variable d1 Contributable3 contribue à cette classe avec un risque de : 0.730
 la variable d1 Contributable4 contribue à cette classe avec un risque de : 0.566

la variable qui contribue le plus à cette classe est d1 Contributable1 avec un risque de : 0.011

Contributions à la classe : d1 Contributable2 dC1 20 dC1 1 dC2 1 (7 10 11)

la variable d1 Contributable1 contribue à cette classe avec un risque de : 0.976
 la variable d1 Contributable2 contribue à cette classe avec un risque de : 0.07
 la variable d1 Contributable3 contribue à cette classe avec un risque de : 7.53e-07
 la variable d1 Contributable4 contribue à cette classe avec un risque de : 0.067

la variable qui contribue le plus à cette classe est d1 Contributable1 avec un risque de : 7.87e-07

Contributions à la classe : dC1 2d dC1 2R dC2 2d dC2 2R dC3 2 dC3 2 (2 4 6 8 17)

la variable d1 Contributable1 contribue à cette classe avec un risque de : 0.797
 la variable d1 Contributable2 contribue à cette classe avec un risque de : 0.0846
 la variable d1 Contributable3 contribue à cette classe avec un risque de : 0.286
 la variable d1 Contributable4 contribue à cette classe avec un risque de : 0.476

la variable qui contribue le plus à cette classe est d1 Contributable2 avec un risque de : 0.0846

Contributions à la classe : d1 Contributable1 dC2 20 d1 Contributable2 dC2 1 (6 9 13)

la variable d1 Contributable1 contribue à cette classe avec un risque de : 1.31e-06
 la variable d1 Contributable2 contribue à cette classe avec un risque de : 0.00786
 la variable d1 Contributable3 contribue à cette classe avec un risque de : 1
 la variable d1 Contributable4 contribue à cette classe avec un risque de : 1

la variable qui contribue le plus à cette classe est d1 Contributable2 avec un risque de : 1.31e-06

Contributions à la classe : d1 Contributable1 dC1 20 dC1 1 dC2 1 d1 Contributable2 dC1 20 dC2 20 (1 2 7 10 11 14)

la variable d1 Contributable1 contribue à cette classe avec un risque de : 0.034
 la variable d1 Contributable2 contribue à cette classe avec un risque de : 0.908
 la variable d1 Contributable3 contribue à cette classe avec un risque de : 0.0777
 la variable d1 Contributable4 contribue à cette classe avec un risque de : 0.0109

la variable qui contribue le plus à cette classe est d1 Contributable2 avec un risque de : 0.0109

Contributions à la classe : d1 Contributable1 dC2 20 d1 Contributable2 dC2 1 dC1 2d dC1 2R dC2 2d dC2 2R dC3 2 dC3 2 (2 4 6 8 9 17 19 16)

la variable d1 Contributable1 contribue à cette classe avec un risque de : 0.00961
 la variable d1 Contributable2 contribue à cette classe avec un risque de : 0.184
 la variable d1 Contributable3 contribue à cette classe avec un risque de : 0.014
 la variable d1 Contributable4 contribue à cette classe avec un risque de : 0.909

la variable qui contribue le plus à cette classe est d1 Contributable1 avec un risque de : 0.00961

Contributions à la classe : d1 Contributable1 dC2 20 d1 Contributable2 dC2 1 dC1 2d dC1 2R dC2 2d dC2 2R dC3 2 d1 Contributable3 dC1 20 dC1 1 dC2 1 d1 Contributable4 dC1 20 dC2 20 (1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16)

la variable d1 Contributable1 contribue à cette classe avec un risque de : 0.77
 la variable d1 Contributable2 contribue à cette classe avec un risque de : 0.803
 la variable d1 Contributable3 contribue à cette classe avec un risque de : 0.286
 la variable d1 Contributable4 contribue à cette classe avec un risque de : 0.106

la variable qui contribue le plus à cette classe est d1 Contributable2 avec un risque de : 0.106