



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

Uma Proposta de Ordenação de Periódicos
Científicos via Teoria da Resposta ao Item

por

Suélvio Alves de Moura

sob orientação da

Prof^a. Dr^a. Cibele Queiroz da Silva

Brasília, abril de 2017

Uma Proposta de Ordenação de Periódicos Científicos via Teoria da Resposta ao Item

Esta é a versão original da dissertação elaborada pelo candidato Suélio Alves de Moura, tal como submetida à Comissão Julgadora.

Uma Proposta de Ordenação de Periódicos Científicos via Teoria da Resposta ao Item

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Comissão Julgadora:

- Prof^a. Dr^a. Cibele Queiroz da Silva (orientadora) - EST/UnB
- Prof. Dr. Dalton Francisco de Andrade (membro externo) - INE/UFSC
- Prof. Dr. Raul Yukihiro Matsushita (membro interno) - EST/UnB
- Prof. Dr. Antonio Eduardo Gomes (suplente) - EST/UnB

Dedicatória

À minha família, em especial aos meus pais
José Moura e Maria de Fátima Moura.

Mensagem

“[...] Ando devagar porque já tive pressa
Levo esse sorriso porque já chorei demais
Hoje me sinto mais forte, mais feliz, quem sabe?
Só levo a certeza de que muito pouco eu sei
Ou nada sei.” (Almir Sater)

Para aqueles que acreditam nos sonhos

“... You may say I’m a dreamer
but I’m not the only one...”
(John Lennon)

“[...] O Sol de cada manhã... Há de brilhar
mesmo nos dias mais nublados.”
(Júlio César Gomes Fonseca)

Agradecimentos

Primeiramente, agradeço a DEUS, Pai Celestial e Criador, pelo dom da vida e por todas as bênçãos que sempre tem derramado sobre mim. Confio nos seus planos, pois sei que tem reservado o melhor para aqueles que acreditam e esperam pelo seu tempo de agir. Sentir a proteção materna da Virgem Maria e a luz suave do Espírito Santo fez com que o meu caminhar, em todos os momentos, se tornasse mais leve.

“Tudo posso naquele que me fortalece!” - Fl. 4; 13.

Família é sinônimo de amor, proteção, porto seguro. Assim defino a minha família. Sou muito feliz e realizado por ter os melhores pais do mundo, minha mainha Maria de Fátima Moura e meu painho José Moura. Duas pessoas simples, humildes, guerreiras e de coração gigante, que sempre batalharam para dar o melhor aos seus seis filhos: eu, Silvio Batista, Silvania Moura, Ciélio Moura, Sileide Moura e Samuel Moura.

A vida nunca foi fácil, mas devido aos ensinamentos dos meus pais, da minha Família, consegui a cada novo dia tornar-me um ser mais forte e acreditar no meu potencial sempre com a certeza de que, sempre que eu precisasse de um refúgio, teria o melhor porto seguro para ancorar o meu barco.

Em 2015, o Senhor Jesus mais uma vez começara a derramar suas bênçãos sobre mim. Neste ano, Ele me concebeu a graça de ser aprovado no Mestrado em Estatística na Universidade de Brasília (UnB). A priori, isso foi um susto e um grande desafio a ser enfrentado, já que eu teria que deixar toda uma vida em Pombal PB, ou melhor, no Sítio Gameleira (zona rural), e vir morar em uma cidade grande, Brasília DF. Cidade esta na qual não tinha família nem tinha amigos. Tudo seria novo para mim, mas, acreditando nas forças do Senhor e que Ele só nos dá o melhor, criei coragem e fiz de Brasília a minha nova cidade.

“Mesmo que já tenha feito uma longa caminhada, sempre haverá mais um caminho a percorrer.” - (Santo Agostinho).

Ao chegar em Brasília, diversas vezes pensei em desistir, isso no primeiro semestre do Mestrado, e voltar para casa, mas ao falar, por meio de uma ligação telefônica, com mainha, painho ou alguém da minha família e quando fazia minhas orações, sentia as forças serem revigoradas e fortalecidas. Não desisti perante as dificuldades ou por qualquer outra coisa, pois sabia que eu estava batalhando para dar uma vida melhor, uma vida mais digna aos meus pais, aos meus irmãos e a toda minha Família.

Não existe nada melhor do que sentir a presença viva desse DEUS que é amor, bondade, proteção e ressurreição.

Durante toda minha pós graduação morei em uma Residência Universitária, local onde vivi grandes experiências e aprendizagens. Para sempre, serei grato a todos que formam a Diretoria de Desenvolvimento Social (DDS) e Decanato de Assuntos Comunitário (DAC); à decana Denise Bom Tempo (gestora em 2015); à diretora do Restaurante Universitário (RU): Ygraine Hartmann Leibovich, à Jesus, Márcia e a Antônia pela disponibilidade em me proporcionarem todas as informações necessárias sobre documentos e trâmites burocráticos para o acesso aos programas de alimentação e assistência estudantil; ao diretor da Moradia Estudantil da Pós Graduação: Senhor Rubéns e todos os funcionários que trabalham na organização, manutenção e limpeza do prédio.

Na Colina, local onde morei durante todo o meu mestrado, conheci pessoas muito especiais, das quais irei levar para sempre comigo. Não tem como deixar de citá-las: minha amiguinha Regina (seu esposo Ailton e sua Família), minha companheira no canto das Santas Missas na Paróquia do Verbo Divino e parceira das mais lindas risadas quando estávamos juntos (“[...] Nossa são tantas as conversas, vamos deixar o que falta para um próximo capítulo, minha amiguinha”) e ao meu colega de quarto haitiano Vogly Nahum Pongnon por tanta compreensão que ele sempre teve comigo (“[...] Vogly estou tão cansado, eu posso apagar a lâmpada do quarto hoje mais cedo?! Pode sim Suêlo?!”).

Na minha vida, encontrei verdadeiros anjos de luz a quem chamo de amigos/irmãos:

“Amigo fiel é proteção poderosa, e quem o encontrar, terá encontrado um tesouro. Amigo fiel não tem preço e o seu valor é incalculável.” - (Ecl. 6,14-15).

Hoje, eu tenho certeza que o Senhor me presenteou com os melhores amigos/irmãos do mundo: Júlio César (JCGF & SAM), Renato Gama, Euriberto Furtado, Aguinaldo Leite e Deivison Urtiga. Irei levar comigo para sempre o jeitinho, a doçura, os aprendizados, os ensinamentos, momentos vividos juntos e a importância de cada um de vocês em minha vida. Sei que o destino se encarregará de nos levar a caminhos diferentes, mas tenho a certeza maior de que nunca iremos esquecer uns dos outros. Creio que, algum dia, o Senhor providenciará nossos reencontros e faremos uma grande festa. Amo muito vocês!

Para sempre serei grato ao Meu Irmãozinho Júlio César, foram tantas e tantas longas noites e finais de semana de muitos e muitos estudos. A cada novo encontro para nossos estudos era sempre uma festa, pois fazíamos tudo com muita empolgação e encarávamos cada desafio com muita aprendizagem e o desejo de superar cada uma das dificuldades, que encontrávamos. Muitas e muitas vezes quando o cansaço batia ou um dos dois pensava em ficar triste ou desanimado, entrava em cena o aconchego de Irmão um para com o outro. Isso tudo foi um suporte do qual recebi muitas forças para seguir a caminhada e não desanimar. O Júlio além de ter me ajudado muito na parte computacional da nossa pesquisa, foi a minha base de Família, Irmão e Amigo aqui em Brasília. Agradeço imensamente a toda sua família pelo carinho e receptividade para comigo; foi deles que recebi o primeiro aconchego

de família aqui em Brasília. E, foi através do Meu Irmãozinho que ganhei uma nova Família, a Família Gomes Fonseca (meu Painho Teteu e minha Mainha Maria, Belo Horizonte - MG).

Em seguida, o Pai Criador me presenteou com mais uma família que, para sempre, irei levar em meu coração; saibam que têm a minha eterna admiração, respeito e carinho. Obrigado por tudo, Padrinho Orlando, Madrinha Cândida e Valter; foi em cada uma dessas famílias que Cristo me fez e faz com que sinta a bênção que é ter um lar, um aconchego.

Agradeço a todos os professores do Programa de Pós Graduação em Estatística (PPGEST), em especial aos professores dos quais tive o prazer de ser seu aluno: Cibele Queiroz da Silva (Modelos Lineares Generalizados), Raul Yukihito Matsushita (Séries Temporais), Cira Etheowalda Guevara Otiniano (Estatística Matemática), Bernardo Borba de Andrade (Técnicas Estatística Computacional), Antônio Eduardo Gomes (Inferência Estatística), Eduardo Yoshio Nakano (Inferência Bayesiana), Gustavo Leonel Gilardoni (Metodologia da Pesquisa Científica). No entanto, meu carinho especial está voltado à Professora Cibele, minha Mãe Científica, pela orientação, paciência e cuidados que sempre teve comigo. À Professora Cibele meu eterno carinho, respeito, admiração e meu muito obrigado por sempre me direcionar aos melhores caminhos e conselhos, assim como uma mãe conduz seu filho.

Peço desculpas pelos momentos em que dei trabalho, mas é que às vezes é preciso errar e cair, para poder levantar-se mais forte. Considero-me uma pessoa de muita sorte, pois além de grandes Professores, aqui também construí grandes amizades. Meu muito obrigado por tudo, Professor Raul, por cada palavra, cada incentivo, por cada uma das muitas conversas agradáveis que tivemos ao longo desse mestrado, para sempre o meu carinho respeito e admiração. Assim como fui ajudado por vocês, quero fazer o mesmo por outras pessoas. Se for da vontade do Senhor, que também sejam filhos acadêmicos.

Agradeço ao Professor Dalton por aceitar integrar a banca de defesa e por todas as suas valiosas contribuições.

À todos os funcionários da Secretaria do PPGEST: Tathyanna, André, Lucas, Edenilson e a Karen por toda a paciência que tiveram sempre comigo (“[...] Tathy tem uma sala de reunião disponível?”) e pela precisão na realização de todas as atividades que contribuem para o sucesso do nosso Departamento de Estatística. A vocês meu muito obrigado, sem falar que comecei a dar trabalho a eles ainda estando na Paraíba - PB (“[...] André, Edenilson, Lucas como está a disponibilidade de bolsas e moradia estudantil aí na UnB?”). Impossível falar do EST e não mencionar a Verinha que nos recebe com tanto carinho e atenção na portaria do nosso prédio e todos os auxiliares da limpeza que conservam tão bem o nosso espaço físico.

À Comissão de Aperfeiçoamento de Pessoal do Nível Superior (CAPES) fica minha gratidão pela concessão da bolsa para a realização do meu mestrado, esta me ajudou a despertar o olhar de pesquisador.

São muitos os meus agradecimentos, infelizmente, não tem como citar todos aqui. Para sempre, irei levar todos em meu coração. Somente as bênçãos divinas poderão recompensar cada um. Para sempre, a minha eterna gratidão. Amém!

Resumo

MOURA, S. A. **Uma Proposta de Ordenação de Periódicos Científicos via Teoria da Resposta ao Item**. 2016. Dissertação (Mestrado) - Programa de Pós-Graduação em Estatística, Departamento de Departamento Universidade de Brasília, Brasília, 2017.

No Brasil ainda não existem metodologias científicas que possibilitem a ordenação de periódicos científicos. Nesta dissertação trabalhou-se com dados da área de Matemática e Estatística para se esboçar um primeiro protótipo de desenvolvimento metodológico para esta finalidade. Utilizando a Teoria da Resposta ao Item para itens politômicos, propõe-se uma metodologia para ordenação dos periódicos que permite classificá-los, em termos dos estratos utilizados pela CAPES, em função dos níveis da escala construída através do modelo. As variáveis que compõem o banco de dados utilizado nas análises consistem dos indicadores bibliométricos fornecidos pelo Scimago (<http://www.scimagojr.com/>), a Web Of Science (<http://wokinfo.com/>) e WebQualis (<http://qualis.capes.gov.br/webqualis>). Tais indicadores são importantes na avaliação da qualidade de publicações científicas. Como as referidas três bases de dados disponibilizam distintos indicadores da qualidade dos periódicos, explorando aspectos distintos, objetiva-se propor um novo indicador da qualidade de periódicos, baseado em uma análise conjunta dos referidos indicadores, via Teoria da Resposta ao Item (TRI). Para tanto, criou-se uma base de dados comum que congrega todos os indicadores disponíveis. A utilização da TRI para criar um “índice de proficiência” ou *valor dos periódicos* nas áreas de Matemática, Estatística e Matemática Aplicada fundamenta-se, principalmente, na possibilidade de criação de uma escala de medida padronizada, que permite realizar comparações entre diferentes periódicos das diferentes áreas da Matemática. Neste trabalho utilizou-se o Modelo de Resposta Gradual de Samejima (1969). Os dados contínuos foram previamente padronizados “dentro” de cada sub-área específica (Matemática, Matemática Aplicada, Estatística e Outras Áreas) na tentativa de eliminar-se efeitos sistemáticos e de criar-se uma “escala universal”. Posteriormente os dados foram discretizados, de modo a permitir o uso dos modelos politômicos de TRI. Os dados se ajustaram razoavelmente bem ao modelo e a interpretação da escala foi rica, principalmente para periódicos com traço latente em determinada região da escala. Por fim, vale ressaltar que os parâmetros utilizados neste trabalho, assim como a escala de habilidade, foram estimados com auxílio dos *softwares* Multilog e R, versão 3.3.1.

Palavras-chave: Modelos Politômicos da Teoria da Resposta ao Item; Indicadores bibliométricos; Avaliação da qualidade de publicações científicas.

Abstract

MOURA, S. A. **A Proposal for Ordering Scientific Periodicals via Item Response Theory**. 2017. Dissertation (Master degree) - Program of Statistics, Department of Statistics University of Brasília, Brasília, 2017.

In Brazil there are still no scientific methodologies that allow ordering of scientific journals. In this dissertation we worked with data from the area of Mathematics and Statistics to sketch a first prototype of methodological development for this purpose. Using Item Response Theory (IRT) for polytomous items we propose a new methodology for ordering the journals. Using the scale or metric built with the help of the IRT model, we also propose some strategy for classifying the journals into groups or strata as the ones used by CAPES. The variables that comprise the database used in the analyzes consist of the bibliometric indicators provided by Scimago (<http://www.scimagojr.com/>), Web of Science (<http://wokinfo.com/>) and WebQualis (<http://qualis.capes.gov.br/webqualis>). Such indicators are important in assessing the quality of scientific publications. As these three databases provide different indicators of journal quality, exploring different aspects, we proposed a new indicator for evaluating journal quality based on a joint analysis of these indicators, via Item Response Theory. To this end, a common database, that incorporates all the indicators available in the different datasets, was created. The use of the IRT to create a “proficiency index” or *journal value* in the areas of Mathematics, Statistics and Applied Mathematics was mainly thought to make possible the creation of a standardized measure or scale which would allow fair comparisons among different periodicals of different areas of Mathematics. In this work, the Gradual Response Model of Samejima (1969) was used. The continuous data were previously standardized within each specific sub-area (Mathematics, Applied Mathematics, Statistics and Other Areas) in an attempt to eliminate systematic effects and create a “universal scale”. Subsequently the data were discretized in order to allow the use of the IRT Polytomous Models. The data provided a reasonable fit and rich scale interpretation, especially for periodicals with latent trait in a certain region of the scale. Finally, it is worth mentioning that the parameters used in this work, as well as the latent scale, were estimated using the software Multilog and R, version 3.3.1.

Key-words: Polytomous Item Response Theory models; Bibliometric indicators; Journal quality of scientific publications.

Sumário

Lista de Abreviaturas	xxi
Lista de Símbolos	xxiii
Lista de Figuras	xxv
Lista de Tabelas	xxix
1 Introdução	1
1.1 O sistema Qualis no Comitê de Matemática	1
1.2 Apresentação do problema de pesquisa	3
1.3 Objetivos	4
1.4 Justificativa e relevância	5
1.5 Estrutura do trabalho	6
2 Apresentação dos bancos de dados	9
2.1 Periódicos no Qualis/CAPES	9
2.2 Web Of Science	11
2.3 Scimago Journal & Country Rank	15
2.4 Journal Indicators of University's Centre for Science and Technology Studies (CWTS)	17
2.5 Identificação e seleção das variáveis para a ordenação dos periódicos científicos	19
2.6 Ano-base em estudo	20
3 Teoria da Resposta ao Item (TRI)	23
3.1 Introdução	23
3.1.1 A Teoria Clássica de Medidas (TCM)	24
3.2 Interpretação dos parâmetros e da Curva Característica do Item (CCI)	26
3.3 Conceitos e Pressupostos da TRI: unidimensionalidade, independência local e escala das habilidades	28
3.4 Principais Modelos da teoria da resposta ao item	28
3.5 Modelos envolvendo um único grupo	29
3.5.1 Modelos para itens dicotômicos ou dicotomizados	29

3.5.1.1	O modelo logístico de 3 parâmetros (ML3)	30
3.5.1.1.1	<i>Escala de Medida na TRI</i>	31
3.5.1.1.2	<i>Função de Informação do Item</i>	33
3.5.1.1.3	<i>Função de Informação do Teste</i>	33
3.5.1.2	O modelo logístico de 2 parâmetros (ML2)	34
3.5.1.3	O modelo logístico de 1 parâmetros (ML1)	34
3.5.2	Modelos para itens não dicotômicos	35
3.5.2.1	Modelo de Resposta Nominal (Nominal Categories Model)	35
3.5.2.2	Modelo de Resposta Gradual (Graded Response Model)	36
3.5.2.3	Modelo de Escala Gradual (Rating Scale Model)	37
3.5.2.4	Modelo de Crédito Parcial (Partial Credit Model)	39
3.5.2.5	Modelo de Crédito Parcial Generealizado (Generalized Partial Credit Model)	39
4	Processo de estimação dos parâmetros na TRI para dados binários	41
4.1	Introdução	41
4.2	Estimação dos parâmetros dos itens	43
4.2.1	Método da Máxima Verossimilhança Condicional	43
4.2.1.1	Aplicação do algoritmo de Newton-Raphson	46
4.2.1.2	Aplicação do método “Scoring” de Fisher	50
4.2.1.2.1	<i>Estimação das habilidades na TRI</i>	51
4.2.1.2.1.1	<i>Aplicação do algoritmo Newton-Raphson</i>	52
4.2.1.2.1.2	<i>Aplicação do método “Scoring” de Fisher</i>	53
4.2.2	Método da Máxima Verossimilhança Conjunta	54
4.2.3	Método da Máxima Verossimilhança Marginal	54
4.2.3.1	Abordagem de Bock & Lieberman	55
4.3	Estimação: aplicação do algoritmo EM	59
4.4	Programas Computacionais	61
5	Processo de estimação dos parâmetros na TRI - modelos politômicos	63
5.1	Introdução	63
5.1.1	Notações e definições	64
5.2	Especificação de itens âncora	64
5.3	Equalização	66
5.4	Modelo de Resposta Gradual: Apresentação e estimação por Máxima Verossimilhança	68
5.4.1	Habilidades conhecidas	68
5.4.1.1	Habilidades conhecidas e agrupadas	71
5.4.2	Habilidades desconhecidas	74
5.4.2.1	Adaptação do algoritmo EM	75

5.4.3	Estimação das habilidades	76
6	Pré-processamento dos dados	79
6.1	Os dados do ano-base de 2012	79
6.1.1	Análise descritiva para a base de dados <i>antes</i> do processo de imputação	81
6.2	Processo de imputação na base de dados	83
6.3	Padronização dos dados	87
6.4	Discretização dos dados	89
6.5	Algumas análises descritivas adicionais para os dados pré-processados	92
6.6	Classificação dos periódicos via CART	95
7	Ordenação de periódicos científicos via Teoria da Resposta ao Item	99
7.1	Etapas da pesquisa	99
7.1.1	Pré processamento dos dados entre as etapas 1 a 5	99
7.1.2	Análise dos dados	100
7.2	Escolha da estratégia de imputação dos dados	100
7.3	Ordenação e classificação dos periódicos via TRI	102
7.3.1	Estimação dos parâmetros TRI	102
7.3.2	Ordenação dos periódicos	107
7.4	Uma proposta para a classificação dos periódicos	110
7.5	Comparação entre as classificações CAPES e via TRI	111
7.6	Analisando a persistência de efeito de grupo	114
7.7	Testando o efeito de DIF	116
7.8	Elaboração da escala de medida	118
7.8.1	Interpretação da escala em ordem decrescente	121
8	Análises via TRI para uma base de dados com mais indicadores	125
8.1	Base de dados resultante dos pareamentos	126
8.1.1	Análise descritiva dos indicadores	127
8.2	Descrição da base $N = 291$	131
8.3	Estimação dos parâmetros TRI	133
9	Considerações finais	143
9.1	Conclusões	143
	Referências Bibliográficas	147
	Apêndice A - Apresentação da linguagem <i>Python</i>	155
A.1	Apresentação da linguagem <i>Python</i>	155
A.2	Características básicas da linguagem	155
A.3	Plataformas disponíveis	156
A.4	Características da linguagem <i>Python</i>	156

A.5	Algumas das vantagens de programar utilizando o <i>Python</i>	156
A.6	Instalação do <i>Python</i>	157
A.7	A linguagem <i>Python</i> na criação de bases de dados	157
A.8	R ou <i>Python</i> para análise de dados?	157
Apêndice B - Comandos utilizando o <i>Python</i>		159
B.1	Programação	159
Apêndice C - Comandos utilizando o <i>Software R</i>		163
C.1	Programação	163
Apêndice D - Bases de dados desta pesquisa		165
D.1	Bases de dados	165
Apêndice E - Algumas análises descritivas das bases de dados desta pesquisa		167
E.1	Introdução	167
E.2	Periódicos no Qualis/CAPES	167
E.3	Distribuição conjunta de frequências	168
E.4	Representações gráficas	169
E.4.1	Gráfico Histograma	169
E.4.2	Gráfico Boxplot	169
E.4.3	Gráfico em Setores	170
E.5	Web Of Science	171
E.6	Scimago Journal & Country Rank	173
Apêndice F - Resultado das simulações para a calibração dos parâmetros dos itens		175
F.1	Resultados - Etapa I	175
F.1.1	Calibração dos parâmetros dos itens para os indicadores AIS, FI, HL e SciNet - base sem imputação	175
F.1.2	Calibração dos parâmetros dos itens para os indicadores AIS, FI e HL - base sem imputação	177
F.1.3	Calibração dos parâmetros dos itens para os indicadores AIS, FI e HL - base após imputação via média	178
Apêndice G - Dados históricos das distribuições dos periódicos por categoria de classificação da CAPES		181
G.1	Resultados e discussão da proposta - Etapa I	181
G.1.1	Comparação do desempenho estimado no modelo da TRI versus estabelecido pela CAPES	181
G.2	Algumas considerações	186

Apêndice H - Comandos desenvolvidas em <i>Visual Basic for Applications</i> (VBA)	187
H.1 Programação em VBA	187
Apêndice I - Sintaxes desenvolvidas no <i>software Parscale</i>, versão 4.1	191
I.1 Sintaxe Parscale - testando o efeito de DIF	191
Apêndice J - Apresentação de algumas árvores de decisão	195
J.1 Árvores de decisão e suas respectivas representações no espaço	195
J.1.1 Modelo $FI \sim AIS + HL + \varepsilon$	195
J.1.2 Modelo $AIS \sim FI + \varepsilon$	198
J.2 Algumas medidas resumo para os modelos: $AIS \sim FI + HL + \varepsilon$, $FI \sim$ $AIS + HL + \varepsilon$ e $AIS \sim FI + \varepsilon$	201
Apêndice K - Pareamento das bases de dados desta pesquisa	203
K.1 Conceito de Pareamento Determinístico	204
K.2 Medidas de similaridade	204
.2.1 Distância de Ratcliff/Obershelp	205
K.2.2 Distância de Jaro-Winkler	206
K.2.3 Distância de Jaro	208
K.2.4 Distância de Jaccard	209
K.2.5 Distância de Levenshtein	210
K.2.6 Distância de Sörensen-Dice	210
K.3 Resultado dos pareamentos	211
K.3.1 Pareamentos estabelecidos através do ISSN	212
K.3.2 Pareamentos estabelecidos através do título do periódico	213
K.3.3 Busca sistemática do ISSN via <i>Web Crawler</i>	216
K.4 Resumo de procedimentos adotados	219
K.5 Base de dados final	220
Apêndice L - Processo de imputação de dados	221
L.1 Imputação de dados	221

Lista de Abreviaturas

ISSN	Número Internacional Normalizado para Publicações Seriadas (<i>International Standard Serial Number</i>)
JCR	Citação de Periódicos por Jornal (<i>Journal Citation Reports</i>)
AI	Influência de um Artigo (<i>Article Influence</i>)
SJR	Indicador/Índice do Jornal Scimago por Rank (<i>Jornal Scimago Indicador por Rank</i>)
HL	Meia Vida (<i>Half-Life</i>)
MCQ	índice MCQ do Mathematical Reviews da American Mathematical Society (<i>Mathematical Citation Quotient</i>)
Q_i	Quantis da Classificação SJR (<i>Quantile classification SJR</i>)

Lista de Símbolos

- U_{ij} é uma variável dicotômica que assume os valores 1 (um), quando o indivíduo, isto é, a unidade pesquisada j responde acertadamente o item i , ou assume 0 (zero) quando o indivíduo j não responde corretamente ao item i
- θ_j representa a habilidade (traço latente) do j -ésimo indivíduo
- b_i é o parâmetro de dificuldade (ou de posição) do item i , medido na mesma escala da habilidade
- a_i é o parâmetro de discriminação (ou de inclinação) do item i , com valor proporcional à inclinação da Curva Característica do Item - CCI no ponto b_i
- c_i é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item i (muitas vezes referido como a probabilidade de acerto casual)
- D representa o fator constante de escala igual a 1. Pode-se utilizar o valor de 1.7 quando se deseja que os resultados da função logística sejam parecidos com os resultados da função ogiva normal
- ζ_i conjunto de parâmetros dos itens, em que $i = 1, \dots, I$
- μ média, medida de posição para o traço latente
- σ desvio-padrão, medida de dispersão para o traço latente
- τ desvio-padrão na escala transformada
- π população composta por n indivíduos

Lista de Figuras

2.1	Visualização da página na internet ao acessar o endereço eletrônico da Web- Qualis. Último acesso em: 24 de fevereiro de 2016.	11
2.2	Visualização da página na internet ao acessar o endereço eletrônico da Web Of Science. Online; último acesso em: 28 de fevereiro de 2016.	14
2.3	Visualização da página na internet ao acessar o endereço eletrônico da Scimago Journal & Country Rank (SJR). Online; último acesso em: 28 de fevereiro de 2016.	17
2.4	Visualização da página na internet ao acessar o endereço eletrônico da CWTS Jornal indicadores. Online; último acesso em: 26 de fevereiro de 2016.	19
3.1	Exemplo politômico de uma Curva Característica do Item - CCI. Adaptado de Turik (2010)	27
3.2	Exemplo de uma Curva Característica do Item - CCI. Adaptado de Andrade et al. (2000)	31
3.3	Exemplos de curvas características e de informação de vários itens. Adaptado de Andrade et al. (2000)	32
3.4	Curva de informação e erro padrão dos testes A e B. Adaptado de Embretson e Reise (2000, p. 168)	35
3.5	Representação gráfica dos Modelos de Escala Gradual e de Resposta Gradual. Adaptado de Andrade et al. (2000)	38
4.1	Proporção observada da resposta correta em função da habilidade. Adaptado de Baker e Kim (2004)	42
6.1	Representações gráficas para os indicadores bibliométricos: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) <i>antes</i> do processo de imputação.	82
6.2	Representações gráficas para os indicadores bibliométricos: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) após o processo de imputação.	88
6.3	Diagramas de dispersão para os indicadores bibliométricos observados: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) <i>após</i> pré-processamento.	92

6.4	Gráfico 3D para os indicadores bibliométricos observados: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) <i>após</i> o processo de imputação via média e padronização.	93
6.5	Gráfico 3D para os indicadores bibliométricos observados: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) <i>após</i> o processo de imputação via mediana e padronização.	93
6.6	Análise de Componentes Principais (PCA) após o processo de imputação via média.	94
6.7	Análise de Componentes Principais (PCA) após o processo de imputação via mediana.	94
6.8	Representação gráfica do CART para a base de dados após imputação pela média.	96
6.9	Representação gráfica das partições do CART após imputação pela média.	97
7.1	Representação gráfica da dispersão do posto TRI (Média versus Mediana)	101
7.2	Representação gráfica da curvas características dos itens. Elaborado pelos autores (2017).	104
7.3	Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	105
7.4	Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	105
7.5	Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	106
7.6	Curva de Informação do Teste (CIT) e o erro padrão de medição entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	106
7.7	Boxplots dos escores de proficiência dos periódicos em todas as áreas de interesse.	109
7.8	Classificação de periódicos CAPES <i>versus</i> TRI, ano base média de 2012 a 2015, $N = 1189$ periódicos. Elaborado pelos autores (2017).	114
7.9	Histograma para os escores de proficiência dos periódicos em todas as áreas, $N = 1189$ periódicos. Elaborado pelos autores (2017).	115
7.10	Histograma para os escores de proficiência dos periódicos por áreas, $N = 1189$ periódicos. Elaborado pelos autores (2017).	115
7.11	Representação gráfica das duas distribuições latentes estimadas, isto é, com grupos e sem grupos. Elaborado pelos autores (2017).	117
7.12	Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	119

7.13	Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	119
7.14	Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	120
7.15	Curva de Informação do Teste (CIT) e o erro padrão de medição entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	120
7.16	Posicionamento dos periódicos A1, A2 e B1 divulgados pela CAPES-2012 na escala de habilidades construída com base nos 882 periódicos. As linhas verticais referem-se às posições dos níveis âncora. Elaborado pelos autores (2017).	123
7.17	Posicionamento dos periódicos B2, B3, B4 e B5 divulgados pela CAPES-2012 na escala de habilidades construída com base nos 882 periódicos. As linhas verticais referem-se às posições dos níveis âncora. Elaborado pelos autores (2017).	124
8.1	Representações gráficas para os indicadores bibliométricos: MPFI, AIS, ES e CHL <i>antes</i> do processo de imputação, $N = 291$ periódicos.	127
8.2	Representações gráficas para os indicadores bibliométricos: II, FI5A, FISC e FIJ <i>antes</i> do processo de imputação, $N = 291$ periódicos.	128
8.3	Scatter Plot para os indicadores bibliométricos observados: MPFI, AIS, ES, CHL, II, FI5A, FISC e FIJ <i>antes</i> o processo de imputação, dados na escala original, $N = 291$ periódicos.	129
8.4	Representações gráficas para os indicadores bibliométricos: FIS, IH, REF/DOC e CITES/DOC <i>antes</i> do processo de imputação, $N = 291$ periódicos.	130
8.5	Scatter Plot para os indicadores bibliométricos observados: FIS, IH, REF/DOC e CITES/DOC <i>antes</i> o processo de imputação, dados na escala original, $N = 291$ periódicos.	131
8.6	Curva Característica do Item (CCI) e a Função de Informação do Item (MPFI) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	135
8.7	Curva Característica do Item (CCI) e a Função de Informação do Item (CITES/DOC) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	136
8.8	Curva Característica do Item (CCI) e a Função de Informação do Item (FIS) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	136

8.9	Curva Característica do Item (CCI) e a Função de Informação do Item (AIS) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	137
8.10	Curva Característica do Item (CCI) e a Função de Informação do Item (IH) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	137
8.11	Curva Característica do Item (CCI) e a Função de Informação do Item (ES) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	138
8.12	Curva Característica do Item (CCI) e a Função de Informação do Item (II) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	138
8.13	Curva Característica do Item (CCI) e a Função de Informação do Item (REF/-DOC) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	139
8.14	Curva Característica do Item (CCI) e a Função de Informação do Item (FI5A) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	139
8.15	Curva Característica do Item (CCI) e a Função de Informação do Item (CHL) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	140
8.16	Curva de Informação do Teste (CIT) e o erro padrão de medição entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).	140

Lista de Tabelas

1.1	Indicadores de qualidade utilizados Classificação de Periódicos no Qualis/-CAPES, ano base de 2012-2014, CAPES (2014).	3
2.1	Representação tabular parcial dos dados coletados na Classificação de Periódicos no Qualis/CAPES, ano base de 2012.	10
2.2	Representação tabular parcial dos dados coletados no Web Of Science, ano base de 2014.	12
2.3	Representação tabular parcial dos dados coletados no Scimago Journal & Country Rank (SJR), ano base de 2014.	15
2.4	Representação tabular parcial dos dados coletados no Journal Indicators da University's Centre for Science and Technology Studies (CWTS), ano base de 2014.	18
2.5	Representação das variáveis que compõem as bases de dados Qualis/CAPES, Web Of Science e Scimago, desta pesquisa.	20
3.1	Classificação dos modelos unidimensionais de acordo com a natureza do item.	40
4.1	Principais vantagens e desvantagens dos métodos de máxima verossimilhança para a TRI.	60
4.2	Programas computacionais com breves descrições.	62
5.1	Exemplo de itens âncoras: Escala de satisfação e definição de itens âncoras e quase-âncoras.	65
5.2	Diferentes tipos de equalização.	67
6.1	Resultados apurados a partir do arquivo PDF de 2012.	80
6.2	Quantitativo dos registros no arquivo pdf com informações ausentes nos indicadores segundo suas classificações por estrato.	80
6.3	Quantitativo dos registros com informações ausentes da Tabela 6.2, segundo suas classificações nos estratos e grupos.	81
6.4	Representação tabular dos periódicos científicos por área de interesse antes do processo de imputação.	82

6.5	Análise descritiva dos indicadores bibliométricos antes da imputação (dados originais).	83
6.6	Quantitativo de registros (entre os 882) utilizados como base para o processo de imputação segundo suas classificações por Área.	84
6.7	Representação tabular dos periódicos científicos, por área de interesse, após o processo de imputação.	85
6.8	Análise descritiva dos indicadores bibliométricos após a imputação via média.	85
6.9	Análise descritiva dos indicadores bibliométricos após a imputação via mediana.	86
6.10	Representação tabular dos NaN's para base de dados após o processo de imputação via média, $N = 1189$ periódicos.	86
6.11	Análise descritiva dos indicadores bibliométricos padronizados (z) após a imputação via média.	89
6.12	Análise descritiva dos indicadores bibliométricos padronizados (z) após a imputação via mediana.	89
6.13	Representação tabular dos indicadores bibliométricos disponíveis na base de dados de 2012. Os dados foram previamente imputados (via média) e padronizados. Cada categoria de 2 a 9, encerra 12.5% da distribuição da variável contínua (AIS, FI ou HL).	90
6.14	Representação tabular dos indicadores bibliométricos disponíveis na base de dados de 2012. Os dados foram previamente imputados (via mediana) e padronizados. Cada categoria de 2 a 9, encerra 12.5% da distribuição da variável contínua (AIS, FI ou HL).	91
6.15	Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo o CART.	97
7.1	Categorias estabelecidas para cada um dos itens ao combinarmos categorias adjacentes considerando-se as categorias preliminares C_1 a C_9	103
7.2	Estimativas dos parâmetros dos itens AIS, FI e HL da base de dados imputada, CAPES/Qualis ano-base 2012, $N = 1189$ periódicos. EP (erro padrão da estimativa).	103
7.3	Representação tabular parcial da ordenação dos periódicos científicos.	107
7.4	Representação tabular da frequência absoluta e relativa dos escores dos periódicos “proficiência dos periódicos” na ordenação dos periódicos científicos por área de interesse.	108
7.5	Representação tabular da classificação dos periódicos científicos por ano em termos percentuais (%) segundo o método apresentado pelo Qualis/CAPES.	110
7.6	Classificação TRI nos estratos A1 a B5 considerando postos/periódicos ordenados em ordem decrescente.	111

7.7	Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área por área - método de classificação via TRI, considerando-se a média das proporções de 2012 a 2015, $N = 1189$ periódicos, e o método proposto nesta dissertação	112
7.8	Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo os critérios considerados pelo Qualis/CAPES, ano base 2012.	112
7.9	Representação tabular da classificação Qualis/CAPES <i>versus</i> classificação proposta pela TRI, media das proporções de 2012 a 2015, $N = 1189$ periódicos.	113
7.10	Representação tabular da distribuição percentual (%) por classificação Qualis/CAPES <i>versus</i> classificação proposta pela TRI, media das proporções de 2012 a 2015, $N = 1189$ periódicos.	113
7.11	Estatística Qui-quadrado (χ^2): teste do efeito de grupos entre o escore dos periódicos.	117
7.12	Estimativas dos parâmetros dos itens AIS, FI e HL da base de dados sem imputação, CAPES/Qualis ano-base 2012, $N = 882$ periódicos. EP (erro padrão da estimativa).	118
7.13	Escala (0,1) dos escores e definição dos níveis âncora e quase-âncora associados às categorias dos itens em estudo.	122
8.1	Informações sobre os periódicos das bases de dados: Classificação de Periódicos no Qualis/CAPES, Web Of Science e Scimago Journal & Country Rank, ano base 2012.	126
8.2	Representação tabular dos periódicos científicos por área de interesse, base de dados antes do processo de imputação, $N = 291$ periódicos.	132
8.3	Análise descritiva dos indicadores bibliométricos após o processo de imputação via média, $N = 291$ periódicos.	132
8.4	Categorias estabelecidas para cada um dos itens ao combinarmos categorias adjacentes considerando-se as categorias preliminares C_1 a C_9	134
8.5	Estimativas dos parâmetros dos itens FI5A, MPFI, ES, AIS, CHL, II, REF/DOC, CITES/DOC, IH e FIS da base de dados imputada, produto dos pareamentos ano-base 2012, $N = 291$ periódicos. EP (erro padrão da estimativa).	134

Capítulo 1

Introdução

O Scimago (<http://www.scimagojr.com/>) e a Web Of Science (<http://wokinfo.com/>), entre outros, são sites que disponibilizam listas, utilizando diferentes métricas, para a classificação de periódicos nas mais diversas áreas. Tais métricas são importantes na avaliação da qualidade dos periódicos científicos. Vale ressaltar que um “periódico” é definido como sendo um veículo periódico de divulgação impressa e/ou eletrônica, que contém trabalhos com conteúdo de Matemática e/ou Probabilidade e/ou Estatística, e suas aplicações, possuindo ISSN, corpo e política editorial e sistema de arbitragem pelos seus pares (CAPES, 2014).

A partir das informações disponíveis nos sites citados estudaremos Modelos de Resposta ao Item politômicos no intuito de propor uma ordenação conjunta dos periódicos científicos nas áreas de Estatística, Matemática, Matemática Aplicada.

O produto desta dissertação pode ser útil para a adoção de políticas para a Pós-graduação no Brasil.

1.1 O sistema Qualis no Comitê de Matemática

O Qualis é definido pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) como o conjunto de procedimentos utilizados para a divulgação da produção intelectual dos programas de Pós-graduação *stricto sensu* (mestrado e doutorado). Até recentemente os periódicos eram classificados quanto ao âmbito de circulação (Local, Nacional, Internacional) e quanto à qualidade (A, B, C), para cada área de avaliação. Assim, este conjunto de informações fornecido pela CAPES tem servido para fundamentar o processo de avaliação do Sistema Nacional de Pós-Graduação da CAPES.

Como supra citado, tem-se que os dados a serem trabalhados nesta pesquisa, referem-se ao ano base de 2014. Tal escolha foi motivada pelo fato da CAPES haver realizado, recentemente, a classificação do Qualis da área de Matemática, Probabilidade e Estatística, relativa aos anos 2013 e 2014. Tal classificação envolveu apenas a produção dos programas de Pós-graduação nesses anos. Destaca-se que esta classificação tem um caráter provisório, e que a classificação final, relativa ao quadriênio de avaliação 2013-2016, será estabelecida no futuro (CAPES, 2014). Devido a essas circunstâncias, nesta pesquisa, os dados a serem trabalhados referem-se ao ano base de 2012.

Vale ressaltar que as regras de classificação fornecidas pelo Qualis CAPES são tais que os periódicos estão divididos em sete estratos: A1, A2, B1, B2, B3, B4 e B5, em ordem decrescente de importância. Há ainda o estrato C, dedicado à produção que não se adequa à área ou à definição de periódico.

No que concerne à distribuição dos periódicos nos diversos estratos, eles devem ficar, em cada ano (2013 e 2014), dentro dos limites estabelecidos pela CAPES: o número de periódicos

A1 deve ser menor que o número de periódicos A2, sendo que a porcentagem de periódicos classificados como A1 e A2 deve ser inferior a 25%. Adicionalmente, o número de periódicos no estrato A deve ser menor que o número de periódicos no estrato B1 e a porcentagem de periódicos classificados como A1, A2 e B1 não deve ser superior a 50% do total de periódicos classificados nos estratos A e B.

Neste contexto, o sistema de classificação foi atualizado pela CAPES (2014). Inicialmente, foram considerados os indicadores: fatores de impacto JCR, Article Influence (AI), SJR, Meia-Vida (HL) e o índice MCQ do Mathematical Reviews da American Mathematical Society. Os índices JCR e SJR foram fornecidos pela CAPES, enquanto os outros índices foram obtidos pela comissão que elaborou o Qualis (CAPES, 2014). As definições detalhadas destes indicadores serão apresentadas no Capítulo 2.

A metodologia de classificação dos periódicos na área da Matemática considerada pela CAPES em 2014 é descrita a seguir. Salienta-se que nos anos de 2013 e 2014 cada sub-área definiu, em separado, os seus critérios de classificação. Claramente, tal procedimento pode gerar assimetrias nas decisões a serem tomadas acerca da classificação geral dos periódicos.

Matemática:

Estrato A1:

$AI \geq 1,0$, $MCQ \geq 0,8$, $HL \geq 10$, $JCR \geq 0,7$.

Estrato A2:

caso 1: $HL \geq 10$, $AI \geq 0,7$, $JCR \geq 0,5$ e $MCQ \geq 0,6$, ou

caso 2: $HL \geq 7$, $AI \geq 1,0$, $JCR \geq 0,7$ e $MCQ \geq 0,7$.

Estrato B1:

$MCQ \geq 0,4$, $AI \geq 0,45$ e

caso 1: $HL \geq 9$, e $JCR \geq 0,34$,

caso 2: $HL \geq 5$ e $JCR \geq 0,55$,

Estrato B2:

$MCQ \geq 0,4$ ou $AI \geq 0,6$ mais análise de HL e outros índices de impacto.

Estrato B3:

$MCQ \geq 0,3$ mais análise de HL e outros índices de impacto e política editorial.

Estrato B4:

$MCQ \geq 0,2$ mais análise de HL e outros índices de impacto e política editorial.

Matemática Aplicada:

Conjuntamente com critérios relativos ao “Article Influence” (AI) : A1: $AI > 1,0$; A2: $AI > 0,7$; B1: $AI > 0,5$; B2: $AI > 0,3$, observação de políticas editoriais e utilização de análise comparativa de acordo Q_i (quartis da classificação SJR), foram considerados índices de Meia-Vida (HL) e Fator de Impacto (JCR) de acordo com a Tabela 1.1:

Tabela 1.1: Indicadores de qualidade utilizados Classificação de Periódicos no Qualis/CAPES, ano base de 2012-2014, CAPES (2014).

Indicadores	Fator de impacto				
	>1,00	0,8 - 1,0	0,6 - 0,8	0,4 - 0,6	< 0,4
$HL > 10$	A1	A2	B1	B2	B3
$6 \leq HL \leq 10$	A2	B1	B2	B3	B4
$4 \leq HL \leq 6$	B1	B2	B3	B4	B4
$3 \leq HL \leq 4$	B2	B3	B3	B4	B5
$0 \leq HL \leq 3$	B3	B3	B4	B5	B5

Fonte: Elaborado por CAPES (2014).

Probabilidade e Estatística:

Estrato A1:

$HL > 10$ e $AI > 1,25$

Estrato A2:

$HL > 10$ e $AI > 0,75$ ou

$6 < HL < 10$ e $AI > 1,0$

As revistas no limite entre A2 e B1 tiveram especial atenção, examinando-se uma combinação dos seguintes índices: SJR, JCR, HL, AI, Q_i (quartis da classificação SJR). Foi proposta mudança do estrato B1 para A2 apenas quando houve forte evidência do conjunto desses índices para promoção.

Estratos B:

Os periódicos foram examinados caso a caso. Foram observados os índices disponíveis entre os seguintes: SJR, JCR, HL, AI e Q_i . Em casos de ausência desses índices, foi feita avaliação a partir do corpo e política editorial da revista e artigos publicados.

No documento CAPES (2014), enfatiza-se que há uma grande heterogeneidade dentro da área no que se refere aos indicadores de qualidade e também que o Qualis da Matemática, Probabilidade e Estatística inclui um grande número de revistas de outras áreas, que provavelmente não poderiam ser avaliadas nos mesmos termos.

Como se observa, os critérios utilizados por cada uma das áreas que constituem o comitê de Matemática e Estatística utilizam métodos “Ad hoc” e não comparáveis entre si. Tal realidade dificulta a categorização dos periódicos nos estratos de maneira satisfatória, de acordo com argumentos não-subjetivos.

1.2 Apresentação do problema de pesquisa

A ordenação das revistas científicas via modelos de classes latentes para dados de resposta ao item politômicos, segundo Bartolucci *et al.* (2015), caracteriza-se como uma estratégia para classificar os periódicos científicos a partir de um conjunto de indicadores quantitativos disponíveis, que representam medidas imperfeitas do valor não observável “da revista de interesse”. Os autores complementam que, depois da discretização dos indicadores disponíveis estima-se um modelo de classe latente para os dados de resposta ao item politômico e, assim utiliza-se o modelo estimado para classificar cada jornal.

Neste contexto, Varin *et al.* (2016), ao abordarem a questão das limitações na gestão do estabelecimento do ranking das revistas científicas, destacam que “devido à disparidade

entre a percepção comum de prestígio das revistas e sua classificação com base em contagem de citações, isso acarreta uma preocupação mais grave, que é gerada pelo uso inadequado de rankings de periódicos para avaliar a influência científica dos pesquisadores”.

Como se pode verificar na exposição dos autores, inexistente uma padronização na medição da classificação e rankings dos periódicos científicos. Há uma lacuna existente no que concerne a uma metodologia única de ordenação para a mensuração e classificação de periódicos. É importante observar também, o potencial existente para o desenvolvimento de modelos que possibilitam a ordenação de revistas científicas que considerem todos os indicadores bibliométricos para tal avaliação.

Neste sentido, a Teoria da Resposta ao Item - TRI, apresenta-se como uma técnica adequada na avaliação desses periódicos científicos, à medida que possibilita o desenvolvimento de uma escala passível de interpretação.

Corroborando com o exposto, *Andrade et al. (2000)*, enfatizam que a TRI oferece modelos para a estimação dos traços latentes e meios de representar a relação entre o desempenho de uma pessoa (ou empresa) em um determinado campo do conhecimento e a probabilidade desta possuir determinada característica que está sendo avaliada.

Vale ressaltar a inexistência de modelos que possibilitem a ordenação de periódicos científicos pela CAPES. A Teoria da Resposta ao Item apresenta-se como ferramenta de análise de modo a mostrar o comportamento e a influência de múltiplos fatores na estimação do *valor* de um periódico. A TRI possibilita classificar os periódicos em função dos níveis da escala construída através do modelo, ou seja, uma escala que é passível de ser interpretada em função dos indicadores bibliométricos analisados.

Dentre as grandes vantagens que a Teoria de Resposta ao Item oferece (*Embretson e Reise, 2000*) destacam-se: (1) possibilita uma melhor análise de cada item que forma o instrumento de medida, pois leva em consideração suas características específicas de construção de escalas; (2) o nível de uma característica que um indivíduo possui pode ser comparado ao nível da característica exigida pelo item; isso facilita a interpretação da escala gerada e permite também conhecer quais itens estão produzindo informação ao longo da escala.

Por fim, vale ressaltar que os modelos apresentados neste trabalho são modelos paramétricos e unidimensionais e cumulativos. No entanto, também existem modelos de TRI não paramétricos e multidimensionais e não cumulativos.

Considerando o exposto, o presente estudo tem como propósito elucidar o seguinte problema de pesquisa: **apresentar uma proposta de ordenação de periódicos científicos via Teoria da Resposta ao Item.**

1.3 Objetivos

Neste trabalho de dissertação objetiva-se criar um “índice de proficiência” ou valor dos periódicos na área de Matemática, Estatística e Matemática Aplicada baseado em Modelos Politômicos da Teoria da Resposta ao Item. Os itens que irão compor o banco de dados a ser utilizado nas análises consistem dos indicadores bibliométricos descritos no Capítulo 2. Tais indicadores serão discretizados de modo a obter-se categorias ordenadas.

No intuito de alcançar o objetivo geral, estabelecem-se os seguintes objetivos específicos:

- i)* identificar os componentes que influenciam a ordenação dos periódicos científicos;
- ii)* investigar o impacto dos indicadores descritos nas bases de dados desta pesquisa no que tange à classificação dos periódicos;

- iii)* desenvolver uma escala de medida padronizada para a classificação dos periódicos científicos nas áreas de Estatística, Matemática e Matemática Aplicada e;
- iv)* avaliar a consistência das classificações apresentadas pelo Qualis CAPES.

1.4 Justificativa e relevância

A presente dissertação é relevante no campo teórico, pois engloba o desenvolvimento de um modelo que possibilita a avaliação e a ordenação de periódicos científicos considerando-se os diversos indicadores bibliométricos disponíveis para análise.

Nesta mesma linha, vale ressaltar que uma vez que os diversos tipos de indicadores bibliométricos não estão descritos em uma base de dados comum, mas sim em quatro bancos de dados distintos, a saber: Qualis, Web Of Science, Scimago e CWTS, faz-se necessário construir um banco de dados que contemple, para um mesmo periódico, a diversidade de indicadores bibliométricos existentes.

Inicialmente a base de dados Journal Indicators CWTS (<http://www.journalindicators.com/>) foi incluída na pesquisa. No entanto, devido a brevidade do tempo para concluir os estudos desta dissertação e por diversas incongruências nesta base, decidiu-se considerá-la apenas para estudos futuros. No Capítulo 2 e no apêndice discorre-se sobre algumas das dificuldades encontradas no uso da base CWTS.

A construção de uma base de dados comum requer a utilização de linguagens adequadas à manipulação de dados massivos, como a linguagem *Python*. Dessa forma, a partir da utilização de bibliotecas específicas para análise de similaridade de strings disponíveis em *Python* (vide breve descrição da linguagem no **Apêndice A**), bem como técnicas de pareamentos e algoritmos de indexação de bases, construiu-se um banco de dados a ser detalhado no **Apêndice K**.

Considerando os indicadores bibliométricos sugeridos pelas bases de dados supra citadas, tem-se como finalidade apresentar um modelo de TRI que, com base nesses indicadores, permita a ordenação dos periódicos e a classificação destes nos extratos A1, A2, B1, B2, B3, B4, B5 e C.

Os modelos de Teoria da Resposta ao Item possibilitam a criação de uma escala de medida interpretável para tal ordenação e classificação desses periódicos, a partir da identificação dos fatores/indicadores utilizados da análise. Por conseguinte, a escala a ser construída pode permitir realizar comparações entre diferentes áreas e identificar os níveis de desempenho nos quais cada periódico se localiza na escala desenvolvida.

A escala formulada possibilitará, também, verificar quais os fatores são mais determinantes na diferenciação entre os periódicos.

Neste contexto, a relevância da pesquisa encontra-se na definição de uma escala ou métrica que seja útil na ordenação de periódicos científicos.

Com o presente estudo também objetiva-se avaliar os fatores que influenciam na classificação de periódicos científicos.

Vale ressaltar que, no início deste trabalho de dissertação, não dispunha-se de uma base de dados que englobasse os principais indicadores existentes, possibilitando uma avaliação da classificação atribuída aos periódicos nas áreas da Matemática, Estatística e Matemática Aplicada. Inicialmente, considerou-se as bases de dados do Journal Indicators CWTS (<http://www.journalindicators.com/>), Scimago (<http://www.scimagojr.com/>) e a Web Of Science (<http://wokinfo.com/>). Tais bases disponibilizam listas, utilizando diferentes métricas, para a classificação de periódicos.

No comitê avaliador de Matemática e Estatística, no ano de 2012, instituiu-se o uso da classificação *Core* para designar os periódicos em que a área prioritária de desenvolvimento da pesquisa é de domínio Matemático/Estatístico. Por exemplo, o periódico *Acta Arithmetica* é considerado *core*, enquanto o periódico *ACM Computation Surveys* é considerado não *core*.

Devido às inconsistências nas informações contidas nas referidas bases de dados e, mesmo, a inexistência de “chave” única, como o ISSN, para muitos dos periódicos, fez-se necessário o uso de técnicas avançadas de identificação de similaridade léxica entre os títulos dos periódicos nas bases.

A base de dados CWTS, por exemplo, não inclui a informação sobre o ISSN dos periódicos. Desta feita, o pareamento das informações dessa base com as das bases, Scimago, Qualis e Web Of Science precisa ser feita através do estudo das similaridades dos títulos dos periódicos. Os erros falso-positivo (um pareamento incorreto) ou falso-negativo (um pareamento correto que não foi estabelecido), são inevitáveis. No entanto, a probabilidade da ocorrência de tais erros é minimizada ao utilizar-se métricas de similaridade adequadas.

Considerando o exposto, uma das contribuições teóricas da pesquisa refere-se à construção de uma base de dados final consistente, composta por diversos indicadores bibliométricos, que atendem aos objetivos desta pesquisa e que, também, contribuirá para outros estudos futuros.

No entanto, dado às limitações de tempo e priorização de alguns objetivos principais da dissertação, optou-se pelo desenvolvimento de análises apenas para uma base de dados de 2012, contendo 1189 periódicos e para uma base de dados contendo 291 periódicos mas que incorpora vários indicadores bibliométricos.

1.5 Estrutura do trabalho

O presente trabalho está estruturado em nove capítulos. O conteúdo do primeiro capítulo, que finaliza neste tópico, foi mencionado na apresentação do capítulo.

No Capítulo 2, apresenta-se os bancos de dados utilizados na dissertação: Periódicos no Qualis/CAPEs, Web Of Science, Scimago Journal & Country Rank e Journal Indicators da University’s Centre for Science and Technology Studies (CWTS). Em cada uma dessas bases de dados, descreve-se suas variáveis, principais características e procedência.

No Capítulo 3 discorre-se sobre a teoria da resposta ao item - TRI, enfatizando-se aspectos como os pressupostos, principais vantagens em relação à teoria clássica dos testes, modelos da teoria e suas classificações. Ressalta-se o modelo de resposta gradual, por ser o modelo a ser aplicado neste trabalho para o desenvolvimento de uma escala de medida objetivando-se a ordenação dos periódicos.

Na sequência, no Capítulo 4, discute-se o processo de estimação dos parâmetros na TRI para dados binários, a estimação das habilidades para os parâmetros dos itens e a estimação conjunta dos parâmetros dos itens e habilidades. Já o Capítulo 5 enfatiza-se o processo de estimação dos parâmetros na TRI para o Modelo de Resposta Gradual (MRG) por máxima verossimilhança.

O Capítulo 6 discorre sobre a metodologia da pesquisa. Neste sentido, o capítulo enfatiza o delineamento metodológico da pesquisa, as definições constitutiva e operacional das variáveis, categorias de análise, os pressupostos da pesquisa, a delimitação da pesquisa, a descrição do procedimento metodológico e alguns procedimentos e critérios adotados na pesquisa de periódicos.

Na Seção 6.6 apresenta-se a classificação dos periódicos científicos via CART. Neste capítulo objetiva-se estimar “pontos de corte” adequados nos indicadores bibliométricos pré-

processados (fases de imputação e padronização) que permitam discriminar os periódicos de uma forma bastante prática e de uso simplificado por parte dos agentes da CAPES.

O Capítulo 7 apresenta-se os resultados e discussão da proposta: ordenação de periódicos científicos via Teoria da Resposta ao Item a sistemática proposta para a ordenação de periódicos científicos via Teoria da Resposta ao Item, cujo foco consiste na descrição dos resultados obtidos na etapa I desta pesquisa (vide Capítulo 6). No Capítulo 8 mostra os resultados e discussão da proposta: ordenação de periódicos científicos via Teoria da Resposta ao Item para uma base final com mais indicadores, em que o foco deste capítulo consiste na descrição dos resultados obtidos na etapa II desta pesquisa (vide Capítulo 6).

É importante mencionar que a presente pesquisa é formada por duas bases de dados finais. Inicialmente, tem-se uma base formada apenas pelos indicadores fornecidos pela CAPES (AIS, FI e HL vide Capítulo 7), no entanto, a segunda base, a que será apresentada no Capítulo 8, é composta por todos os indicadores que compoem as três fontes de dados supra citadas. Mas detalhes sobre estas etapas, encontram-se no escopo deste texto.

Considerando os critérios expostos e o procedimento de imputação apresentada na Seção 6.2, a base um é constituída por 1189 periódicos e formada por três indicadores bibliométricos (AIS, FI e HL), já a base dois é formada por 291 periódicos com 12 indicadores, a saber:

- FISC: fator de impacto sem auto citações (Web Of Science);
- FIJ: fator de impacto (Web Of Science);
- FI5A: fator de impacto 5 anos (Web Of Science);
- MPFI: média percentual do fator de impacto (Web Of Science);
- ES: eigenfactor score (Web Of Science);
- AIS: Article Influence Score (Web Of Science);
- CHL: cited half life (Web Of Science);
- II: immediacy index (Web Of Science);
- REF/DOC: número médio de referências por documentos em um dado ano (Scimago);
- CITES/DOC: média de citações por documento no período de 2 anos (Scimago);
- IH: índice H (Scimago);
- FIS: fator de impacto (Scimago).

O perfil dessas bases, isto é, as suas características, estão apresentadas no Capítulo 2.

No Capítulo 9, destacam-se as considerações finais da pesquisa, enfatizando-se as principais conclusões quanto aos objetivos e as sugestões para futuras pesquisas.

Enfatiza-se, na sequência, logo após ao Capítulo 9, os apêndices desta pesquisa. Nesta etapa destaca-se: uma apresentação sobre a linguagem *Python*, a programação em *Python* e em R, ambos os códigos, criados pelos autores, para serem utilizados neste trabalho de dissertação, na sequência, algumas análises descritivas sobre as bases de dados que foram apresentadas no Capítulo 2, algumas definições e referências bibliográficas sobre análise multivariada de dados, resultado das simulações para a calibração dos parâmetros dos itens, dados históricos das distribuições dos periódicos por categoria de classificação da CAPES, Comandos desenvolvidas em *Visual Basic for Applications (VBA)*, sintaxes desenvolvidas

no *software* Parscale, versão 4.1, a apresentação de algumas árvores de decisão e, por fim, apresenta-se o conceito de pareamento determinístico, as medidas de similaridade empregadas nos pareamentos e bases de dados resultantes considerando a CWTS.

Capítulo 2

Apresentação dos bancos de dados

Neste capítulo descreve-se alguns bancos de dados que trazem indicadores bibliométricos de periódicos científicos: Periódicos no Qualis/CAPES, Web Of Science, Scimago e CWTS, respectivamente, que são alvos desta pesquisa. Apresentaremos cada um destes bancos de dados, descrevendo suas variáveis, principais características e procedência dos mesmos.

Vale ressaltar que os referidos bancos de dados foram verificados até mesmo manualmente, na investigação de problemas de inconsistência de natureza diversa.

Tal procedimento foi motivado devido às diferentes formatações dos dados e às necessidades específicas deste trabalho.

2.1 Periódicos no Qualis/CAPES

O Qualis é uma lista de periódicos organizada pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) com o intuito de mensurar a qualidade da produção científica dos programas de Pós-graduação. Na verdade, trata-se de uma classificação indireta, visto que não avalia a qualidade das pesquisas ou dos artigos produzidos, e sim dos periódicos científicos em que eles são publicados. Como resultado, o Qualis fornece uma lista com a classificação dos periódicos nos quais os pesquisadores nacionais publicam os seus trabalhos científicos. A cada periódico incluído no Qualis é atribuído uma categoria de classificação indicativa da importância ou qualidade científica do periódico, sendo tal classificação baseada em uma gama de informações bibliométricas.

No que concerne a classificação dos periódicos segundo a CAPES, tem-se que a mesma é realizada pelas áreas de avaliação e passa por um processo de atualização anual (SIC - Serviço de Informação ao Cidadão). Os periódicos são enquadrados em estratos indicativos de qualidade. O periódico deve ter fator de impacto medido pelo **Institute for Scientific Information (ISI)**. As classificações dos periódicos em ordem decrescente de valor, são dadas por: A1, o mais elevado; A2; B1; B2; B3; B4; B5; C (com peso zero). No entanto, o mesmo periódico, ao ser classificado em duas ou mais áreas distintas, pode receber diferentes avaliações.

No comitê avaliador de Matemática e Estatística, no ano de 2012 instituiu-se o uso da classificação *Core* para designar os periódicos em que a área prioritária de desenvolvimento da pesquisa é de domínio Matemático/Estatístico.

A classificação Qualis/CAPES divulgada em 2012 é composta de 1189 periódicos e contém 9 colunas contendo informações sobre o ISSN das revistas, classificação Qualis dos jornais, entre outras. A Tabela 2.1 é apenas ilustrativa da representação tabular dos dados. Utilizou-se esta base de dados da CAPES durante todo o desenvolvimento deste trabalho.

O banco de dados Qualis 2012 é formado pelas seguintes variáveis:

Tabela 2.1: Representação tabular parcial dos dados coletados na Classificação de Periódicos no Qualis/CAPES, ano base de 2012.

ISSN	Título	Qualis 2012	...	Core
1085-3375	Abstract and Applied Analysis	B5	...	Matematica Aplicada
1525-9293	ACIS International Journal of Computer and Information S	B5	...	Outra area
0360-0300	ACM Computing Surveys	A2	...	Outra area
1084-4309	ACM Transactions on Design Automation of Electronic Sy	B1	...	Outra area
0730-0301	ACM Transactions on Graphics	A1	...	Matematica
0098-3500	ACM Transactions on Mathematical Software	A1	...	Matematica Aplicada
0167-8019	Acta Applicandae Mathematicae	B2	...	Matematica
0065-1036	Acta Arithmetica	A2	...	Matematica
⋮	⋮	⋮	⋮	⋮
0020-7748	International Journal of Theoretical Physics	B5	...	Outra area

Fonte: Banco de dados obtido da Classificação de Periódicos no Qualis/CAPES, ano base de 2012.

ISSN: Número Internacional Normalizado para Publicações Seriadas (International Standard Serial Number). É o identificador de publicações seriadas aceito internacionalmente. Tem como função identificar unicamente o título de uma publicação seriada. Por exemplo, jornais, revistas, anuários, etc, em circulação, futuras e encerradas, também;

Título: Nome da revista ou jornal em que os periódicos foram publicados;

Qualis 2010: Avaliação dos periódicos divulgada pela CAPES em 2010. Quanto aos tipos de extratos: A1, o mais elevado; A2; B1; B2; B3; B4; B5; C, com peso zero;

Qualis 2012: Avaliação dos periódicos divulgada pela CAPES em 2012. Quanto aos tipos de extratos: A1, o mais elevado; A2; B1; B2; B3; B4; B5; C, com peso zero;

AIS: Escore de Influência de um artigo (Article Influence Score - AIS). Determina a influência média de artigos de um periódico durante os primeiros cinco anos após a publicação. Esta medida é calculada dividindo-se o *Eigenfactor Score* (vide Seção 2.2), pelo número de artigos da revista, normalizado como uma fração de todos os artigos em todas as publicações. Esta medida é aproximadamente análoga ao Fator de Impacto Jornal de 5 anos, que será definido nas seções 2.2 e 2.3;

Meia vida: O principal aspecto da obsolescência da literatura é a meia-vida, do inglês Half-Life, das citações. O termo “meia-vida das citações” teve origem no conceito físico de meia-vida dos materiais radioativos. O conceito foi transposto para a área de ciência da informação, pela primeira vez, por [Burton e Kebler \(1960\)](#), para expressar o período em que uma pesquisa alcança a metade de sua vida útil. Em 1970, [Brookes](#) idealizou um método alternativo de medida da meia-vida dos periódicos científicos, que foi logo adotado, sem qualquer ajuste, pelo ISI. O novo critério estabelece que “a meia-vida é o tempo (em anos) para que 50% das citações recebidas por um periódico apareçam na literatura (JCR, 1998)”.

Core: Identifica a natureza das áreas dos periódicos, especialmente com ênfase em Matemática, Matemática Aplicada e Estatística. Devido às especificidades das sub-áreas foi decidido manter a distinção entre os periódicos intra área (denominadas revistas “core”) e publicações em revistas prevalentes em outras áreas (denominadas revistas “não core”).

Na sequência, apresenta-se algumas sugestões de acesso aos mecanismos de busca do Qualis da CAPES. Vejamos:

- (i) busca por título do periódico - no site do WebQualis (<http://qualis.capes.gov.br/webqualis>), pode-se realizar a pesquisa pelo título do periódico: Consultar ⇒ Periódico ⇒ Por título do periódico. Vale a pena lembrar que o mesmo periódico pode ter classificações diferentes em áreas diferentes;
- (ii) criação de relatórios - para criar um relatório de interesse em uma planilha do excel com todas as publicações da área e suas respectivas estratificações deve-se seguir a sequência: Lista Completa ⇒ Periódico ⇒ Selecione a área desejada ⇒ Gerar relatório.

Na Figura 2.1 apresenta-se a visualização da página na internet ao acessar a WebQualis, utilizando o seguinte link <http://qualis.capes.gov.br/webqualis>

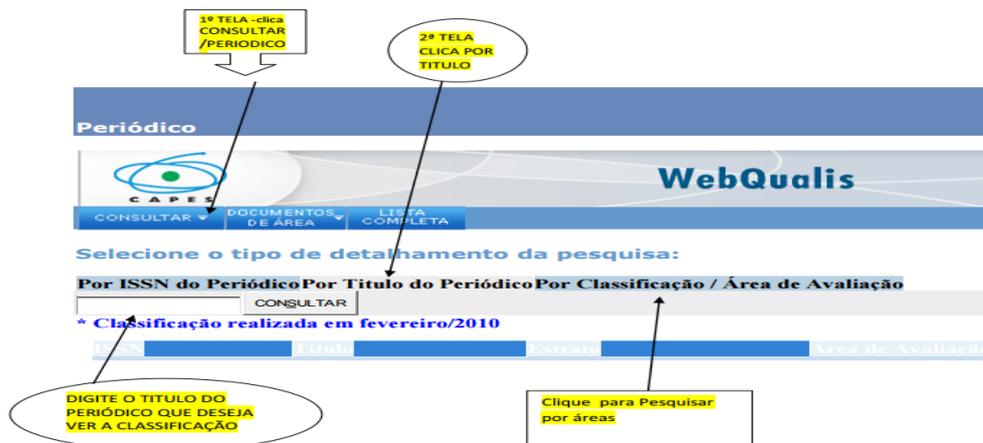


Figura 2.1: Visualização da página na internet ao acessar o endereço eletrônico da WebQualis. Último acesso em: 24 de fevereiro de 2016.

2.2 Web Of Science

O Web of Science é uma base de dados que disponibiliza acesso a mais de 55 milhões de registros dos principais periódicos, possibilitando a identificação de artigos de periódicos em diversas áreas do conhecimento.

Essa base de dados pode ser utilizada para uma pesquisa de alta qualidade, associando registros, buscando por autores na área de interesse, por referência citada, recebendo alertas de publicação para assuntos, publicações ou autores, processando os resultados de uma busca e salvando-os ou enviando-os por e-mail, salvando o histórico de buscas, entre outras possibilidades.

A título de informação, o Web of Science não relaciona termos, só busca o que for digitado. As possibilidades de pesquisa são bastante ampliadas com o uso dos símbolos de truncagem, operadores booleanos e operadores de proximidade. A seguir, tem-se as definições para essas possibilidades

- Os símbolos de truncagem são usados para encontrar palavras no singular ou plural, e variações de escrita das palavras;
- Operadores booleanos: O operador **AND** busca por todos os termos digitados. Quando utilizamos **OR**, pelo menos um dos termos digitados tem que ser encontrado. Já o operador **NOT** exclui os termos digitados da sua pesquisa.

- Os operadores de proximidade são utilizados para pesquisar por uma frase exata, deverá colocá-la entre aspas (“ ”).

Para exemplificar os indicadores que são utilizados na base da Web Of Science, ano base de 2014, tem-se a seguir uma representação tabular completa deste banco de dados. A base de dados é formada por 11.813 linhas e 17 colunas. A Tabela 2.2 traz uma representação tabular parcial dos mesmos.

Tabela 2.2: Representação tabular parcial dos dados coletados no Web Of Science, ano base de 2014.

Rank	Full Journal Title	JCR Abbreviated Title	ISSN	...	Normalized Eigenfactor
1	Ca-A Cancer Journal For Clinicians	Ca-Cancer J Clin	0007-9235	...	7.02626
2	New England Journal Of Medicine	New Engl J Med	0028-4793	...	75.74992
3	Chemical Reviews	Chem Rev	0009-2665	...	25.08968
4	Lancet	Lancet	0140-6736	...	44.30131
5	Nature Reviews Drug Discovery	Nat Rev Drug Discov	1474-1776	...	6.73858
6	Nature Biotechnology	Nat Biotechnol	1087-0156	...	16.70407
7	Nature	Nature	0028-0836	...	167.85370
8	Annual Review Of Immunology	Annu Rev Immunol	0732-0582	...	5.10245
9	Nature Reviews Molecular Cell Biology	Nat Rev Mol Cell Bio	1471-0072	...	12.59153
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮
11770	npj Primary Care Respiratory Medicine	NPJ PRIM CARE RESP M	2055-1010	...	0.00000

Fonte: Banco de dados obtido na Web of Science, ano base de 2014.

Atualmente, o Web of Science disponibiliza, além do ISSN dos periódicos, os seguintes indicadores:

Posto: Refere-se à posição ocupada (rank) pelo periódico dentro da classificação geral realizada pela Web Of Science;

Título completo: Nome completo da revista, ou jornal, em que os periódicos foram publicados;

Título abreviado: Fornece o título abreviado de um periódico incluído na base de dados JCR;

Total de citações: Todas as referências citadas são indexadas e podem ser encontradas na Busca de referências citadas. Clique no link “Cited References” da Rede de citações para passar para a visualização de referências citadas;

Fator de impacto: O fator de impacto (Journal impact factor) é uma medida da frequência com que, em média, os artigos de um dado periódico foram citados em jornais nos últimos anos em um ano específico. Esta medida objetiva eliminar alguns dos vieses, tais como esclarecer a frequência absoluta do total de citações, que favorecem os grandes jornais em relação aos menores, ou revistas que apresentam alta ou baixa frequência de citações ao longo do tempo, e por fim, das revistas mais antigas sobre as mais novas. Por exemplo, um jornal hipotético publica 52 artigos em 2012 e 58 em 2013. Em 2014, há 100 citações aos artigos de 2012 e 120 citações aos de 2013. O fator de impacto é $(100 + 120)/(58 + 52) = 220/110 = 2$;

Fator de impacto sem as auto citações (Impact factor): Corresponde a uma métrica que calcula o fator de impacto do jornais sem considerar suas auto citações;

Fator de impacto de 5 anos (5 Year Impact Factor): Corresponde ao cálculo do fator de impacto relacionado ao total de citações de um periódico em artigos publicados nos 5 anos anteriores ao ano base. Por exemplo, se o ano base for 2014, as citações dos anos 2013 a 2009 são considerados;

Immediacy Index: Mede todas as referências que informa quão rapidamente os artigos de um dado jornal são citados. Por exemplo, um jornal publica 40 artigos em 2012, que são citados 30 vezes em 2012. Logo, $30/40 = 0.75$;

Itens citáveis: O total de “Artigos citáveis” (Citable Items) mostra a lista de citações de todos os anos e todas as edições na Coleção Principal da Web of Science, incluindo anos e edições que não são cobertos por sua assinatura.

Cited half life (CHL): Indica a idade mediana dos artigos publicados em um jornal que foram citados em um ano base do JCR. Dessa forma, metade das citações ao jornal no ano base referem-se a itens publicados durante o tempo decorrido do CHL. Por exemplo, se um jornal possui CHL de 8,0 em 2013, isso significa que 50% das citações que o jornal recebeu em 2013 foram de artigos publicados nos últimos 8 anos, isto é, de 2006 a 2013. As citações remanescentes, perfazendo 50% do total, foram citações de artigos publicados em 2005 ou anteriormente;

Citing half life (CIHL): Similar ao CHL mas referente a idade mediana dos artigos citados pelo jornal em um ano base do JCR;

Eigenfactor Score: É calculado usando o número de vezes que artigos de um jornal que foram publicados nos últimos cinco anos foram citados em um dado ano base. No entanto, as auto-citações ao próprio jornal são removidas, e o escore também considera influência dos jornais em que as citações são feitas. A soma dos escores eigenfactor de todos os jornais da base JCR é 100;

Article Influence Score (AIS): É uma medida da influência média dos artigos de um jornal cinco anos após a publicação. Os AIS’s são normalizados de modo que o valor médio de AIS dos periódicos na base JCR sejam iguais à unidade, isto é, 1,00. Um escore maior do que 1,00 indica que os artigos de um jornal têm influência acima da média e um escore menor do que 1,00 indica que os artigos daquele jornal têm influência baixa da média.

O AIS é calculado pela equação

$$AIS = \frac{0.01 \times \text{eigenfactor score}}{x},$$

em que

x = número de artigos publicados pelo jornal nos últimos 5 anos dividido pelo número de artigos publicados por todos os jornais na base JCR nos últimos 5 anos;

Percentual de artigos em itens citáveis: Informa, em termos percentuais, o montante de artigos na revista que trazem pesquisa original. Por exemplo, em 2013, dado jornal publicou 700 artigos originais e 30 artigos de revisão de teoria, em um total de 730 itens citáveis. Portanto, tal forma possui 99,9% de artigos originais entre todos os itens citáveis;

Average Journal Impact Factor Percentile: Corresponde ao cálculo da média percentual do fator de impacto sobre a publicação de uma revista;

Normalized Eigenfactor: O Eigenfactor Normalizado converte o Eigenfactor de uma revista em uma pontuação multiplicativa centrada em torno de 1, de tal forma que, se um jornal recebe uma pontuação Eigenfactor Normalizado de 2, tal jornal é considerado como sendo duas vezes mais influente que um jornal típico da base JCR.

O acesso pode ser feito de várias maneiras, recomenda-se as seguintes:

- i) Acesse a página da Biblioteca da Universidade de Brasília - UnB (<http://www-periodicos-capes-gov-br.ez54.periodicos.capes.gov.br/>), digite Web of Science;
- ii) Acesse a página da UnB (<http://www.unb.br>), clique em Biblioteca, a seguir clique em Bases de Dados e no local de busca digite Web of Science;
- iii) Acessando diretamente a página do ISI Web of Knowledge (<http://isiknowledge.com/>) e selecionando a aba Web of Science.

Na Figura 2.2 tem-se a visualização da página da internet ao acessar a Web Of Science, utilizando o seguinte link <http://apps-webofknowledge.ez54.periodicos.capes.gov.br/>

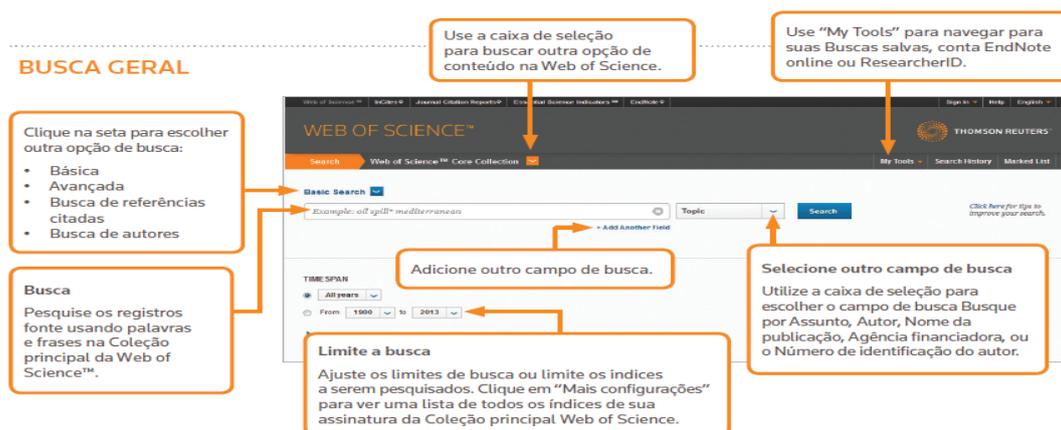


Figura 2.2: Visualização da página na internet ao acessar o endereço eletrônico da Web Of Science. Online; último acesso em: 28 de fevereiro de 2016.

Para mais detalhes sobre algumas informações específicas referentes a Figura 2.2, acesse a home page da Thomson Reuters, ip-science.thomsonreuter.com/contact.

2.3 Scimago Journal & Country Rank

O Scimago Journal & Country Rank (Scimago) é um portal que inclui as revistas e indicadores científicos dos países desenvolvidos a partir das informações contidas na base de dados *Scopus*, Elsevier. Os indicadores fornecidos pelo Scimago podem ser utilizados para avaliar e analisar os domínios científicos. Para mais detalhes, sobre este indicador, ver [Guerrero-Botea e Moya-Anegónb \(2012\)](#). Os indicadores mostram a visibilidade das revistas contidas no banco de dados *Scopus* (vide Seção 2.4).

Uma das grandes novidades apresentadas ao acessar a página da Scimago é que, além do portal do Scimago, a empresa disponibiliza uma Atlas da Ciência, cujo objetivo principal é fornecer uma representação gráfica da pesquisa Ibero-Americana. Tal apresentação é constituída por uma coleção de mapas interativos, permitindo funções de navegação nos espaços semânticos formados pelos mapas.

Para exemplificar os indicadores que são utilizados na base de dados do Scimago, no ano base de 2014, a Tabela 2.3 ilustra uma representação tabular deste banco de dados. A base de dados é formada por 1.284 linhas e 14 colunas. Nas linhas da tabela registram-se os periódicos e, nas colunas, os valores dos indicadores qualitativos/descriptivos destes periódicos.

Tabela 2.3: Representação tabular parcial dos dados coletados no Scimago Journal & Country Rank (SJR), ano base de 2014.

Rank	Title	Type	SJR	...	Country
1	IEEE Trans. on Pattern Analysis and Machine Intell.	j	8,741	...	United States
2	Molecular Systems Biology	j	8,571	...	United Kingdom
3	Annals of Mathematics	j	8,551	...	United States
4	Vital and hea. stat. Series 10, Data from the Nat. Hea. Sur.	k	7,801	...	United States
5	Foundations and Trends in Com. and Infor. Theory	j	7,602	...	United States
6	Journal of the Amer. Mathe. Society	j	6,605	...	United States
7	Acta Numerica	j	6,385	...	United Kingdom
8	Inventiones Mathematicae	j	6,38	...	United States
9	Journal of the Royal Stat. Soci.. Series B: Stat. Metho.	j	6,148	...	United Kingdom
:	:	:	:	:	:
:	:	:	:	:	:
1284	Konstruktion	j	0,1	...	Germany

Fonte: Banco de dados obtido no Scimago Journal & Country Rank, ano base de 2014.

Atualmente, o Scimago disponibiliza, além do título e ISSN do jornal, os seguintes indicadores:

Posto: Refere-se à posição ocupada pelo periódico dentro da classificação geral realizada pela Scimago;

Tipo: Corresponde aos tipos de publicações (b: livro, k: livro de séries, p:conferências e processos);

SJR: É uma medida de influência científica dos periódicos que leva em conta tanto o número de citações recebidas pelo jornal quanto a importância e o prestígio dos jornais dos quais as citações são originárias. Dessa forma, o SJR representa uma métrica de prestígio baseado na ideia de que as citações recebidas por um jornal têm pesos ou valores distintos dependendo do prestígio do jornal citante. O algoritmo utilizado no SJR é similar ao aplicado no *google pagerank*.

Índice H: O índice H expressa o número de artigos “h” de um jornal, que receberem pelo menos h citações. O índice H quantifica tanto a produtividade científica de um jornal quanto o impacto científico. Este índice também é aplicável a cientistas, países, etc;

Total Docs.: Total de publicações num período em todos os tipos de documentos (citáveis ou não) de um jornal nos três últimos anos;

Total Docs. (3 years): Total de publicações para o ano base de 2014, todos os tipos de publicações dos anos de 2011 à 2013 são considerados;

Total Refs.: Corresponde ao número total de referências de um jornal de um jornal em período;

Total Cites (3 years): Corresponde ao número de citações recebidas por um jornal nos últimos três anos;

Citable Docs. (3 years): Corresponde ao número de documentos citáveis publicados por um jornal nos últimos 3 anos;

Cites/Doc. (2 years): Equivale a média de citações por documento no período de 2 anos. A medida é calculada considerando-se o número de citações recebidas por um jornal em dado ano dividido pelo número total de documentos do jornal nos dois anos anteriores;

Ref./Doc.: Número médio de referências por documentos em um dado ano;

Country: País em que fica localizado o Jornal (Revista) da publicação.

Na sequência, apresenta-se algumas sugestões de como utilizar a página da Scimago Journal & Country Rank (SJR). Para iniciar a navegação, via web, utilizar o link <http://www.scimagojr.com/journalrank.php>.

Sugestão 1: As revistas podem ser pesquisadas por título da revista, ISSN ou editor. Nas buscas pelo ISSN, não incluir hífen. Nas buscas por título da revista ou editora do jornal, podem ser usadas palavras individuais ou palavras exatas/frase;

Sugestão 2: Os conjuntos de Jornais listados no ranking podem ser personalizados através da filtragem e da classificação por critérios e, assim, aplicar um limite mais baixo para o ranking. Cada revista pode ser analisada separadamente, basta clicar no link do título da revista para se obter um perfil individual, incluindo tabelas de séries temporais e gráficos para analisar métricas significativas de desempenho diário. Ranking e opções de filtragem incluem todas as variáveis listadas anteriormente.

Na Figura 2.3 tem-se a visualização da página na internet ao acessar o Scimago pelo link <http://www.scimagojr.com/journalrank.php>

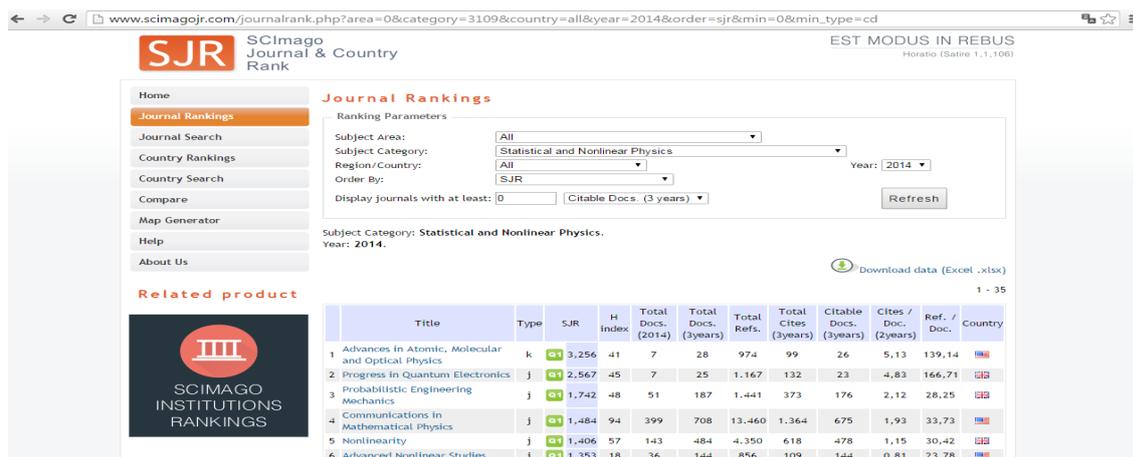


Figura 2.3: Visualização da página na internet ao acessar o endereço eletrônico da Scimago Journal & Country Rank (SJR). Online; último acesso em: 28 de fevereiro de 2016.

2.4 Journal Indicators of University's Centre for Science and Technology Studies (CWTS)

O Journal Indicators da University's Centre for Science and Technology Studies, ou apenas, CWTS, oferece uma série de indicadores bibliométricos em revistas científicas. Estes indicadores foram calculados com base no banco de dados bibliográficos *Scopus* produzido pela Elsevier.

Neste contexto, o *Scopus* é a maior base de dados de resumos e citações de literatura revisada por pares: revistas científicas, livros e anais de eventos. O CWTS oferece ferramentas inteligentes para acompanhar, analisar e visualizar pesquisas. Assim, o *Scopus* oferece uma visão ampla, a informação científica interdisciplinar global que pesquisadores, professores e alunos precisam, para se manterem informados.

Segundo Moed (2010), um indicador chave oferecido pelo CWTS é o SNIP, que significa fator de impacto normalizado por publicação. Waltman *et al.* (2013), propuseram algumas modificações na forma de calcular o SNIP. Para mais detalhes, ver Waltman *et al.*, 2013.

Segundo Waltman *et al.* (2013), o SNIP usa uma fonte normalizada para corrigir diferenças nas práticas de citação entre os campos científicos.

Comparativamente ao fator de impacto (IF), o SNIP tem a vantagem de ser calculado de modo a corrigir por diferenças em práticas ou padrões de citação em diferentes áreas do conhecimento. Dessa forma, o SNIP proporciona maior acurácia na comparabilidade entre áreas no impacto da citação. Sendo uma medida corrigida por área, é especialmente útil para pesquisadores que publicam em áreas interdisciplinares.

Objetivando exemplificar os itens abordados nesta seção considerou-se o banco de dados CWTS, no ano base de 2014. Como a representação tabular completa deste banco de dados possui 1.000 linhas e 4 colunas, a Tabela 2.4 é apenas uma representação tabular parcial dos mesmos.

Tabela 2.4: *Representação tabular parcial dos dados coletados no Journal Indicators da University's Centre for Science and Technology Studies (CWTS), ano base de 2014.*

Rank	Title	P	SNIP
809	Abstract and Applied Analysis	2450	0.62
3	ACM Computing Surveys	140	8.48
98	ACM Transactions on Algorithms	117	1.91
547	ACM Transactions on Applied Perception	65	0.89
82	ACM Transactions on Computational Logic	89	2.00
1	ACM Transactions on Intelligent Systems and Technology	190	10.61
12	ACM Transactions on Mathematical Software	75	3.99
97	ACM Transactions on Modeling and Computer Simulation	61	1.92
546	Acta Applicandae Mathematicae	240	0.89
⋮	⋮	⋮	⋮
888	Zeitschrift fur Naturforschung-Section A Journal of Physical Sciences	252	0.54

Fonte: Banco de dados obtido da CWTS Jornal Indicadores, ano base de 2014.

Além do SNIP, o CWTS oferece três indicadores, o posto (rank), P e IPP:

Posto: Refere-se à posição ocupada pelo periódico dentro da classificação geral realizada pela CWTS;

Título: Nome da revista ou jornal em que os periódicos foram publicados;

P: O número de publicações de uma fonte, nos últimos três anos;

IPP: Impacto por publicação, mede a razão entre o número de citações em um dado ano (A) dos artigos publicados de um jornal nos três anos anteriores (A-1, A-2, A-3) e o número total de artigos publicados desse jornal nos mesmos anos A-1, A-2 e A-3.

A seguir, apresenta-se algumas sugestões relativas aos mecanismos de busca do CWTS. Vejamos:

- (i) no site do CWTS (www.journalindicators.com), clique na opção “Indicators”. Para ter acesso às informações sobre os periódicos, basta seguir a seguinte ordem de cliques e preenchimento das devidas informações: Select subject area (Main area ou Subarea) ⇒ Select sources and sort order (Search for ou Order by) ⇒ Advanced parameters (Year ou Source type ou Minimum number of publications);
- (ii) Todas as estatísticas fornecidas pela CWTS podem ser baixados em um arquivo de Excel. Para cada combinação de uma fonte e de um ano, o ficheiro Excel fornece os valores dos indicadores correspondentes, no caso de P e SNIP.

O site do CWTS ainda fornece os intervalos de estabilidade, que informam sobre a estabilidade ou a confiabilidade de um indicador.

Na Figura 2.4 ilustra-se a visualização da página na internet ao acessar o CWTS, via link www.journalindicators.com



Figura 2.4: Visualização da página na internet ao acessar o endereço eletrônico da CWTS Jornal indicadores. Online; último acesso em: 26 de fevereiro de 2016.

A título de informação, como já mencionado anteriormente no Capítulo 1, ressalta-se que a www.journalindicators.com (CWTS) não será considerada nesta pesquisa.

Após a apresentação dos bancos de dados serem analisados, é necessário realizar a identificação e seleção das variáveis para a ordenação e classificação dos periódicos científicos dentro de cada uma dessas bases.

2.5 Identificação e seleção das variáveis para a ordenação dos periódicos científicos

Complementando o exposto, fundamentando-se nas informações fornecida pelos sites Web-Qualis <http://qualis.capes.gov.br/webqualis> (Qualis/CAPES), <http://apps-webofknowledge.ez54.periodicos.capes.gov.br/> (Web Of Science) e <http://www.scimagojr.com/journalrank.php> (Scimago), na Tabela 2.5 apresentam-se as variáveis que compõem as bases de dados citadas anteriormente. Na Tabela 2.5 “S” indica a presença da referida variável e “N” a ausência da mesma.

Tabela 2.5: Representação das variáveis que compõem as bases de dados Qualis/CAPES, Web Of Science e Scimago, desta pesquisa.

Variável	Base CAPES (1189)	Web Of Science	Scimago
Título do Jornal	S	S	S
ISSN	S	S	S
Classificação Qualis 2012	S	N	N
AIS	S	S	N
Half Life	S	S	N
Área (Mat., Mat. Apl. Est. Out. área)	S	N	N
Total Cites	N	S	N
Journal Impact Factor	S	S	N
Impact Factor Journal Self Cites	N	S	N
5 Year Impact Factor	N	S	N
Immediacy Index	N	S	N
Citable Items	N	S	N
Citing Half life	N	S	N
Eigenfactor Score	N	S	N
Perc Articles Citable Items	N	S	N
Average Journal Impact Factor (%)	N	S	N
Normalized Eigenfactor	N	S	N
Type	N	N	S
SJR	N	N	S
H index	N	N	S
Total Docs. (2014)	N	N	S
Total Docs. (3years)	N	N	S
Total Refs.	N	N	S
Total Cites (3years)	N	N	S
Citable Docs. (3years)	N	N	S
Cites/Doc. (2years)	N	N	S
Ref./Doc.	N	N	S
Country	N	N	S

Fonte: Elaborado pelos autores (2017).

Conforme pode ser visualizado na Tabela 2.5, algumas bases de dados apresentam variáveis em comum, enfatizando-se que, o *Título do Jornal* e o valor do *ISSN* estão presentes em todas as bases.

Na tabela acima, as variáveis em destaque correspondem às que foram consideradas nas análises iniciais da pesquisa. As análises iniciais correspondem a Etapa I que será devidamente apresentada no Capítulo 6.

No **Apêndice K**, apresentam-se os pareamentos que foram realizados nas bases de dados supra citadas, bem como algumas medidas de similaridade utilizadas nesta pesquisa, necessárias à construção de uma base de dados comum. Por fim, o resultado de tais pareamentos.

2.6 Ano-base em estudo

Os dados prioritários a serem trabalhados nesta dissertação referem-se ao ano-base de 2012. Embora existam arquivos de anos mais recentes disponíveis no WebQualis (<http://qualis.capes.gov.br/webqualis>), tais arquivos não trazem informações sobre a área de concentração principal (área) de um dado periódico, isto é, não há informação sobre se determinado periódico incluso, em uma dada lista é de Matemática, Matemática Aplicada, Estatística ou de Outra Área. Somente um documento de 2012, disponível somente em PDF, traz esta informação. Esse arquivo necessitou ser transformado em arquivo EXCEL.

A informação sobre a área é imprescindível para que se possa proceder a análises de dados que incorporem e tratem conhecidos efeitos sistemáticos nos indicadores bibliométricos, que são devidos a área de concentração dos periódicos a que estes pertencem. Por exemplo, os fatores de impacto de periódicos em áreas de genética, medicina, entre outras, são inegavelmente superiores àqueles observados nas áreas de Matemática. No entanto, também observa-se considerável heterogeneidade nos fatores de impacto das revistas científicas

inclusas na grande área de Matemática, observando as sub-áreas de Matemática Pura, Matemática Computacional e Estatística.

Nesta dissertação a informação sobre a *área* ao qual um dado periódico pertence (Matemática Pura, Matemática Computacional, Estatística ou Outra Área) foi utilizada na tentativa de obter-se escalas universais (isentas de efeitos sistemáticos devidos a área) de modo a permitir a construção de uma métrica única para ordenar os periódicos. Nos Capítulos 3, 4 e 5 são descritos aspectos teóricos da Teoria da Resposta ao Item.

Capítulo 3

Teoria da Resposta ao Item (TRI)

Na literatura, encontram-se vários conceitos sobre Teoria da Resposta ao Item (TRI). Dentre eles, destaca-se o apresentado por Lord *et al.* (1968), “[...] a Teoria da Resposta ao Item é um conjunto de modelos estatísticos que permitem mensurar um traço latente através de um conjunto de respostas a itens de um instrumento de avaliação ou de pesquisa.”

No que concerne mensurar aos traços latentes, Araujo *et al.* (2009), ressaltam que a busca por informações acerca de características psicológicas de indivíduos levou muitos pesquisadores a desenvolver modelos que pudessem estimá-las. Tais propriedades psicológicas, também referidas por traços latentes, são características individuais que não podem ser observadas diretamente, tais como: proficiência em determinado conteúdo na avaliação educacional, atitude em relação à mudança organizacional, nível de estresse, satisfação do consumidor, nível de depressão, nível de satisfação, qualidade de vida, capacidade de empatia, propensão à ansiedade, propensão à depressão etc..

Neste Capítulo objetiva-se introduzir os principais conceitos, modelos e resultados que podem ser obtidos a partir da aplicação da TRI. Os modelos de TRI para dados dicotômicos e politômicos são de especial interesse nesta dissertação. Desse modo, tais modelos são descritos com detalhes nas próximas seções. Dessa forma, a descrição inclui o modelo logístico de dois parâmetros para itens dicotômicos e o Modelo de Resposta Gradual (Graded Response Model), proposto por Samejima em 1969, para itens politômicos. Na Seção 3.1 é apresentada uma breve história da TRI e suas aplicações em diversas áreas do conhecimento científico.

3.1 Introdução

Os primeiros modelos para variáveis latentes foram apresentados nos estudos de Lawley (1943), Guttman *et al.* (1950) e Lazarsfeld *et al.* (1950), marcando, assim, o início da TRI.

Segundo Andrade *et al.* (2000), o interesse pelo estudo de modelos de TRI vem crescendo em várias áreas do conhecimento, particularmente em avaliação educacional. A Teoria da Resposta ao Item inclui modelos para a estimação de traços latentes, ou seja, características do indivíduo que não podem ser observadas diretamente. Esse tipo de variável deve ser inferida a partir da observação de variáveis secundárias, que estejam relacionadas à variável latente. A metodologia de TRI descreve formas funcionais de representar a relação entre a probabilidade de acerto de um item e determinados níveis de traços latentes, proficiências ou habilidades de um indivíduo na área de conhecimento avaliada.

Inicialmente, é pertinente apresentar um histórico da TRI. Os primeiros modelos de resposta ao item surgiram em meados da década de 50, com as publicações de Lord (1952), envolvendo modelos em que se considerava uma única habilidade, de um único grupo de indivíduos, sendo medida por um teste construído com itens dicotômicos. Estes modelos

foram primeiramente desenvolvidos de modo a representar a probabilidade de acerto, a um item dicotômico, em termos da função de probabilidade, isto é, com curvas expressa na forma de uma ogiva normal. Posteriormente, foram descritos numa forma matemática mais conveniente, e que vem sendo usada até então: a logística. As publicações de Lord também contribuíram para o desenvolvimento de programas computacionais.

Lord (1952), foi o primeiro a apresentar o modelo unidimensional de 2 parâmetros e, na sequência, desenvolveu o modelo de 3 parâmetros (vide equação 3.2), considerando a questão do acerto casual Rasch (1960), sugeriu o modelo unidimensional de 1 parâmetro com a função de ogiva normal. Birnbaum (1968), propôs a substituição, nos modelos de 2 e 3 parâmetros, da função ogiva normal pela função logística. Wright (1968), sugeriu a substituição, no modelo de Rasch, da função normal pela função logística.

No que diz respeito aos modelos para itens não dicotômicos, Samejima (1969), apresentou o Modelo de Resposta Gradual (Graded Response Model), no qual assume-se que as categorias de resposta de um item podem ser ordenadas entre si. Como exemplo, tem-se as pesquisas de marketing, visto os itens que compõem o instrumento são, em geral, caracterizadas por um conjunto ordenado de respostas, e não pela resposta binária (0 e 1), sim ou não, concorda ou discorda etc, o que caracteriza o modelo dicotômico. Bock (1972) propôs o Modelo de Resposta Nominal (Nominal Categories Model), no qual as categorias de resposta não são, necessariamente, ordenadas. Por exemplo, tem-se os testes de múltipla escolha, em que as categorias não são ordenadas. Muraki (1992), sugeriu uma modificação no modelo de resposta gradual proposto por Samejima. Andrich (1978), apresentou o Modelo de Escala Gradual (Rating Scale Model) que, a princípio, seria um caso particular do modelo de resposta gradual de Samejima. No entanto, Andrich fez uma suposição a mais: a de que os escores das categorias são igualmente espaçados.

3.1.1 A Teoria Clássica de Medidas (TCM)

Na Teoria Clássica dos Testes (TCT), também conhecida como Teoria Clássica das Medidas (TCM), são avaliadas as propriedades psicométricas dos itens (questões), especificamente, a confiabilidade (consistência interna da escala), a análise do item e a validade do construto.

Lord *et al.* (1968) enfatizam que na TCM, os parâmetros/características observados num teste são compostos de um escore verdadeiro mais o erro de medida, este último assumido como aleatório. A equação básica conhecida como modelo clássico de medida, descreve a relação entre os escores observados, os escores verdadeiros e o erro:

$$X_{ij} = T_i + \epsilon_{ij}, \quad (3.1)$$

em que

X_{ij} : é o escore observado para o indivíduo i no teste j ,

T_i : é o escore verdadeiro para o indivíduo i no teste j ,

ϵ_{ij} : é o erro de medida para o indivíduo i no teste j .

Assume-se que $E(\epsilon_{ij}) = 0$, $Cov(T_i, \epsilon_{ij}) = 0$, $Cov(\epsilon_{ij}, \epsilon_{ik}) = 0$ e $Cov(T_i, \epsilon_{kj}) = 0$ para $k \neq i$. Portanto, $Var(X_{ij}) = Var(T_i) + Var(\epsilon_{ij})$, $Var(T_i) = 0$.

Vale ressaltar que o índice j pode indexar formas de testes paralelos ou tempos de administração do mesmo teste. Duas formas são ditas paralelas se elas possuem escores verdadeiros idênticos e se seus erros têm variâncias iguais.

A fidedignidade de um teste pode ser definida por três modos, que são equivalentes no modelo clássico: pela razão da variância dos escores verdadeiros em relação a variância dos escores observados, pela correlação dos escores observados nas duas formas paralelas de um teste ou pelo quadrado da correlação entre os escores observados e os verdadeiros (Bechger *et al.*, 2003).

Lord *et al.* (1968), Bechger *et al.* (2003), van Belle e Arnold (2000) e Hambleton e Jones (1993) enfatizam que a TCM, quando comparada à TRI, apresenta uma série de limitações teóricas, como:

- (i) Os parâmetros clássicos dos itens (dificuldade e discriminação) dependem diretamente da amostra selecionada, e se a mesma não for representativa da população, os parâmetros dos itens não podem ser considerados válidos para esta população;
- (ii) A avaliação de aptidões ou habilidades também depende do teste utilizado. Assim, testes diferentes que medem a mesma aptidão irão produzir escores diferentes da mesma aptidão para sujeitos idênticos. O mesmo acontece com testes com índices de dificuldade diferentes;
- (iii) Outro problema da teoria clássica é sua orientação para o teste total e não para o item individual. Toda a informação do item deriva de considerações do teste geral, não podendo, assim, determinar como o examinando se comportaria diante de cada item individual, como acontece na TRI, onde o importante é o item e não o escore total da prova ou teste.
- (iv) Os métodos clássicos de análise necessitam de dados completos ou técnicas de imputação para dados faltantes, ao passo que na TRI isso não é necessário.

Neste contexto, vale ressaltar que a TRI foi desenvolvida, principalmente, para suprir limitações da Teoria Clássica de Medidas (TCM). Uma das limitações da TCM é que o instrumento de medida depende das características dos respondentes que se submetem ao teste ou ao questionário (Hambleton *et al.*, 1991).

De acordo com Araujo *et al.* (2009), a TRI surgiu como uma forma de considerar cada item particularmente, sem relevar os escores totais. Portanto, as conclusões não dependem do teste ou questionário como um todo, como é feito na TCM, mas sim de cada item que o compõe.

Andrade *et al.* (2000), ressaltam que, inicialmente, a estimação era feita através do método da máxima verossimilhança conjunta, que envolve um número muito grande de parâmetros a serem estimados simultaneamente, causando grandes problemas computacionais.

Para solucionar tais problemas, Bock e Lieberman (1970) introduziram o método da máxima verossimilhança marginal para a estimação dos parâmetros, que consiste em

Etapa 1: inicialmente, estimam-se os parâmetros dos itens, assumindo-se uma certa distribuição para as habilidades;

Etapa 2: assumindo os parâmetros dos itens conhecidos, estimam-se as habilidades.

Apesar do avanço que esse método trouxe para o problema, ele requer que todos os parâmetros dos itens sejam estimados simultaneamente.

Neste contexto, Bock e Aitkin (1981) propuseram uma modificação no método acima, utilizando o algoritmo EM de Dempster, Dempster *et al.* (1977), de modo a permitir que os itens pudessem ter seus parâmetros estimados em separado, facilitando, em muito, o

aspecto computacional do processo de estimação. Mais recentemente, métodos bayesianos foram propostos para, entre outras coisas, resolver o problema de estimação dos parâmetros dos itens respondidos corretamente ou incorretamente por todos os respondentes e, também, o problema da estimação das habilidades dos respondentes que acertaram ou erraram todos os itens da prova, (Andrade *et al.*, 2000).

No que concerne às vantagens da Teoria de Resposta ao Item sobre a Teoria Clássica de Medidas destacam os seguintes aspectos:

- i) a TRI possibilita fazer comparações entre traço latentes de indivíduos de populações diferentes quando são submetidos a testes ou questionários que tenham alguns itens comuns e permite, ainda, a comparação de indivíduos da mesma população submetidos a testes totalmente diferentes. Isto é possível porque a TRI tem como elementos centrais os itens e não o teste ou questionário como um todo, (Andrade *et al.*, 2000);
- ii) possibilita uma melhor análise de cada item que forma o instrumento de medida, pois leva em consideração suas características específicas de construção de escalas, uma vez que os itens e os indivíduos estão na mesma escala, assim o nível de uma característica que um indivíduo possui pode ser comparado ao nível da característica exigida pelo item; isso facilita a interpretação da escala gerada e permite também conhecer quais itens estão produzindo informação ao longo da escala, (Embretson e Reise, 2000);
- iii) a TRI permite tratamento para um conjunto de dados faltantes, utilizando para isso somente os dados respondidos, o que não pode acontecer na Teoria Clássica de Medidas. Outro benefício da TRI é o princípio da invariância, isto é, os parâmetros dos itens não dependem do traço latente do respondente e os parâmetros dos indivíduos não dependem dos itens apresentados (Hambleton *et al.*, 1991).

A partir das respostas obtidas de respondentes a um conjunto de itens, a TRI possibilita a estimação dos parâmetros dos itens e dos indivíduos em uma escala de medida. Por exemplo, considere-se que a variável latente de interesse é o nível de satisfação: uma análise feita por meio da TRI pode estimar o nível de satisfação do respondente (um parâmetro do indivíduo) e, também, os parâmetros dos itens, de modo a criar uma escala de medida do nível de satisfação (Bortolotti *et al.*, 2012).

Descreve-se, a seguir, o conceito de Curva Característica do Item.

3.2 Interpretação dos parâmetros e da Curva Característica do Item (CCI)

Embretson e Reise (2000), descrevem que os modelos utilizados na TRI requerem dois elementos imprescindíveis: a *Curva Característica do Item* - CCI, pois há uma forma específica para cada mecanismo do processo de resposta utilizado, e os pressupostos da *independência local* e à unidimensionalidade. A forma de uma CCI descreve como a mudança do traço latente relaciona-se com a mudança na probabilidade de uma resposta específica.

Complementarmente ao exposto, tem-se que no gráfico de CCI produzido no modelo politômico as curvas geradas são denominadas Curvas de Categoria de Resposta (CCR). A Figura 3.1 representa, graficamente, as quatro curvas diferentes geradas por um item com quatro possibilidades de resposta, cada curva representa uma alternativa. No eixo horizontal são representados os valores do traço latente e no eixo vertical, a probabilidade de escolha das alternativas, que varia de 0 a 1.

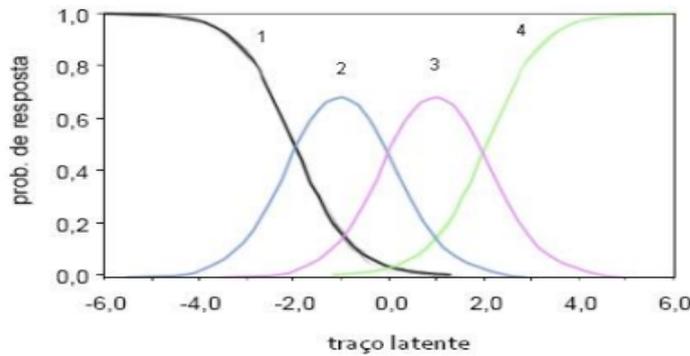


Figura 3.1: Exemplo politômico de uma Curva Característica do Item - CCI. Adaptado de Turik (2010).

Interpreta-se pela leitura do gráfico, que respondentes com nível de traço latente até -2 possuem uma maior probabilidade de optar pela primeira categoria de resposta. Respondentes com nível de traço latente entre -2 e 0 têm maior probabilidade de responderem à segunda categoria de resposta. Já respondentes com nível de traço latente entre 0 e 2, tem maior probabilidade de escolher a categoria três. Por conseguinte, para respondentes com nível de traço latente acima de 2, a maior probabilidade é que escolham a última categoria de resposta (Turik, 2010).

No que concerne a Figura 3.1, Nunes *et al.* (2008), enfatizam que pode-se perceber uma associação entre o acréscimo do traço latente e correspondente incremento na probabilidade de escolha de respostas com pontuações mais altas na escala, como se cada resposta indicasse um passo mais adiante na escala subjacente. Demonstrando, dessa forma, que à medida que se avança no nível do traço latente medido, maior a probabilidade sucessiva que pontuações mais altas sejam escolhidas.

No que concerne *independência local* e à unidimensionalidade, Embretson e Reise (2000), enfatizam que a independência local é obtida quando, controlados pelo nível do traço latente, os itens do teste são independentes. Dessa forma, para o respondente s e os itens i e j , a independência condicional local assegura que $P(U_{is} = 1, U_{js} = 1 | \theta_s) = P(U_{is} = 1 | \theta_s)P(U_{js} = 1 | \theta_s)$. Assim a probabilidade de responder um item é precisamente determinada pelo nível do traço latente do respondente e não por suas respostas a outros itens do conjunto.

Corroborando com os autores supra citados, Lord *et al.* (1968), enfatiza que a independência local é vista como consequência da correta determinação da dimensionalidade dos dados.

Lord *et al.* (1968) por Embretson e Reise (2000) esclarecem que a dimensionalidade consiste no número de fatores responsáveis para expressar o traço latente. A dimensionalidade pode ser verificada através de uma Análise Fatorial apropriada para dados categorizados.

No caso de item politômicos com categorias ordenadas, como ilustrados pela Figura 3.1, vale ressaltar que os modelos acumulativos da TRI são modelos em que a probabilidade de um indivíduo selecionar categorias mais altas de um item aumenta com o aumento do traço latente, isto é, níveis maiores de traço latente conduzem a valores mais altos de probabilidade de resposta a categorias apresentando um comportamento monotônico na CCI (Araujo *et al.*, 2009).

Por fim, no que tange à aplicação da Teoria da Resposta ao Item, Andrade *et al.* (2000), destacam que a primeira aplicação da TRI no Brasil, na área de educação, foi no ano de 1995, para a avaliação dos dados do Sistema Nacional de Ensino Básico - SAEB. Os autores destacam, também, que a TRI vem se configurando como a técnica preponderantemente

utilizada em diversos campos do conhecimento, sendo que a ênfase na aplicação desta teoria está na área de avaliação educacional.

Apresenta-se, a seguir, os principais conceitos e suposições da TRI, que se refere à unidimensionalidade, a independência local e à escala de habilidade.

3.3 Conceitos e Pressupostos da TRI: unidimensionalidade, independência local e escala das habilidades

Segundo [Andrade et al. \(2000\)](#), o pressuposto da unidimensionalidade do teste, indica a homogeneidade do conjunto de itens que, supostamente, estão medindo um *único* traço latente. Em outras palavras, nos modelos unidimensionais deve haver apenas uma habilidade responsável pela realização de todos os itens da prova.

Em se tratando da escala das habilidades, segundo [Andrade et al. \(2000\)](#), está teoricamente, pode assumir qualquer valor real entre $-\infty$ e $+\infty$. Assim, precisa-se estabelecer uma *origem* e uma *unidade de medida* para a definição da mesma. Esses valores são escolhidos de modo a representar, respectivamente, o valor médio e o desvio-padrão das habilidades dos indivíduos da população em estudo.

Segundo [Andrade et al. \(2000\)](#) e [Embretson e Reise \(2000\)](#), os modelos de TRI devem atender a dois pressupostos: a unidimensionalidade e a independência local ou independência condicional. Objetivando satisfazer a primeira suposição, é fundamental que os itens sejam homogêneos e mensurem somente um traço latente, ou um único desempenho; ou seja, é necessário que os itens de um teste/instrumento caracterizem um fator predominante. Em se tratando da segunda suposição - a independência local - como já relatado, é imprescindível que as respostas aos itens de um instrumento, sejam independentes. Tal suposição é fundamental para o processo de estimação dos parâmetros do modelo, destacam os mesmos autores.

A TRI oferece modelos para representar os traços latentes e meios de representar a relação entre o desempenho de uma pessoa em um determinado campo de conhecimento e a probabilidade desta atribuir uma resposta correta a um certo item, ([Andrade et al., 2000](#)).

Os mesmos autores apontam que as vantagens da utilização da TRI dependendo, fundamentalmente, da adequação (ajuste) dos modelos e seus pressupostos. Por exemplo, somente a partir de modelos com bom ajuste é que pode-se garantir a obtenção de itens e habilidades invariantes.

Nesta mesma linha, tem-se que a probabilidade de um indivíduo responder corretamente a um certo item é sempre a mesma, independentemente da escala utilizada para medir a sua habilidade, ou ainda, a habilidade de um indivíduo é *invariante* à escala de medida ([Andrade et al., 2000](#)).

Na sequência, apresenta-se os principais Modelos da Teoria da Resposta ao Item.

3.4 Principais Modelos da teoria da resposta ao item

Segundo [Andrade et al. \(2000\)](#), a TRI é um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo atribuir certa resposta a um item como função dos parâmetros do item e da habilidade (ou habilidades) do respondente. Essa relação é sempre expressa de tal forma que quanto maior a habilidade, maior a probabilidade de acerto no item. Os vários modelos propostos na literatura dependem, fundamentalmente, de três fatores:

- i) da natureza do item: dicotômicos ou não dicotômicos;
- ii) do número de populações envolvidas: apenas uma ou mais de uma;
- iii) da quantidade de traços latentes que está sendo medida: apenas um ou mais de um (unidimensionais ou multidimensionais).

Considerando o exposto, discutem-se os modelos unidimensionais por se tratar do escopo e foco deste trabalho e pela amplitude do tema em questão. Uma discussão dos modelos multidimensionais, bem como de suas características e aplicações é apresentada em Embretson e Reise (2000).

Na sequência, tem-se a Seção 3.5, na qual apresenta-se descrições detalhadas dos itens dicotômicos e politômicos da TRI, em função da natureza do item, fundamentando-se na abordagem de Andrade *et al.* (2000), isto é, no que concerne aos modelos que envolvem um único grupo.

Nas Seções 3.5.1 e 3.5.2, apresentam-se uma análise dos modelos para itens dicotômicos e não dicotômicos, respectivamente, propostos pelos autores mencionados.

3.5 Modelos envolvendo um único grupo

Nesta seção, será utilizada com bastante frequência, o conceito de grupo. Neste contexto, Andrade *et al.* (2000), ressaltam que, o conceito de grupo único de respondentes refere-se a uma amostra de indivíduos retirada de uma mesma população. Por exemplo, população de alunos de escolas públicas que cursam o 3º ano do ensino fundamental.

3.5.1 Modelos para itens dicotômicos ou dicotomizados

Nesta subseção, apresenta-se os modelos que podem ser utilizados tanto para itens dicotômicos (corrigidos como certo ou errado) quanto para os caracterizados como dicotomizados. Os modelos avaliados de forma dicotomizada consistem na análise de itens abertos, isto é, de resposta livre, mas expressos como certo = 1 e errado = 0.

De acordo com Andrade *et al.* (2000), os modelos logísticos para itens dicotômicos são os modelos de resposta ao item mais utilizados, sendo que há, basicamente, três tipos, que se diferenciam pelo número de parâmetros utilizados para descrever o item. Estes são conhecidos como os modelos logísticos de 1, 2 e 3 parâmetros, que consideram, respectivamente:

- (i) somente a *dificuldade* do item;
- (ii) a *dificuldade* e a *discriminação*;
- (iii) a *dificuldade*, a *discriminação* e a probabilidade de resposta correta dada por indivíduos de baixa habilidade (*probabilidade de acerto ao acaso*).

Os parâmetros de discriminação, dificuldade e de acerto casual serão descritos na Seção 4.5.1.1.

Nos modelos dicotômicos, as respostas dadas aos itens devem ser restritas a duas opções, ou as mesmas devem ser dicotomizadas posteriormente para fins de análise. Estes modelos se diferenciam pelo número de parâmetros analisados: os modelos de um parâmetro consideram a dificuldade do item (Rasch, 1960; Wright, 1977), os modelos de dois parâmetros consideram a dificuldade e a discriminação do item (Lord, 1952; Birnbaum, 1968) e os modelos de três parâmetros consideram, além destas, a probabilidade de acerto casual (Lord, 1980).

Neste contexto, tem-se a seguir, a classificação dos modelo logístico unidimensional de 3 parâmetros (ML3), modelo logístico de 2 parâmetros (ML2) e de modelo logístico de 1 parâmetro (ML1), que representam, respectivamente: a dificuldade do item, a discriminação e a probabilidade de acerto ao acaso.

3.5.1.1 O modelo logístico de 3 parâmetros (ML3)

Definição 1: Dos modelos propostos pela TRI, o *modelo logístico unidimensional de 3 parâmetros (ML3)* é dado por (Lord, 1980):

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + \exp\{-Da_i(\theta_j - b_i)\}}, \quad (3.2)$$

com $i = 1, 2, \dots, I$, e $j = 1, 2, \dots, n$ em que

U_{ij} : é uma variável dicotômica que assume os valores 1 (um), quando o indivíduo, isto é, a unidade pesquisada j responde acertadamente o item i , ou assume 0 (zero) quando o indivíduo j não responde corretamente ao item i ;

θ_j : representa a habilidade (traço latente) do j -ésimo indivíduo;

$P(U_{ij} = 1|\theta_j)$: é a probabilidade de um indivíduo j com habilidade θ_j responder corretamente o item i e é chamada de Função de Resposta do Item - FRI;

b_i : é o parâmetro de dificuldade (ou de posição) do item i , medido na mesma escala da habilidade;

a_i : é o parâmetro de discriminação (ou de inclinação) do item i , com valor proporcional à inclinação da Curva Característica do Item - CCI no ponto b_i ;

c_i : é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item i (muitas vezes referido como a probabilidade de acerto casual);

D : representa o fator constante de escala igual a 1. Pode-se utilizar o valor de 1.7 quando se deseja que os resultados da função logística sejam próximos usando os parâmetros da Ogiva.

No que concerne a interpretação do modelo logístico de 3 parâmetros, tem-se que a $P(U_{ij} = 1|\theta_j)$ é considerada como a proporção de respostas corretas ou proporção de respostas concordo ou proporção de respostas que satisfazem o item i dentre todos os indivíduos da população com um traço latente, θ_j .

Neste contexto, na Figura 3.2, ilustra a Curva Característica do Item - CCI para o caso do modelo de TRI de 3 parâmetros (vide equação 3.2).

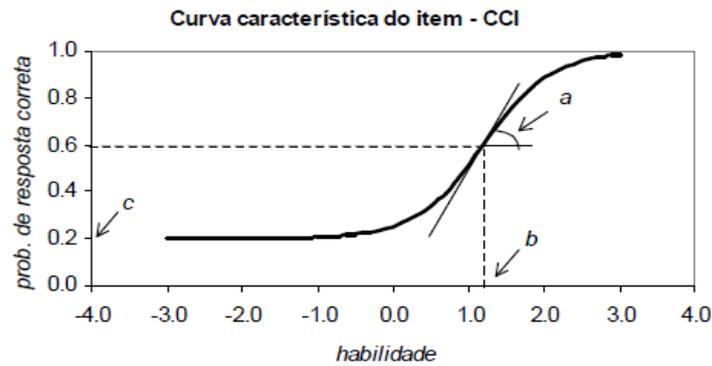


Figura 3.2: Exemplo de uma Curva Característica do Item - CCI. Adaptado de *Andrade et al. (2000)*.

Na Figura 3.2 pode-se observar que o modelo de TRI em questão é não linear e, quanto maior a habilidade (traço latente) maior a probabilidade do item ser respondido corretamente. A CCI tem o formato de uma curva em “S”, com deslocamento na escala do traço latente e descrita pelos parâmetros dos itens descritos acima.

A seguir apresentam-se os conceitos de Escala de Medida na TRI, da função de Informação do Item e da função de Informação do Teste.

3.5.1.1.1 Escala de Medida na TRI

Neste contexto, vale ressaltar que, segundo *Andrade et al. (2000)*, o desempenho descrito pelo parâmetro θ , dado o conjunto de itens pré-determinados, pode, na teoria, atingir qualquer valor real entre $-\infty$ e $+\infty$ e faz-se necessário definir um ponto de partida e também uma unidade de medida para a elaboração da escala de θ . A escolha desses valores é realizada de forma que represente o valor médio e o desvio-padrão da variável latente em análise. Apesar da frequente utilização da escala (0,1) para θ , implicando em média igual a zero e desvio-padrão igual a um, em termos práticos, não faz a menor diferença estabelecer-se estes valores ou outros quaisquer. O importante são as relações de ordem existentes entre seus pontos.

No caso da escala (0,1), os valores do parâmetro b variam empiricamente entre -2 e +2. No caso de a , espera-se valores entre 0 e +2, sendo mais adequados os valores de a maiores que 1. Contudo, esses valores podem ser modificados - a partir da transformação de todos os parâmetros - para valores que melhor caracterizem o desempenho a ser medido, desde que as posições entre seus pontos permaneçam inalteradas.

Por exemplo, $a(\theta - b)$ pode ser reescrito como:

$$\begin{aligned} a(\theta - b) &= \frac{a}{\gamma}[(\gamma\theta + \alpha) - (\gamma b + \alpha)] \\ &= a^*(\theta^* - b^*). \end{aligned}$$

Assim, tem-se:

1. $\theta^* = \gamma\theta + \alpha$;
2. $b^* = \gamma b + \alpha$;
3. $a^* = a/\gamma$;

$$4. P(U_i = 1|\theta) = P(U_i = 1|\theta^*),$$

em que

θ^* : consiste a habilidade (traço latente) na escala transformada;

a^* : refere-se ao parâmetro de discriminação na escala transformada;

b^* : é o parâmetro de dificuldade na escala transformada;

μ : é a média desejada na escala transformada;

τ : é o desvio padrão na escala transformada.

Nesse sentido, ressalta-se que a probabilidade, bem como o desempenho (habilidade) não se modifica, apenas são transformados mantendo-se a relação existente entre os valores.

De acordo com o exposto, alguns exemplos de curvas características e de curvas de informação (traçado pontilhado) de itens com diferentes combinações de valores dos parâmetros a e b são apresentados na Figura 3.3.

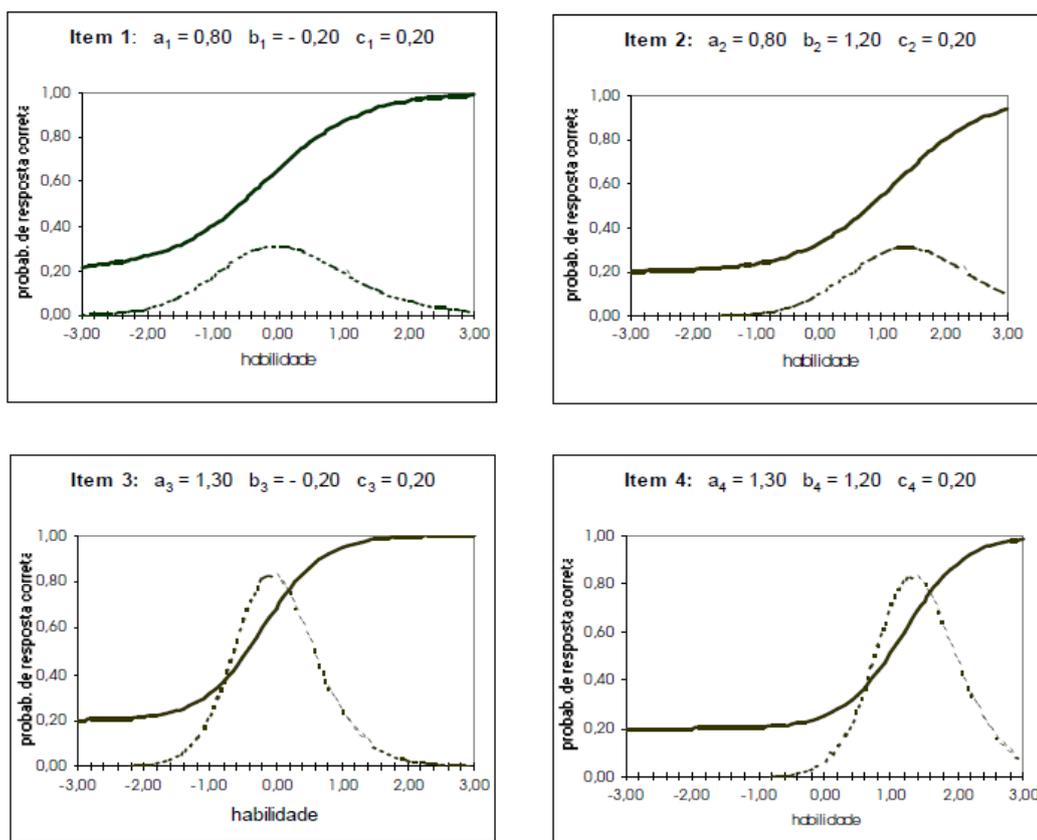


Figura 3.3: Exemplos de curvas características e de informação de vários itens. Adaptado de Andrade et al. (2000).

Com base na Figura 3.3, conforme ressaltam Andrade et al. (2000), observa-se que: *i*) comparando-se os itens 2 e 4 (e também os itens 1 e 3) percebe-se que os itens com maior valor do parâmetro a têm a curva característica com inclinação mais acentuada; *ii*) neste contexto, o item 4 é mais apropriado para discriminar estes dois indivíduos do que o item 2. Neste

contexto, o parâmetro a é denominado de *parâmetro de discriminação (ou de inclinação)* do item; *iii)* outra questão importante levantada pelos autores é que, itens com maior valor do parâmetro b exigem maior habilidade para a obtenção de resposta correta; *iv)* desta feita, o item 2 é mais difícil do que o item 1. Assim, o parâmetro b é denominado de *parâmetro de dificuldade (ou de posição)* do item.

3.5.1.1.2 Função de Informação do Item

A *função de informação do item* corresponde a uma medida bastante utilizada em conjunto com a Curva Característica do Item - CCI. Nesta mesma temática, [Andrade et al. \(2000\)](#) destacam que ela permite analisar quanto um item (ou teste) contém de informação para a medida de habilidade. Segundo os autores, a função de informação de um item é dada por:

$$I_i(\theta) = \frac{\left[\frac{d}{d\theta}P_i(\theta)\right]^2}{P_i(\theta)Q_i(\theta)}, \quad (3.3)$$

em que

$I_i(\theta)$ é a “informação” fornecida pelo item i no nível de habilidade θ ;

$P_i(\theta) = P(X_{ij} = 1|\theta)$ e $Q_i(\theta) = 1 - P_i(\theta)$.

Outra questão importante levantada pelos autores supra citados, consiste em reescrever o modelo logístico de 3 (três) parâmetros, expressão 3.2, a partir da equação 3.3 como

$$I_i(\theta) = D^2 a_i^2 \frac{Q_i(\theta)}{P_i(\theta)} \left[\frac{P_i(\theta) - c_i}{1 - c_i} \right]^2. \quad (3.4)$$

No que concerne a expressão 3.4, conforme ressaltam [Andrade et al. \(2000\)](#), vale ressaltar a importância que têm os três parâmetros sobre o montante de informação do item. Isto é, a informação é maior: *(i)* quando b_i se aproxima de θ ; *(ii)* quanto maior for o valor de a_i ; *(iii)* e quanto mais c_i se aproxima de 0 (zero).

3.5.1.1.3 Função de Informação do Teste

[Andrade et al. \(2000\)](#) destacam que a informação fornecida pelo teste é simplesmente a soma das informações fornecidas por cada item que compõe o mesmo:

$$I_i(\theta) = \sum_{i=1}^I I_i(\theta). \quad (3.5)$$

Nesta mesma linha, conforme ressaltam [Andrade et al. \(2000\)](#), uma outra maneira de representar esta função de informação do teste é através do erro-padrão de medida, chamado na TRI de erro-padrão de estimação, que é dado por:

$$EP(\theta) = \frac{1}{\sqrt{I_i(\theta)}}. \quad (3.6)$$

É importante notar que as medidas 3.5 e 3.6 de informação dependem do valor do parâmetro θ . Desta feita, a amplitude do intervalo de confiança para θ dependerá também do seu valor.

Portanto, após a apresentação do ML3, da escala de medida na TRI, da função de informação do item e da função de informação do teste, na sequência, apresenta-se o modelo logístico de 2 parâmetros (ML2) e de 1 parâmetro (ML1), que constituem variações deste modelo.

3.5.1.2 O modelo logístico de 2 parâmetros (ML2)

Definição 2: No modelo logístico de 2 parâmetros (ML2), o parâmetro c é igual a 0 (zero), ou seja, não há a possibilidade de acerto casual, segundo [Andrade et al. \(2000\)](#). Assim, tem-se o chamado *Modelo logístico unidimensional de 2 parâmetros (ML2)*, dado por ([Lord, 1952](#)):

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + \exp\{-Da_i(\theta_j - b_i)\}}, \quad (3.7)$$

com $i = 1, 2, \dots, I$, e $j = 1, 2, \dots, n$.

As variáveis da expressão 3.7 do ML2 estão descritas na fórmula 3.2, quando esta foi apresentada nesta pesquisa.

No que concerne a aplicação do ML2, [Embretson e Reise \(2000, p. 70\)](#), ressaltam que o ML2 é recomendado para medidas nas quais os itens possuem diferentes discriminações para determinado traço latente. Reforça-se, também, que o ML3 é utilizado para essa mesma finalidade, porém, quando há a possibilidade de acerto ao acaso.

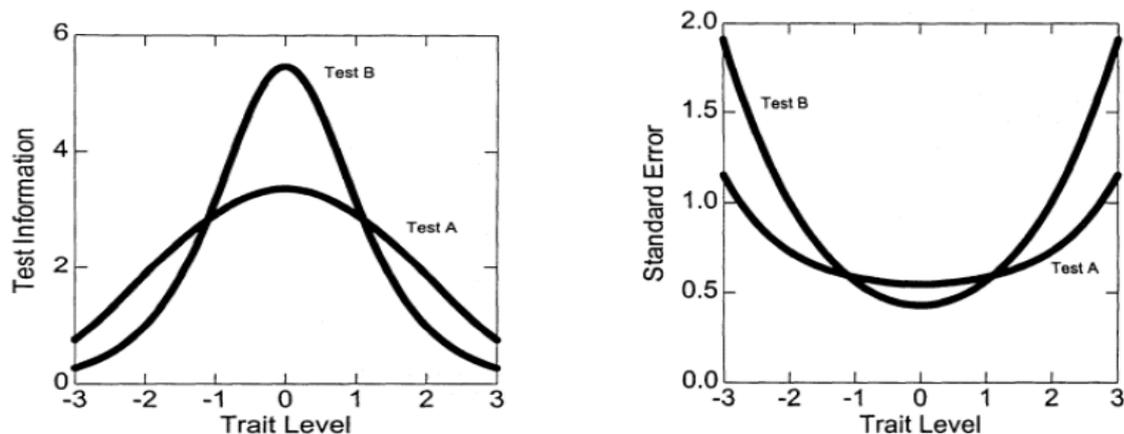
3.5.1.3 O modelo logístico de 1 parâmetros (ML1)

Definição 3: O *modelo logístico unidimensional de 1 parâmetro (ML1)*, também conhecido como modelo de Rasch, caracteriza-se por apresentar, além do parâmetro c igual a zero, o parâmetro de discriminação a igual a 1 (um). Assim, todos os itens possuem poder de discriminação igual, de acordo com [Rasch \(1960\)](#) e [Wright \(1977\)](#). A fórmula para este modelo é apresentada pelos autores:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + \exp\{-D(\theta_j - b_i)\}}, \quad (3.8)$$

com $i = 1, 2, \dots, I$, e $j = 1, 2, \dots, n$.

Um exemplo de curva de informação do teste e de erro padrão de medida é apresentado na Figura 3.4.



(a) Curva de informação dos testes A e B.

(b) Curva do erro padrão dos testes A e B.

Figura 3.4: Curva de informação e erro padrão dos testes A e B. Adaptado de Embretson e Reise (2000, p. 168).

Na Figura 3.4, nota-se que o erro padrão da medida, Gráfico (b), é menor no intervalo no qual a curva de informação do item atinge maior quantidade de informação. Nesse sentido, para o teste A, que atinge maior quantidade de informação entre o intervalo de -1,5 a 1,5, aproximadamente, o erro padrão também é menor neste intervalo. Para o teste B, cuja quantidade maior de informação se concentra no intervalo entre -1 a 1, o erro padrão é menor que 0,5 neste intervalo.

Na sequência, apresentam-se, os modelos para itens politômicos.

3.5.2 Modelos para itens não dicotômicos

De acordo com Andrade *et al.* (2000), tem-se que os modelos para os itens não dicotômicos, ou politômicos, referem-se aos modelos para a avaliação de itens que possuem respostas abertas ou itens que possuem várias alternativas de respostas (múltipla escolha).

Araujo *et al.* (2009), ressaltam que os modelos para itens politômicos dependem da natureza das categorias de resposta. Em testes de múltipla escolha, em que as categorias não são ordenadas, o modelo é denominado Modelo Nominal e, nos casos em que as categorias são ordenadas, o modelo é denominado Modelo Ordinal, por exemplo, quando as categorias dos itens são dadas conforme uma escala de Likert¹.

Conforme a abordagem de Andrade *et al.* (2000) e Embretson e Reise (2000), serão apresentados nesta seção os seguintes modelos: Modelo de Resposta Nominal, Modelo de Resposta Gradual, Modelo de Escala Gradual, Modelo de Crédito Parcial e Modelo de Crédito Parcial Generalizado.

3.5.2.1 Modelo de Resposta Nominal (Nominal Categories Model)

O *modelo de Resposta Nominal*, desenvolvido por Bock (1972), pode ser aplicado a testes de múltipla escolha. Dessa forma, as categorias são não ordenadas. O modelo proposto tem

¹A *escala Likert* ou *escala de Likert* é um tipo de escala de resposta psicométrica usada habitualmente em questionários, e é a escala mais usada em pesquisas de opinião. Ao responderem a um questionário baseado nesta escala, os perguntados especificam seu nível de concordância com uma afirmação, (Likert, 1932)

como intuíto “maximizar a precisão da habilidade estimada usando toda a informação contida nas respostas dos indivíduos, e não apenas se o item foi respondido corretamente ou não” (Andrade *et al.*, 2000).

Nesta mesma linha, Bock descreveu que a probabilidade com que um indivíduo j selecionaria uma opção particular k (de m_i opções avaliáveis) do item i seria representada por:

$$P_{i,k}(\theta_j) = \frac{\exp\{a_{i,k}^+(\theta_j - b_{i,k}^+)\}}{\sum_{h=1}^{m_i} \exp\{a_{i,h}^+(\theta_j - b_{i,h}^+)\}}, \quad (3.9)$$

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$ e $k = 1, 2, \dots, m_i$.

No que tange a expressão 3.9, vale ressaltar os seguintes itens:

- i) em cada θ_j , a soma das probabilidades sobre as m_i opções, $\sum_{k=1}^{m_i} P_{i,k}(\theta_j)$ é 1 e, que;
- ii) as quantidades $(b_{i,k}^+; a_{i,k}^+)$ são parâmetros do item i relacionado a k -ésima opção

3.5.2.2 Modelo de Resposta Gradual (Graded Response Model)

Na mesma linha, o modelo politômico a ser apresentado é o *Modelo de Resposta Gradual* de Samejima (1969). Este modelo, pressupõe ordenação entre as alternativas de respostas de um item, de acordo com Andrade *et al.* (2000), sendo possível captar mais informação referente à resposta fornecida, não se limitando somente à análise se a resposta foi correta ou incorreta.

Embretson e Reise (2000) ainda, ressaltam que este modelo é uma extensão do modelo logístico de dois parâmetros (ML2), já apresentado anteriormente.

Andrade *et al.* (2000), destacam que “os escores das categorias de um item i são arranjados em ordem do menor para o maior e denotados por $k = 1, 2, \dots, m_i$ onde $(m_i + 1)$ é o número de categorias do i -ésimo item. A probabilidade $P_{i,k}^+(\theta_j)$ de um indivíduo j escolher *uma particular categoria ou outra mais alta* do item i pode ser dada por uma extensão do modelo logístico de 2 parâmetros”:

$$P_{i,k}^+(\theta_j) = \frac{1}{1 + \exp\{-Da_i(\theta_j - b_{i,k})\}}, \quad (3.10)$$

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$ e $k = 1, 2, \dots, m_i$, em que $b_{i,k}$ é o parâmetro de dificuldade da k -ésima categoria do item i .

Andrade *et al.* (2000) destacam que no caso de modelos para itens não dicotômicos, a discriminação de uma categoria específica de resposta depende tanto do parâmetro de inclinação, comum a todas as categorias do item, quanto da distância das categorias de dificuldade adjacentes.

Cabe ressaltar que, da definição, deve-se ter (Andrade *et al.*, 2000):

$$b_{i,1} \leq b_{i,2} \leq \dots \leq b_{i,m_i},$$

ou seja, deve-se ter, necessariamente, uma ordenação entre o nível de dificuldade das categorias de um dado item, de acordo com a classificação de seus escores.

Complementando o exposto, a probabilidade de um indivíduo j receber um escore k no item i , de acordo com os autores supra citados, é dada, então, pela expressão:

$$P_{i,k}(\theta_j) = P_{i,k}^+(\theta_j) - P_{i,k+1}^+(\theta_j).$$

Samejima (1969) também define $P_{i,0}^+(\theta_j)$ e $P_{i,m_i+1}^+(\theta_j)$, de modo que:

$$P_{i,0}^+(\theta_j) = 1$$

e

$$P_{i,m_i+1}^+(\theta_j) = 0.$$

Desta feita,

$$P_{i,0}(\theta_j) = P_{i,0}^+(\theta_j) - P_{i,1}^+(\theta_j) = 1 - P_{i,1}^+(\theta_j)$$

e

$$P_{i,m}(\theta_j) = P_{i,m}^+(\theta_j) - P_{i,m_i+1}^+(\theta_j) = P_{i,m}^+(\theta_j).$$

Portanto, tem-se que:

$$P_{i,k}(\theta_j) = \frac{1}{1 + \exp\{-Da_i(\theta_j - b_{i,k})\}} - \frac{1}{1 + \exp\{-Da_i(\theta_j - b_{i,k+1})\}}. \quad (3.11)$$

No que concerne a expressão 3.11 nota-se que em um item com $(m_i + 1)$ categorias, m_i valores de dificuldade necessitam ser estimados, além do parâmetro de inclinação do item. Assim, para cada item, o número de parâmetros a ser estimado será dado pelo seu número de categorias de resposta. No caso de um teste com I itens, cada um com $(m_i + 1)$ categorias de resposta, tem-se então $\left[\sum_{i=1}^I m_i + I \right]$ parâmetros de item a serem estimados.

3.5.2.3 Modelo de Escala Gradual (Rating Scale Model)

O terceiro modelo politômico a ser apresentado é o Modelo de Escala Gradual. Este modelo é um caso particular do modelo de resposta gradual de Samejima (1969) e também pressupõe uma ordenação entre as alternativas de respostas de um item.

No entanto, faz-se a resolução adicional das categorias ordenadas serem igualmente espaçadas.

Este modelo, foi proposto por Andrich (1978) e dado por

$$P_{i,k}(\theta_j) = \frac{1}{1 + \exp\{-Da_i(\theta_j - b_i + d_k)\}} - \frac{1}{1 + \exp\{-Da_i(\theta_j - b_i + d_{k+1})\}}, \quad (3.12)$$

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$ e $k = 1, 2, \dots, m$, em que

b_i é agora o parâmetro de locação do item i e

d_k é o parâmetro de categoria.

Como $P_{i,k}^+(\theta_j) - P_{i,k+1}^+(\theta_j) \geq 0$, então $d_k - d_{k+1} \geq 0$ e $d_1 \geq d_2 \geq \dots \geq d_m$. Sob a suposição de escores equidistantes das categorias de resposta, o parâmetro $b_{i,k}$ no modelo de escala gradual fica

$$b_{i,k} = b_i - d_k.$$

Embretson e Reise (2000) destacam que este modelo é indicado para itens com categorias ordenadas e que este modelo é uma extensão do modelo logístico de dois parâmetros.

Andrade *et al.* (2000) destacam que os parâmetros de categoria d_k não dependem do item, isto é, são *comuns* a todos os itens do teste. Logo, se os itens que compõem a prova tiverem suas próprias categorias de resposta, que podem diferir no número, então este modelo não é adequado.

Na Figura 3.5, tem-se a representação do Modelo de Escala Gradual de Andrich, bem como o de Resposta Gradual de Samejima, de acordo com a abordagem dos autores, considerando 4 categorias de resposta para cada item.

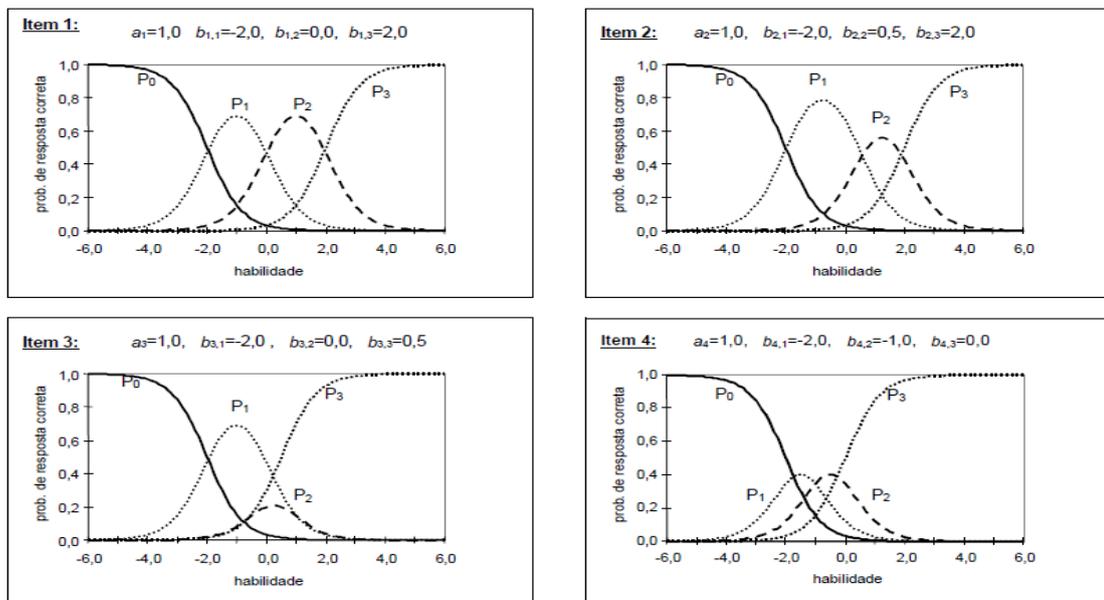


Figura 3.5: Representação gráfica dos Modelos de Escala Gradual e de Resposta Gradual. Adaptado de Andrade *et al.* (2000).

De acordo com Figura 3.5, destaca-se que o pressuposto complementar do modelo de escala gradual, quando comparado ao de Samejima, é de que, além de possuírem categorias

de respostas ordenadas, os escores das categorias são *igualmente espaçados*, (Andrade *et al.*, 2000).

Dessa forma, observa-se que os itens 1 e 4 podem representar o modelo de escala gradual - por este configurar-se como uma extensão do modelo de resposta gradual de Samejima - com base na abordagem dos autores, pois apresentam os parâmetros de categoria igualmente espaçados. Neste contexto, vale ressaltar que, todos os itens apresentados na Figura 3.5 podem representar o modelo de resposta gradual, por possuírem categorias de respostas ordenadas.

3.5.2.4 Modelo de Crédito Parcial (Partial Credit Model)

O modelo de Crédito Parcial foi desenvolvido por Masters (1982), sendo uma extensão do Modelo de Rasch (ML1) para itens dicotômicos. No modelo de Crédito Parcial todos os parâmetros são de locação e assume-se que o poder de discriminação é igual para todos os itens. Este modelo, é indicado para itens cujas categorias de respostas são ordenadas.

Supondo que o item i tem $(m_i + 1)$ categorias de resposta ordenáveis ($k = 0, 1, \dots, m_i$), tem-se que o modelo de crédito parcial é dado por (Andrade *et al.*, 2000):

$$P_{i,k}(\theta_j) = \frac{\exp \left\{ \sum_{u=0}^k (\theta_j - b_{i,u}) \right\}}{\sum_{u=0}^{m_i} \exp \left\{ \sum_{v=0}^u (\theta_j - b_{i,v}) \right\}}, \quad (3.13)$$

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$, $k = 1, 2, \dots, m_i$ e $b_{i,0} \equiv 0$, em que

$P_{i,k}(\theta_j)$ é a probabilidade de um indivíduo com habilidade θ_j escolher a categorias k , dentre as $(m_i + 1)$ categorias do item i ;

$b_{i,k}$ é o parâmetro de item que regula a probabilidade de escolher a categoria k em vez da categoria adjacente $(k - 1)$ no item i . Cada parâmetro $b_{i,k}$ corresponde ao valor de habilidade em que o indivíduo tem a mesma probabilidade de responder à categoria k e à categoria $(k - 1)$, isto é, onde $P_{i,k}(\theta_j) = P_{i,k-1}(\theta_j)$.

Assim, segundo Andrade *et al.* (2000) para itens com $(m_i + 1)$ categorias de resposta, será necessário estimar m_i parâmetros de item. Note que, para itens com apenas 2 categorias de resposta, este modelo é análogo ao de Rasch para itens dicotômicos.

3.5.2.5 Modelo de Crédito Parcial Generalizado (Generalized Partial Credit Model)

De Masters (1982), Muraki (1992) desenvolveu o *modelo de crédito parcial generalizado - MCPG*, cuja principal diferença em relação ao modelo de Masters consiste no relaxamento do pressuposto de homogeneidade do poder de discriminação dos itens.

O modelo de crédito parcial generalizado - MCPG é expressado por:

$$P_{i,k}(\theta_j) = \frac{\exp \left\{ \sum_{u=0}^k Da_i(\theta_j - b_{i,u}) \right\}}{\sum_{u=0}^{m_i} \exp \left\{ \sum_{v=0}^u Da_i(\theta_j - b_{i,v}) \right\}}, \quad (3.14)$$

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$, $k = 1, 2, \dots, m_i$.

Segundo Embretson e Reise (2000), o modelo generalizado do modelo de crédito parcial propiciou que os itens dentro de uma escala diferissem no parâmetro de inclinação.

Segundo Embretson e Reise (2000), os parâmetros de inclinação, a_i , não são interpretados da mesma forma que os modelos dicotômicos de TRI. Isto porque nos modelos politômicos a discriminação depende da combinação dos parâmetros de inclinação e de dispersão nos limites entre as categorias.

Na Tabela 3.1 apresenta-se a classificação dos modelos unidimensionais da TRI, em função da natureza do item.

Tabela 3.1: Classificação dos modelos unidimensionais de acordo com a natureza do item.

Natureza do item	Modelos	Parâmetros considerados
Dicotômico (um grupo ou mais)	Modelo logístico de 3 parâmetros	Dificuldade do item Discriminação do item Acerto casual
	Modelo logístico de 2 parâmetros	Dificuldade do item Discriminação do item
	Modelo logístico de 1 parâmetro (Modelo de Rasch)	Dificuldade do item
Politômico (não - dicotômico)		
Nominal	Modelo de Resposta Nominal (Bock, 1972)	Dificuldade do item Discriminação do item ML2
Ordinal	Modelo de Resposta Gradual (Samejima, 1969)	Dificuldade do item Discriminação do item ML2
	Modelo de Escala Gradual (Andrich, 1978)	Dificuldade do item Discriminação do item ML2
	Modelo de crédito Parcial (Masters, 1982)	Dificuldade do item ML1
	Modelo de crédito Parcial Generalizado (Muraki, 1982)	Dificuldade do item Discriminação do item ML2
Modelos que englobam duas ou mais populações	Bock e Zimowski (1997) Extensão dos modelos ML1, ML2 e ML3	ML1 ML2 ML3

Fonte: Elaborado pelos autores com base em Andrade *et al.* (2000) e Embretson e Reise (2000).

Considerando a Tabela 3.1, vale ressaltar que estes modelos caracterizam-se como modelos cumulativos, apresentando até três parâmetros (de dificuldade, de discriminação e de acerto casual).

Na sequência o Capítulo 4, apresenta os principais métodos de estimação dos parâmetros dos modelos para uma única população.

Capítulo 4

Processo de estimação dos parâmetros na TRI para dados binários

4.1 Introdução

O presente capítulo apresenta a estimação dos parâmetros dos itens para dados binários. Na sequência, abordam-se, também, os métodos de máxima verossimilhança e suas principais características, a estimação das habilidades quando os parâmetros dos itens são conhecidos e, a estimação conjunta: quando tanto os parâmetros dos itens quanto das habilidades necessitam ser estimados.

Baker e Kim (2004) enfatizam que uma das etapas mais relevantes da aplicação da Teoria da Resposta ao Item refere-se à estimação dos parâmetros dos itens sob a suposição de que os valores das habilidades dos examinandos sejam conhecidos.

Nesta mesma conjectura, de acordo com Baker e Kim (2004, p. 23-24), supondo a existência de vários grupos de examinandos, com valores de habilidade conhecidos e distribuídas ao longo da escala de habilidade variando desde valores muito baixos até valores muito altos. Nos níveis baixos de habilidade, uma pequena proporção dos examinandos em um grupo, responderá corretamente aos itens dicotômicos. Nos níveis altos de habilidade, uma grande proporção responderá corretamente a tais itens. Estas proporções observadas de respostas corretas podem, então, ser representadas graficamente como uma função dos valores das habilidades, como está representado na Figura 4.1. Em seguida, a tarefa básica é encontrar a curva característica, sob um determinado dado modelo, que se ajuste às proporções observadas de respostas corretas. Na Figura 4.1, a curva representa o ajuste de um modelo para este conjunto de dados.

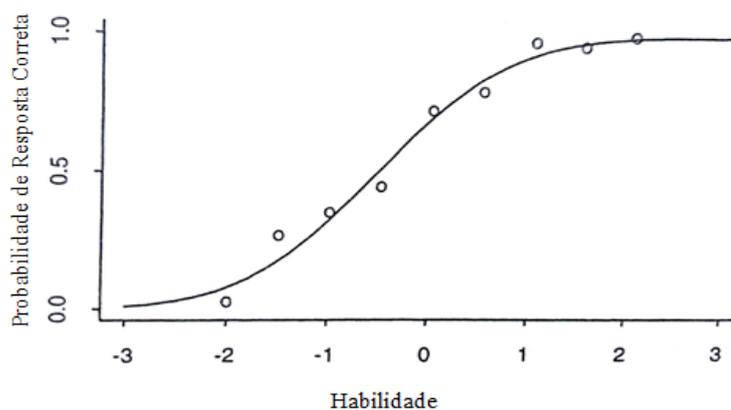


Figura 4.1: Proporção observada da resposta correta em função da habilidade. Adaptado de Baker e Kim (2004).

Conforme pode ser observado na Figura 4.1, quanto maior a habilidade maior a probabilidade de responder o item corretamente. O modelo de TRI considera esse caráter acumulativo da escala de proficiência, notadamente na forma que a probabilidade de sucesso é descrita, ou através de um modelo logístico ou de um probito.

A próxima etapa consiste na estimação dos parâmetros dos itens e dos traços latentes. Os métodos mais empregados são o Método de Máxima Verossimilhança e Métodos Bayesianos. Nesta mesma linha, vale ressaltar que o processo de estimação dos parâmetros dos itens da TRI é denominado de *calibração*.

O processo de estimação na TRI normalmente é realizado pelo método da máxima verossimilhança, por meio da utilização de algum processo iterativo, como, por exemplo, o algoritmo de *Newton-Raphson* (Issac e Keller, 1966), o “*Scoring*” de Fisher (Rao, 1973), ou algum procedimento bayesiano (Mislevy, 1986).

No que concerne aos métodos de estimação mais frequentemente utilizados, os autores Embretson e Reise (2000) destacam: *i*) o método da máxima verossimilhança condicional; *ii*) o método da máxima verossimilhança, utilizando-se, geralmente, o método de máxima verossimilhança conjunta; *iii*) o método da máxima verossimilhança marginal. Os métodos de estimação bayesianos podem ser utilizados se as informações sobre os parâmetros dos itens estiverem disponíveis.

Notação

Utilizou-se nesta seção a notação e desenvolvimentos descritos em Andrade *et al.* (2000).

Sejam θ_j a habilidade do indivíduo j e U_{ji} a variável aleatória que representa a resposta (binária) do indivíduo j ao item i , com

$$U_{ji} = \begin{cases} 1, & \text{se o indivíduo } j \text{ responder corretamente ao item } i, \\ 0, & \text{caso contrário.} \end{cases} \quad (4.1)$$

Sejam $\mathbf{U}_j = (\mathbf{U}_{j1}, \mathbf{U}_{j2}, \dots, \mathbf{U}_{jI})$ o vetor aleatório de respostas do indivíduo j e $\mathbf{U}_\cdot = (\mathbf{U}_{1\cdot}, \mathbf{U}_{2\cdot}, \dots, \mathbf{U}_{n\cdot})$ o conjunto integral de respostas. De forma similar, representa-se as observações por u_{ji} , \mathbf{u}_j , e \mathbf{u}_\cdot . Vale ressaltar que, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ representará o vetor de habilidades dos n indivíduos e $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_I)$ o conjunto de parâmetros dos itens.

Suposições básicas do modelo

Apresenta-se a seguir, as duas principais suposições que usaremos nos modelos de TRI:

- (S1) as respostas oriundas de indivíduos diferentes são independentes;
- (S2) os itens são respondidos de forma independente por cada indivíduo (Independência Local), fixada sua habilidade.

Lord *et al.* (1968) enfatizam que a suposição (S2) garante que, para cada valor de θ , ao tomar-se um conjunto de indivíduos com habilidade θ , as covariâncias entre as respostas para cada par de itens serão nulas. Entretanto, se for considerado um conjunto de indivíduos com habilidades variadas, estas covariâncias não, necessariamente, serão nulas. Na verdade, elas serão positivas.

Vale ressaltar que, nos modelos da TRI, pode-se verificar um problema denominado *falta de identificabilidade do modelo*. Esta não-identificabilidade ocorre porque mais de um conjunto de parâmetros produz o mesmo valor na probabilidade dada pelos modelos. Essa não-identificabilidade pode ser eliminada, por exemplo, fixando alguns valores para o traço latente.

Neste contexto, Andrade *et al.* (2000) enfatizam que essa não-identificabilidade está fortemente relacionada às características da população em estudo. Os autores complementam que, para resolver este problema, basta especificar/fixar uma medida de posição (média, por exemplo) e outra de dispersão (desvio-padrão, por exemplo) para o traço latente. Deste modo, ao definir uma métrica (unidade de medida) para o traço latente e, naturalmente para os parâmetros dos itens, elimina-se o problema da não-identificabilidade. É usual definir esta métrica na forma (μ, σ) com $\mu = 0$ e $\sigma = 1$.

Apresenta-se, a seguir, a estimação dos parâmetros dos itens da TRI, pelo método da máxima verossimilhança, quando as habilidades são conhecidas.

4.2 Estimação dos parâmetros dos itens

Nesta seção apresenta-se o Método da Máxima Verossimilhança Condicional na estimação dos parâmetros dos itens, pelo método da máxima verossimilhança, quando as habilidades são conhecidas.

4.2.1 Método da Máxima Verossimilhança Condicional

Conforme a abordagem de Embretson e Reise (2000, p. 214), o Método da máxima verossimilhança condicional (CML, do inglês *Conditional Maximum Likelihood*): “modela as probabilidades de resposta a partir das probabilidades de vários padrões de resposta que conduz ao mesmo escore total”. Neste método, de acordo com os autores, os parâmetros dos itens são estimados por um processo iterativo de busca, como o Newton-Raphson.

Complementando o exposto e levando em consideração a expressão 4.1, enfatiza-se que, pela independência entre as respostas dos diferentes indivíduos (S1), e a independência local (S2), de acordo com Hambleton e Swaminathan (1985, Cap. 7) e considerando a notação matemática abordada por Andrade *et al.* (2000), pode-se escrever a verossimilhança, $L(\zeta) = P(\mathbf{U}_{..} = \mathbf{u}_{..} | \boldsymbol{\theta}, \zeta)$, como

$$L(\zeta) = P(U_{j1} = \mathbf{u}_{j1} | \boldsymbol{\theta}_1, \zeta, \dots, U_{jn} = \mathbf{u}_{jn} | \boldsymbol{\theta}_n, \zeta)$$

$$\begin{aligned}
&= P(\mathbf{U}_{j1} = \mathbf{u}_{j1} | \boldsymbol{\theta}_1, \boldsymbol{\zeta}), \dots, P(\mathbf{U}_{jn} = \mathbf{u}_{jn} | \boldsymbol{\theta}_n, \boldsymbol{\zeta}) \\
&= \prod_{j=1}^n P(\mathbf{U}_j = \mathbf{u}_j | \boldsymbol{\theta}_j, \boldsymbol{\zeta}),
\end{aligned}$$

em que $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_I)$ é o conjunto de parâmetros dos itens. Dessa forma,

$$L(\boldsymbol{\zeta}) = \prod_{j=1}^n \prod_{i=1}^I P(\mathbf{U}_{ji} = \mathbf{u}_{ji} | \boldsymbol{\theta}_j, \boldsymbol{\zeta}_i). \quad (4.2)$$

Verifica-se que, na última igualdade, considera-se que a distribuição de \mathbf{U}_{ji} só depende de $\boldsymbol{\zeta}$ através de $\boldsymbol{\zeta}_i$. Desta feita, utilizando-se a notação $P_{ji} = P(\mathbf{U}_{ji} = 1 | \boldsymbol{\theta}_j, \boldsymbol{\zeta}_i)$ e $Q_{ji} = 1 - P_{ji}$, tem-se que

$$\begin{aligned}
P(\mathbf{U}_{ji} = \mathbf{u}_{ji} | \boldsymbol{\theta}_j, \boldsymbol{\zeta}_i) &= P(\mathbf{U}_{ji} = 1 | \boldsymbol{\theta}_j, \boldsymbol{\zeta}_i)^{\mathbf{u}_{ji}} P(\mathbf{U}_{ji} = 0 | \boldsymbol{\theta}_j, \boldsymbol{\zeta}_i)^{1-\mathbf{u}_{ji}} \\
&= P_{ji}^{\mathbf{u}_{ji}} Q_{ji}^{1-\mathbf{u}_{ji}}.
\end{aligned} \quad (4.3)$$

Logo,

$$L(\boldsymbol{\zeta}) = \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{\mathbf{u}_{ji}} Q_{ji}^{1-\mathbf{u}_{ji}}. \quad (4.4)$$

Desta feita, segue-se que a função de log-verossimilhança é dada por

$$\ln L(\boldsymbol{\zeta}) = \sum_{j=1}^n \sum_{i=1}^I \{ \mathbf{u}_{ji} \ln P_{ji} + (1 - \mathbf{u}_{ji}) \ln Q_{ji} \}. \quad (4.5)$$

Neste contexto, os Estimadores de Máxima Verossimilhança (EMV) de $\boldsymbol{\zeta}_i$, $i = 1, \dots, I$, são os valores que maximizam a log-verossimilhança (4.5), ou equivalentemente, são as soluções da equação

$$\frac{\partial \ln L(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i} = \mathbf{0}, \quad i = 1, \dots, I.$$

Nota-se que

$$\begin{aligned}
\frac{\partial \ln L(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i} &= \sum_{j=1}^n \left\{ \frac{\partial \ln(L(\boldsymbol{\zeta}))}{\partial P_{ji}} \frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right\} \\
&= \sum_{j=1}^n \left\{ \mathbf{u}_{ji} \frac{\partial(\ln P_{ji})}{\partial \boldsymbol{\zeta}_i} + (1 - \mathbf{u}_{ji}) \frac{\partial(\ln Q_{ji})}{\partial \boldsymbol{\zeta}_i} \right\} \\
&= \sum_{j=1}^n \left\{ \mathbf{u}_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) + (1 - \mathbf{u}_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) \right\}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^n \left\{ \mathbf{u}_{ji} \frac{1}{P_{ji}} - (1 - \mathbf{u}_{ji}) \frac{1}{Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) \\
&= \sum_{j=1}^n \left\{ \frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right). \tag{4.6}
\end{aligned}$$

Por conveniência, considera-se a ponderação:

$$W_{ji} = \frac{P_{ji}^* Q_{ji}^*}{P_{ji} Q_{ji}}, \tag{4.7}$$

em que

$$P_{ji}^* = \{1 + \exp\{-D_{a_i}(\theta_j - b_i)\}\}^{-1} \quad \text{e} \quad Q_{ji}^* = 1 - P_{ji}^*. \tag{4.8}$$

Desta feita, pode-se reescrever a expressão (4.6) como

$$\frac{\partial \ln L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right). \tag{4.9}$$

Complementando o exposto, vale ressaltar que cada uma destas equações é especificada em termos de ζ_i , $i = 1, \dots, I$.

No caso do modelo de TRI de 3 parâmetros, $\zeta_i = (a_i, b_i, c_i)$, $i = 1, \dots, I$, e em uma abordagem proposta por Hambleton e Swaminathan (1985), que também está descrita em Andrade *et al.* (2000), tem-se que, para obter as equações de estimação, precisa-se das seguintes expressões:

$$\frac{\partial P_{ji}}{\partial a_i} = D(1 - c_i)(\theta_j - b_i) P_{ji}^* Q_{ji}^*, \tag{4.10}$$

$$\frac{\partial P_{ji}}{\partial b_i} = -D a_i (1 - c_i) P_{ji}^* Q_{ji}^*, \tag{4.11}$$

$$\frac{\partial P_{ji}}{\partial c_i} = Q_{ji}^*. \tag{4.12}$$

Para o parâmetro de discriminação, a_i , de (4.9) e (4.10), tem-se

$$\begin{aligned}
\frac{\partial \ln L(\zeta)}{\partial a_i} &= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial a_i} \right) \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) D(1 - c_i)(\theta_j - b_i) P_{ji}^* Q_{ji}^* \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= D(1 - c_i) \sum_{j=1}^n \{ (\mathbf{u}_{ji} - P_{ji})(\theta_j - b_i) W_{ij} \}. \tag{4.13}
\end{aligned}$$

Analogamente, para o parâmetro de dificuldade, b_i , de (4.9) e (4.11), tem-se

$$\begin{aligned}
\frac{\partial \ln L(\zeta)}{\partial b_i} &= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial b_i} \right) \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) (-1) D a_i (1 - c_i) P_{ji}^* Q_{ji}^* \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= -D a_i (1 - c_i) \sum_{j=1}^n \{ (\mathbf{u}_{ji} - P_{ji}) W_{ij} \}.
\end{aligned} \tag{4.14}$$

Finalmente, para o parâmetro de acerto ao acaso, c_i , de (4.9) e (4.12), tem-se

$$\begin{aligned}
\frac{\partial \ln L(\zeta)}{\partial c_i} &= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial c_i} \right) \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) Q_{ji}^* \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) \frac{W_{ij}}{P_{ji}^*} \right\}.
\end{aligned} \tag{4.15}$$

Por fim, as equações de estimação para os parâmetros a_i , b_i e c_i são, respectivamente,

$$a_i : D(1 - c_i) \sum_{j=1}^n \{ (\mathbf{u}_{ji} - P_{ji}) (\boldsymbol{\theta}_j - b_i) W_{ij} \} = 0, \tag{4.16}$$

$$b_i : -D a_i (1 - c_i) \sum_{j=1}^n \{ (\mathbf{u}_{ji} - P_{ji}) W_{ij} \} = 0, \tag{4.17}$$

$$c_i : \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) \frac{W_{ij}}{P_{ji}^*} \right\} = 0. \tag{4.18}$$

Desta feita, segue a descrição do algoritmo de Newton-Raphson e do Método “Scoring” de Fisher na estimação dos parâmetros. Estes métodos iterativos são utilizados para a obtenção das estimativas de máxima verossimilhança dos parâmetros dos itens, visto que as expressões (4.16), (4.17) e (4.18) não possuem solução explícita.

4.2.1.1 Aplicação do algoritmo de Newton-Raphson

Seja $l(\zeta) = \ln L(\zeta)$ a log-verossimilhança, em que $\zeta = (\zeta_1, \dots, \zeta_I)$, com $\zeta_i = (a_i, b_i, c_i)'$. Para valores iniciais $\widehat{\zeta}_i^{(0)} = (a_i^{(0)}, b_i^{(0)}, c_i^{(0)})'$, obtem-se uma estimativa de ζ_i . No passo (1) do processo iterativo, tem-se: $\widehat{\zeta}_i^{(1)} = \widehat{\zeta}_i^{(0)} + \Delta \widehat{\zeta}_i^{(0)}$, ou seja,

$$\begin{aligned}
\widehat{\mathbf{a}}_i^{(1)} &= \widehat{\mathbf{a}}_i^{(0)} + \Delta \widehat{\mathbf{a}}_i^{(0)}, \\
\widehat{\mathbf{b}}_i^{(1)} &= \widehat{\mathbf{b}}_i^{(0)} + \Delta \widehat{\mathbf{b}}_i^{(0)}, \\
\widehat{\mathbf{c}}_i^{(1)} &= \widehat{\mathbf{c}}_i^{(0)} + \Delta \widehat{\mathbf{c}}_i^{(0)},
\end{aligned} \tag{4.19}$$

em que $\Delta\hat{\mathbf{a}}_i^{(0)}$, $\Delta\hat{\mathbf{b}}_i^{(0)}$ e $\Delta\hat{\mathbf{c}}_i^{(0)}$ são erros de aproximação. Usando-se a expressão (4.19) expandida em série de Taylor de $\partial l(\boldsymbol{\zeta})/\partial\boldsymbol{\zeta}_i$ no ponto $\hat{\boldsymbol{\zeta}}_i^{(0)}$, tem-se:

$$\begin{aligned}\frac{\partial l(\boldsymbol{\zeta})}{\partial a_i} &= \frac{\partial l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial a_i} + \Delta\hat{\mathbf{a}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial a_i^2} + \Delta\hat{\mathbf{b}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial a_i \partial b_i} + \Delta\hat{\mathbf{c}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial a_i \partial c_i} + \mathbf{R}_{\mathbf{a}_i}(\hat{\boldsymbol{\zeta}}_i^{(0)}), \\ \frac{\partial l(\boldsymbol{\zeta})}{\partial b_i} &= \frac{\partial l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial b_i} + \Delta\hat{\mathbf{b}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial b_i^2} + \Delta\hat{\mathbf{a}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial b_i \partial a_i} + \Delta\hat{\mathbf{c}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial b_i \partial c_i} + \mathbf{R}_{\mathbf{b}_i}(\hat{\boldsymbol{\zeta}}_i^{(0)}), \\ \frac{\partial l(\boldsymbol{\zeta})}{\partial c_i} &= \frac{\partial l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial c_i} + \Delta\hat{\mathbf{c}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial c_i^2} + \Delta\hat{\mathbf{a}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial c_i \partial a_i} + \Delta\hat{\mathbf{b}}_i^{(0)} \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial c_i \partial b_i} + \mathbf{R}_{\mathbf{c}_i}(\hat{\boldsymbol{\zeta}}_i^{(0)}),\end{aligned}$$

em que $\partial l(\hat{\boldsymbol{\zeta}}_i)/\partial\alpha_i$ representa a função $\partial l(\boldsymbol{\zeta}_i)/\partial\alpha_i$ avaliada no ponto $\boldsymbol{\zeta}_i = \hat{\boldsymbol{\zeta}}_i$. Vale ressaltar que nessas expressões o uso de $\partial l(\boldsymbol{\zeta})/\partial\boldsymbol{\zeta}_i$ está em função apenas de $\boldsymbol{\zeta}_i$, não dependendo de $\boldsymbol{\zeta}_l$ para $l \neq i$. Desta feita, pode-se representá-la, de forma simplificada, por $\partial l(\boldsymbol{\zeta}_i)/\partial\boldsymbol{\zeta}_i$. Fazendo-se

$$\frac{\partial l(\hat{\boldsymbol{\zeta}}_i)}{\partial a_i} = \frac{\partial l(\hat{\boldsymbol{\zeta}}_i)}{\partial b_i} = \frac{\partial l(\hat{\boldsymbol{\zeta}}_i)}{\partial c_i} = 0,$$

utilizando-se a notação

$$\begin{aligned}L_1 &= \frac{\partial l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial a_i} & L_{11} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial a_i^2} & L_{12} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial a_i \partial b_i} & L_{13} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial a_i \partial c_i}, \\ L_2 &= \frac{\partial l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial b_i} & L_{21} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial b_i \partial a_i} & L_{22} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial b_i^2} & L_{23} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial b_i \partial c_i}, \\ L_3 &= \frac{\partial l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial c_i} & L_{31} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial c_i \partial a_i} & L_{32} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial c_i \partial b_i} & L_{33} &= \frac{\partial^2 l(\hat{\boldsymbol{\zeta}}_i^{(0)})}{\partial c_i^2},\end{aligned}$$

e desprezando-se os restos $\mathbf{R}_{\mathbf{a}_i}(\hat{\boldsymbol{\zeta}}_i^{(0)})$, $\mathbf{R}_{\mathbf{b}_i}(\hat{\boldsymbol{\zeta}}_i^{(0)})$ e $\mathbf{R}_{\mathbf{c}_i}(\hat{\boldsymbol{\zeta}}_i^{(0)})$, tem-se

$$\begin{aligned}0 &= L_1 + L_{11}\Delta\hat{\mathbf{a}}_i^{(0)} + L_{12}\Delta\hat{\mathbf{b}}_i^{(0)} + L_{13}\Delta\hat{\mathbf{c}}_i^{(0)}, \\ 0 &= L_2 + L_{12}\Delta\hat{\mathbf{a}}_i^{(0)} + L_{22}\Delta\hat{\mathbf{b}}_i^{(0)} + L_{23}\Delta\hat{\mathbf{c}}_i^{(0)}, \\ 0 &= L_3 + L_{31}\Delta\hat{\mathbf{a}}_i^{(0)} + L_{32}\Delta\hat{\mathbf{b}}_i^{(0)} + L_{33}\Delta\hat{\mathbf{c}}_i^{(0)}.\end{aligned}$$

Agora, escrevendo-se esse resultado em forma matricial, tem-se

$$-\begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \begin{pmatrix} \Delta\hat{\mathbf{a}}_i^{(0)} \\ \Delta\hat{\mathbf{b}}_i^{(0)} \\ \Delta\hat{\mathbf{c}}_i^{(0)} \end{pmatrix}.$$

Resolvendo-se o sistema para $\Delta\hat{\boldsymbol{\zeta}}_i^{(0)}$, tem-se

$$\begin{pmatrix} \Delta \widehat{\mathbf{a}}_i^{(0)} \\ \Delta \widehat{\mathbf{b}}_i^{(0)} \\ \Delta \widehat{\mathbf{c}}_i^{(0)} \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{21} & L_{22} & L_{23} \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix},$$

e finalmente, por (4.19)

$$\begin{pmatrix} \widehat{\mathbf{a}}_i^{(1)} \\ \widehat{\mathbf{b}}_i^{(1)} \\ \widehat{\mathbf{c}}_i^{(1)} \end{pmatrix} = \begin{pmatrix} \widehat{\mathbf{a}}_i^{(0)} \\ \widehat{\mathbf{b}}_i^{(0)} \\ \widehat{\mathbf{c}}_i^{(0)} \end{pmatrix} - \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{21} & L_{22} & L_{23} \end{pmatrix}^{-1} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}.$$

Após obtido $\widehat{\boldsymbol{\zeta}}_i^{(1)}$, este é considerado um novo ponto inicial para a obtenção de $\widehat{\boldsymbol{\zeta}}_i^{(2)}$, e assim por diante. Este processo é repetido até que algum critério de parada seja alcançado. Por exemplo, até que $\Delta \widehat{\boldsymbol{\zeta}}_i^{(t)} = \widehat{\boldsymbol{\zeta}}_i^{(t)} - \widehat{\boldsymbol{\zeta}}_i^{(t-1)}$ seja suficientemente pequeno ou que um número pré-definido, t_{max} , de iterações seja cumprido.

Neste mesmo contexto, tem-se que as expressões L_k , $k = 1, 2, 3$ são dados por (4.13) a (4.15), respectivamente e as expressões L_{kl} , $k, l = 1, 2, 3$, são obtidas de

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} &= \sum_{j=1}^n \left\{ \left[\frac{\partial}{\partial \boldsymbol{\zeta}_i} \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \right] \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right)' + \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \left(\frac{\partial^2 P_{ji}}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} \right) \right\} \\ &= \sum_{j=1}^n \left\{ \left[\frac{\partial \nu_{ji}}{\partial \boldsymbol{\zeta}_i} \right] \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right)' + \nu_{ji} \left(\frac{\partial^2 P_{ji}}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} \right) \right\}, \end{aligned} \quad (4.20)$$

em que

$$\nu_{ji} = \frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \quad (4.21)$$

e

$$\begin{aligned} \frac{\partial \nu_{ji}}{\partial \boldsymbol{\zeta}_i} &= \frac{\partial}{\partial \boldsymbol{\zeta}_i} \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \\ &= \frac{1}{(P_{ji} Q_{ji})^2} \left\{ -P_{ji} Q_{ji} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) - (\nu_{ji} - P_{ji}) \left(\frac{\partial P_{ji} Q_{ji}}{\partial \boldsymbol{\zeta}_i} \right) \right\} \\ &= -\frac{1}{(P_{ji} Q_{ji})^2} \left\{ P_{ji} Q_{ji} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) + (\nu_{ji} - P_{ji}) \left[\left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) - 2P_{ji} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) \right] \right\} \\ &= -\frac{1}{(P_{ji} Q_{ji})^2} \{ P_{ji} Q_{ji} + (\mathbf{u}_{ji} - P_{ji})(1 - 2P_{ji}) \} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) \\ &= -\frac{1}{(P_{ji} Q_{ji})^2} (\mathbf{u}_{ji} - P_{ji})^2 \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) \\ &= -\nu_{ji}^2 \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right). \end{aligned} \quad (4.22)$$

Nesta mesma linha, vale ressaltar que a última igualdade segue do fato que $\mathbf{u}_{ji} = \mathbf{u}_{ji}^2$.

Por fim, complementando o exposto, considerando-se $\hat{\zeta}_i^{(t)}$ a estimativa de ζ_i na iteração t , então na iteração $t + 1$ do algoritmo de Newton-Raphson tem-se

$$\hat{\zeta}_i^{(t+1)} = \hat{\zeta}_i^{(t)} - [\mathbf{H}(\hat{\zeta}_i^{(t)})]^{-1} \mathbf{h}(\hat{\zeta}_i^{(t)}), \quad (4.23)$$

em que

$$\begin{aligned} \mathbf{h}(\zeta_i) &\equiv \frac{\partial \ln L(\zeta)}{\partial \zeta_i} \\ &= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} (P_{ji}^* Q_{ji}^*) \mathbf{h}_{ji} \\ &= \sum_{j=1}^n (\mathbf{u}_{ji} - P_{ji}) W_{ji} \mathbf{h}_{ji}, \end{aligned} \quad (4.24)$$

e

$$\begin{aligned} \mathbf{H}(\zeta_i) &\equiv \frac{\partial \ln L(\zeta)}{\partial \zeta_i \partial \zeta_i'} \\ &= \sum_{j=1}^n \left\{ \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) (P_{ji}^* Q_{ji}^*) \mathbf{H}_{ji} - \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right)^2 (P_{ji}^* Q_{ji}^*)^2 \mathbf{h}_{ji} \mathbf{h}_{ji}' \right\} \\ &= \sum_{j=1}^n (\mathbf{u}_{ji} - P_{ji}) W_{ji} \{ \mathbf{H}_{ji} - (\mathbf{u}_{ji} - P_{ji}) W_{ji} \mathbf{h}_{ji} \mathbf{h}_{ji}' \}, \end{aligned} \quad (4.25)$$

com

$$\begin{aligned} \mathbf{h}_{ji} &= (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) \\ &= \begin{pmatrix} D(1 - c_i)(\theta_j - b_i) \\ -Da_i(1 - c_i) \\ \frac{1}{P_{ji}^*} \end{pmatrix}, \end{aligned} \quad (4.26)$$

e

$$\begin{aligned} \mathbf{H}_{ji} &= (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial^2 P_{ji}}{\partial \zeta_i \partial \zeta_i'} \right) \\ &= \begin{pmatrix} D^2(1 - c_i)(\theta_j - b_i)^2(1 - 2P_{ji}^*) & \cdot & \cdot \\ -Da_i(1 - c_i)\{1 + Da_i(\theta_j - b_i)(1 - 2P_{ji}^*)\} & D^2 a_i^2(1 - c_i)(1 - 2P_{ji}^*) & \cdot \\ -D(\theta_j - b_j) & Da_i & 0 \end{pmatrix} \end{aligned} \quad (4.27)$$

Para obter as expressões (4.24) e (4.25), utilizadas em (4.23), verifica-se que, de (4.20), tem-se

$$\frac{\partial \ln L(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}'_i} = \sum_{j=1}^n \left\{ \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \left(\frac{\partial^2 P_{ji}}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}'_i} \right) - \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right)^2 \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right)' \right\} \quad (4.28)$$

Desta feita, tem-se as seguintes expressões

$$\begin{aligned} \frac{\partial P_{ji}^* Q_{ji}^*}{\partial \alpha_i} &= (1 - 2P_{ji}^*) \frac{\partial P_{ji}^*}{\partial \alpha_i}, \quad \alpha_i \in \{a_i, b_i, c_i\}, \\ \frac{\partial^2 P_{ji}}{\partial a_i^2} &= D^2 (1 - c_i) (\theta_j - b_i)^2 P_{ji}^* Q_{ji}^* (1 - 2P_{ji}^*), \\ \frac{\partial^2 P_{ji}}{\partial a_i \partial b_i} &= -D (1 - c_i) P_{ji}^* Q_{ji}^* \{1 + D a_i (\theta_j - b_i) (1 - 2P_{ji}^*)\}, \\ \frac{\partial^2 P_{ji}}{\partial a_i \partial c_i} &= -D (\theta_j - b_i) P_{ji}^* Q_{ji}^*, \\ \frac{\partial^2 P_{ji}}{\partial b_i^2} &= D^2 a_i^2 (1 - c_i) P_{ji}^* Q_{ji}^* (1 - 2P_{ji}^*), \\ \frac{\partial^2 P_{ji}}{\partial b_i \partial c_i} &= D a_i P_{ji}^* Q_{ji}^*, \\ \frac{\partial^2 P_{ji}}{\partial c_i^2} &= \frac{\partial Q_{ji}^*}{\partial c_i} = 0. \end{aligned}$$

Portanto, com estas expressões obtêm-se $\partial^2 P_{ji} / (\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}'_i)$ e, por consequência, obtendo-se as expressões (4.24) e (4.25), já apresentadas nesta seção.

Na sequência, apresenta-se o método “Scoring” de Fisher, método de estimação iterativo, normalmente, utilizado na TRI.

4.2.1.2 Aplicação do método “Scoring” de Fisher

No que concerne a aplicação do método “Scoring” de Fisher, basta substituir os componentes da matriz de derivadas segundas utilizadas no processo iterativo de Newton-Raphson pelos seus valores esperados. Vale ressaltar que a variável U_{ji} só pode assumir dois valores: 1 (um), com probabilidade P_{ji} e 0 (zero) com probabilidade Q_{ji} . Desta feita, tem-se que $U_{ji} \sim \text{Bernoulli}(P_{ji})$.

Por definição, segue-se que $E(U_{ji}) = P_{ji}$ e $E(U_{ji} - P_{ji})^2 = \text{Var}(U_{ji}) = P_{ji} Q_{ji}$. Assim, de (4.25), tem-se que

$$\begin{aligned} \Delta(\boldsymbol{\zeta}_i) &\equiv E(\mathbf{H}(\boldsymbol{\zeta}_i)) \\ &= \sum_{j=1}^N \{E(U_{ji} - P_{ji}) \mathbf{W}_{ji} \mathbf{H}_{ji} - E(U_{ji} - P_{ji})^2 \mathbf{W}_{ji}^2 \mathbf{h}_{ji} \mathbf{h}'_{ji}\} \\ &= \sum_{j=1}^N \{-P_{ji} Q_{ji} \mathbf{W}_{ji}^2 \mathbf{h}_{ji} \mathbf{h}'_{ji}\} \\ &= - \sum_{j=1}^N \{P_{ji}^* Q_{ji}^* \mathbf{W}_{ji}^2 \mathbf{h}_{ji} \mathbf{h}'_{ji}\}. \end{aligned} \quad (4.29)$$

Desta feita, de (4.29), a expressão para as estimativas de ζ_i na iteração $t + 1$ é dada por

$$\hat{\zeta}_i^{(t+1)} = \hat{\zeta}_i^{(t)} - [\Delta(\hat{\zeta}_i^{(t)})]^{-1} \mathbf{h}(\hat{\zeta}_i^{(t)}). \quad (4.30)$$

Enfatiza-se, na sequência, anteriormente à discussão do Método da máxima verossimilhança conjuntado e do Método da máxima verossimilhança marginal na TRI, a estimação das habilidades.

4.2.1.2.1 Estimação das habilidades na TRI

Nesta etapa da pesquisa enfatiza-se a importância de tratar da estimação das habilidades quando os parâmetros dos itens são conhecidos. Nesta mesma linha, [Andrade et al. \(2000\)](#) alegam que na prática, essa situação ocorre quando os itens já foram calibrados (estimados) em outros testes. Como a calibração dos itens deve ser feita com um número grande de indivíduos, a estimação das habilidades de um grupo pequeno de indivíduos é mais confiável se forem utilizados itens já calibrados.

Considerando-se a independência entre as respostas de diferentes indivíduos (S1) e a independência local (S2), pode-se escrever a função de log-verossimilhança como em (4.5), agora como função de θ e não de ζ . Logo, facilmente, tem-se

$$\ln L(\theta) = \sum_{j=1}^n \sum_{i=1}^I \{ \mathbf{u}_{ji} \ln P_{ji} + (1 - \mathbf{u}_{ji}) \ln Q_{ji} \}. \quad (4.31)$$

O EMV de θ_j é o valor que maximiza a verossimilhança, ou equivalentemente, é a solução da equação

$$\frac{\partial \ln L(\theta)}{\partial \theta_j} = \mathbf{0}, \quad j = 1, \dots, n. \quad (4.32)$$

Na sequência, de (4.32), tem-se

$$\begin{aligned} \frac{\partial \ln L(\theta)}{\partial \theta_i} &= \sum_{j=1}^n \left\{ \frac{\partial \ln(L(\theta))}{\partial P_{ji}} \frac{\partial P_{ji}}{\partial \theta_i} \right\} \\ &= \sum_{j=1}^n \left\{ \mathbf{u}_{ji} \frac{\partial(\ln P_{ji})}{\partial \theta_i} + (1 - \mathbf{u}_{ji}) \frac{\partial(\ln Q_{ji})}{\partial \theta_i} \right\} \\ &= \sum_{j=1}^n \left\{ \mathbf{u}_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \theta_i} \right) + (1 - \mathbf{u}_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial P_{ji}}{\partial \theta_i} \right) \right\} \\ &= \sum_{j=1}^n \left\{ \mathbf{u}_{ji} \frac{1}{P_{ji}} - (1 - \mathbf{u}_{ji}) \frac{1}{Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_i} \right) \\ &= \sum_{j=1}^n \left\{ \frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_i} \right) \end{aligned} \quad (4.33)$$

$$= \sum_{j=1}^n \left\{ (\mathbf{u}_{ji} - P_{ji}) \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_i} \right), \quad (4.34)$$

em que (4.34) segue de (4.7). Como

$$\frac{\partial P_{ji}}{\partial \theta_j} = Da_i(1 - c_i)P_{ji}^*Q_{ji}^* \quad (4.35)$$

obtêm-se

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\theta})}{\partial \theta_j} &= \sum_{i=1}^I \left\{ (\mathbf{u}_{ji} - P_{ji}) Da_i(1 - c_i) P_{ji}^* Q_{ji}^* \frac{W_{ij}}{P_{ji}^* Q_{ji}^*} \right\} \\ &= D \sum_{i=1}^I \{ a_i(1 - c_i)(\mathbf{u}_{ji} - P_{ji}) W_{ij} \}. \end{aligned} \quad (4.36)$$

Portanto, a equação de estimação (4.32) para θ_j , $j = 1, \dots, n$, é dada por

$$\theta_j : D \sum_{i=1}^I \{ a_i(1 - c_i)(\mathbf{u}_{ji} - P_{ji}) W_{ij} \} = 0. \quad (4.37)$$

Por fim, vale ressaltar que a equação (4.37) não apresenta solução explícita para θ_j e, por isso, precisa-se de algum método iterativo para obter as estimativas desejadas. Nesta mesma linha, enfatiza-se que os processos iterativos de Newton-Raphson e “Scoring” de Fisher, devidamente apresentados no escopo deste capítulo, são algoritmos que possibilitam obter as estimativas de (4.37). Para mais detalhes, bem como as expressões necessárias para aplicação destes processos iterativos, ver Hambleton e Swaminathan (1985), Baker e Kim (2004), Embretson e Reise (2000) e Andrade *et al.* (2000).

4.2.1.2.1.1 Aplicação do algoritmo Newton-Raphson

Analogamente ao que foi feito na Seção 4.2.1.1, e considerando-se $\hat{\boldsymbol{\theta}}_j^{(t)}$ a estimativa de $\boldsymbol{\theta}_j$ na iteração t , então na iteração $t + 1$ do algoritmo de Newton-Raphson tem-se

$$\hat{\boldsymbol{\theta}}_j^{(t+1)} = \hat{\boldsymbol{\theta}}_j^{(t)} - [\mathbf{H}(\hat{\boldsymbol{\theta}}_j^{(t)})]^{-1} \mathbf{h}(\hat{\boldsymbol{\theta}}_j^{(t)}), \quad (4.38)$$

em que

$$\begin{aligned} \mathbf{h}(\boldsymbol{\theta}_j) &\equiv \frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j} \\ &= \sum_{i=1}^I \left\{ (\mathbf{u}_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} (P_{ji}^* Q_{ji}^*) \mathbf{h}_{ji} \\ &= \sum_{i=1}^I (\mathbf{u}_{ji} - P_{ji}) W_{ji} \mathbf{h}_{ji}, \end{aligned} \quad (4.39)$$

e

$$\begin{aligned}
\mathbf{H}(\boldsymbol{\theta}_j) &\equiv \frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j^2} \\
&= \sum_{i=1}^I \left\{ \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji}Q_{ji}} \right) (P_{ji}^*Q_{ji}^*) \mathbf{H}_{ji} - \left(\frac{\mathbf{u}_{ji} - P_{ji}}{P_{ji}Q_{ji}} \right)^2 (P_{ji}^*Q_{ji}^*)^2 \mathbf{h}_{ji}^2 \right\} \\
&= \sum_{i=1}^I (\mathbf{u}_{ji} - P_{ji}) W_{ji} \{ \mathbf{H}_{ji} - (\mathbf{u}_{ji} - P_{ji}) W_{ji} \mathbf{h}_{ji}^2 \}, \tag{4.40}
\end{aligned}$$

com

$$\begin{aligned}
\mathbf{h}_{ji} &= (P_{ji}^*Q_{ji}^*)^{-1} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\theta}_j} \right) \\
&= Da_i(1 - c_i) \tag{4.41}
\end{aligned}$$

e

$$\begin{aligned}
\mathbf{H}_{ji} &= (P_{ji}^*Q_{ji}^*)^{-1} \left(\frac{\partial^2 P_{ji}}{\partial \boldsymbol{\theta}_j^2} \right) \\
&= D^2 a_i^2 (1 - c_i) (1 - 2P_{ji}^*). \tag{4.42}
\end{aligned}$$

4.2.1.2.1.2 Aplicação do método “Scoring” de Fisher

No que concerne a aplicação do método “Scoring” de Fisher, basta substituir os componentes da matriz de derivadas segundas utilizadas no processo iterativo de Newton-Raphson pelos seus valores esperados. Por (4.40), tem-se que

$$\begin{aligned}
\Delta(\boldsymbol{\theta}_j) &\equiv E(\mathbf{H}(\boldsymbol{\theta}_j)) \\
&= \sum_{i=1}^I \{ E(U_{ji} - P_{ji}) W_{ji} \mathbf{H}_{ji} - E(U_{ji} - P_{ji})^2 \mathbf{W}_{ji}^2 \mathbf{h}_{ji}^2 \} \\
&= \sum_{i=1}^I \{ -P_{ji} Q_{ji} \mathbf{W}_{ji}^2 \mathbf{h}_{ji}^2 \} \\
&= - \sum_{i=1}^I \{ P_{ji}^* Q_{ji}^* \mathbf{W}_{ji}^2 \mathbf{h}_{ji}^2 \}. \tag{4.43}
\end{aligned}$$

Desta feita, de (4.43), a expressão para as estimativas de $\boldsymbol{\theta}_j$, $j = 1, \dots, n$ na iteração $t + 1$ é dada por

$$\widehat{\boldsymbol{\theta}}_j^{(t+1)} = \widehat{\boldsymbol{\theta}}_j^{(t)} - [\Delta(\widehat{\boldsymbol{\theta}}_j^{(t)})]^{-1} \mathbf{h}(\widehat{\boldsymbol{\theta}}_j^{(t)}). \tag{4.44}$$

4.2.2 Método da Máxima Verossimilhança Conjunta

Método da máxima verossimilhança conjunta: (JML, do inglês Joint Maximum Likelihood): “modela as probabilidades de padrão de resposta a partir da utilização de estimativas provisórias de estimativas de nível do traço” (Embretson e Reise, 2000, p. 208). Estas estimativas provisórias são, então, atualizadas e sucessivamente melhoradas com a estimação sequencial dos parâmetros dos itens, ou seja, dadas as estimativas das habilidades, estima-se os parâmetros dos itens. Posteriormente, dadas as estimativas atualizadas destes parâmetros, atualiza-se as estimativas das habilidades.

Vale ressaltar que o método da máxima verossimilhança conjunta, na TRI, consiste em um dos casos mais comuns, visto que, nem os parâmetros dos itens e nem as habilidades são conhecidos. Nas expressões (4.16) a (4.18) e (4.37) tem-se as equações de estimação para os parâmetros dos itens e habilidades.

De acordo com Andrade *et al.* (2000), a principal diferença da estimação conjunta se dá no tratamento da métrica (escala) em que todos os parâmetros são estimados. Os autores completam que na estimação conjunta não há uma métrica definida e, portanto, deve-se estabelecê-la. A explicação formal para a necessidade do estabelecimento da métrica dos parâmetros consiste em um problema denominado *falta de identificabilidade do modelo*. Essa não-identificabilidade ocorre porque mais de um conjunto de parâmetros produz o mesmo valor no ML3, e conseqüentemente, na verossimilhança.

Conforme já apresentado no Capítulo 3, se $\theta_j^* = \alpha\theta_j + \beta$, $b_i^* = \alpha b_i + \beta$, $a_i^* = a_i/\alpha$ e $c_i^* = c_i$, em α e β são constantes reais com $\alpha > 0$, então

$$\begin{aligned} P(U_{ji} = 1 | \theta_j^*, \zeta_i^*) &= c_i^* + (1 - c_i^*) \{1 + \exp[-Da_i^*(\theta_j^* - b_i^*)]\}^{-1} \\ &= c_i + (1 - c_i) \left\{1 + \exp \left[-D \frac{a_i}{\alpha} (\alpha\theta_j + \beta - (\alpha b_i + \beta)) \right]\right\}^{-1} \\ &= c_i + (1 - c_i) \{1 + \exp[-Da_i(\theta_j - b_i)]\}^{-1} \\ &= P(U_{ji} = 1 | \theta_j, \zeta_i). \end{aligned}$$

Tal não-identificabilidade pode ser eliminada de várias formas, como fixando alguns valores para as habilidades, por exemplo. Outra maneira de lidar com a falta de identificabilidade, é feita através da especificação dos valores populacionais da média, μ , e desvio-padrão, σ , da escala das habilidades, como já mencionado na Seção 4.5.1.1.1, é usual fixar $\mu = 0$ e $\sigma = 1$.

Na Tabela 4.1, Seção 4.2.3.1 deste capítulo, tem-se as principais vantagens e desvantagens da utilização deste modelo na TRI. Para um aprofundamento das discussões referentes ao método de estimação, destacam-se as abordagens de Hambleton e Swaminathan (1985), Baker e Kim (2004) e Embretson e Reise (2000).

Apresenta-se, na sequência, o método da máxima verossimilhança marginal.

4.2.3 Método da Máxima Verossimilhança Marginal

De acordo com Embretson e Reise (2000, p. 210), o Método da máxima verossimilhança marginal (MML, do inglês Marginal Maximum Likelihood): “modela as probabilidades de padrão de resposta em decorrência da amostragem aleatória de uma população com uma distribuição conhecida do nível do traço latente”. O método de estimação iterativa proposto por Bock e Aitkin (1981) é baseado no algoritmo EM.

Andersen (1980) argumenta que se considerar uma população π composta por n indivíduos com habilidades θ_j , $j = 1, \dots, n$, e construir-se a distribuição de frequência acumulada

$G(\theta) = (\text{número de } j : \theta_j < \theta)/n$, então, se n for suficientemente grande os θ_j estarão bastante próximos, de forma que $G(\theta)$ pode ser aproximada por uma distribuição contínua. A densidade $g(\theta)$, relativa à $G(\theta)$, pode realmente ser considerada a função densidade para θ no experimento de retirar um indivíduo ao acaso da população π e observar seu parâmetro θ . Nesta mesma linha, vale ressaltar que, quando atribui-se uma distribuição de probabilidade para θ *não se está aplicando nenhum argumento bayesiano*. A distribuição de θ realmente existe, no sentido explicado acima, como a densidade relativa à distribuição $G(\theta)$.

Bock e Lieberman (1970) elaboraram um artifício para eliminar as habilidades da verossimilhança que consiste em marginalizar a verossimilhança, integrando-a com relação à distribuição da habilidade. De forma geral, considera-se que as habilidades, θ_j , $j = 1, \dots, n$, são realizações de uma variável aleatória θ , com distribuição contínua e função densidade de probabilidade (f.d.p.) $g(\theta|\boldsymbol{\eta})$, duplamente diferenciável, com as componentes de $\boldsymbol{\eta}$ conhecidas e finitas. Para o caso em que θ tem distribuição normal, tem-se que $\boldsymbol{\eta} = (\mu, \sigma^2)$, em que μ é a média e σ^2 a variância das habilidades dos indivíduos de π . Desta feita, para que os itens sejam estimados na métrica (0,1), deve-se adotar $\mu = 0$ e $\sigma = 1$.

No que concerne o método da máxima verossimilhança marginal, na sequência, apresenta-se a abordagem de Bock e Lieberman (1970) para este método. Vale ressaltar que esta apresenta algumas vantagens em comparação ao método da máxima verossimilhança conjunta.

4.2.3.1 Abordagem de Bock & Lieberman

Condicionalmente a $\boldsymbol{\zeta}$ e $\boldsymbol{\eta}$, seja a distribuição marginal de \mathbf{U}_j , dada por (Bock e Lieberman, 1970):

$$\begin{aligned} P(\mathbf{u}_j|\boldsymbol{\zeta}, \boldsymbol{\eta}) &= \int_{\mathbb{R}} P(\mathbf{u}_j|\boldsymbol{\theta}, \boldsymbol{\zeta}, \boldsymbol{\eta})g(\theta|\boldsymbol{\eta})d\theta \\ &= \int_{\mathbb{R}} P(\mathbf{u}_j|\boldsymbol{\theta}, \boldsymbol{\zeta})g(\theta|\boldsymbol{\eta})d\theta, \end{aligned} \quad (4.45)$$

em que a última igualdade usa-se a distribuição de \mathbf{U}_j , não é função de $\boldsymbol{\eta}$ e \mathbb{R} representa o conjunto dos números reais. Usando-se a independência entre as respostas de diferentes indivíduos, pode-se escrever a probabilidade associada ao vetor de respostas $\mathbf{U}_.$ como

$$P(\mathbf{u}_j|\boldsymbol{\zeta}, \boldsymbol{\eta}) = \prod_{j=1}^n P(\mathbf{u}_j|\boldsymbol{\zeta}, \boldsymbol{\eta}). \quad (4.46)$$

Vale ressaltar que, embora a verossimilhança possa ser escrita como (4.46), tem sido frequente utilizar a abordagem de *Padrões de Resposta*. Como tem-se I itens ao total, com 2 possíveis repostas para cada item (0 ou 1), há $S = 2^I$ possíveis repostas (padrões de resposta).

De acordo com Bock e Lieberman (1970), segundo as notações matemáticas expresas pelos autores Andrade *et al.* (2000), seja r_j o número de ocorrências distintas do padrão de resposta j , e ainda $s \leq \min(n, S)$ o número de padrões de resposta com $r_j > 0$. Segue-se

$$\sum_{j=1}^s r_j = n. \quad (4.47)$$

Pela independência entre as respostas dos diferentes indivíduos, tem-se que os dados seguem uma distribuição *Multinomial*, isto é,

$$L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \frac{n!}{\prod_{j=1}^s r_j!} \prod_{j=1}^s [P(\mathbf{u}_j | \boldsymbol{\zeta}, \boldsymbol{\eta})]^{r_j}, \quad (4.48)$$

e, portanto, a log-verossimilhança é

$$\ln L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \ln \left\{ \frac{n!}{\prod_{j=1}^s r_j!} \right\} + \sum_{j=1}^s r_j \ln [P(\mathbf{u}_j | \boldsymbol{\zeta}, \boldsymbol{\eta})]. \quad (4.49)$$

Desta feita, as equações de estimação para os parâmetros dos itens são dadas por

$$\frac{\partial \ln L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \zeta_i} = 0, \quad i = 1, \dots, I, \quad (4.50)$$

com

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \zeta_i} &= \frac{\partial}{\partial \zeta_i} \left\{ \sum_{j=1}^s r_j \ln [P(\mathbf{u}_j | \boldsymbol{\zeta}, \boldsymbol{\eta})] \right\} \\ &= \sum_{j=1}^s r_j \frac{1}{P(\mathbf{u}_j | \boldsymbol{\zeta}, \boldsymbol{\eta})} \frac{\partial P(\mathbf{u}_j | \boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}}. \end{aligned} \quad (4.51)$$

No entanto,

$$\begin{aligned} \frac{\partial P(\mathbf{u}_j | \boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \zeta_i} &= \frac{\partial}{\partial \zeta_i} \int_{\mathbb{R}} P(\mathbf{u}_j | \theta, \boldsymbol{\zeta}) g(\theta | \boldsymbol{\eta}) d\theta \\ &= \int_{\mathbb{R}} \left(\frac{\partial}{\partial \zeta_i} P(\mathbf{u}_j | \theta, \boldsymbol{\zeta}) \right) g(\theta | \boldsymbol{\eta}) d\theta \\ &= \int_{\mathbb{R}} \left(\frac{\partial}{\partial \zeta_i} \prod_{l=1}^I P(\mathbf{u}_{jl} | \theta, \zeta_l) \right) g(\theta | \boldsymbol{\eta}) d\theta \\ \frac{\partial P(\mathbf{u}_j | \boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \zeta_i} &= \int_{\mathbb{R}} \left(\prod_{l \neq i}^I P(\mathbf{u}_{jl} | \theta, \zeta_l) \right) \left(\frac{\partial}{\partial \zeta_i} P(\mathbf{u}_{ji} | \theta, \zeta_i) \right) g(\theta | \boldsymbol{\eta}) d\theta \\ &= \int_{\mathbb{R}} \left(\frac{\partial P(\mathbf{u}_{ji} | \theta, \zeta_i) / \partial \zeta_i}{P(\mathbf{u}_{ji} | \zeta_i)} \right) P(\mathbf{u}_j | \theta, \boldsymbol{\zeta}) g(\theta | \boldsymbol{\eta}) d\theta. \end{aligned} \quad (4.52)$$

A ordem da derivada e da integral em (4.52) pôde ser permutada com base no Teorema da Convergência Dominada de Lebesgue (Teicher e Chow, 1978).

Agora, reescrevendo $P(\mathbf{u}_{ji} | \theta, \zeta_i)$ como em (4.3), tem-se:

$$\begin{aligned}
\frac{\partial}{\partial \zeta_i} P(\mathbf{u}_{ji}|\theta, \zeta_i) &= \frac{\partial}{\partial \zeta_i} \left(P_i^{\mathbf{u}_{ji}} Q_i^{1-\mathbf{u}_{ji}} \right) \\
&= \mathbf{u}_{ji} P_i^{\mathbf{u}_{ji}-1} \left(\frac{\partial P_i}{\partial \zeta_i} \right) Q_i^{1-\mathbf{u}_{ji}} + (1 - \mathbf{u}_{ji}) Q_i^{-\mathbf{u}_{ji}} \left(-\frac{\partial}{\partial \zeta_i} P_i \right) P_i^{\mathbf{u}_{ji}} \\
&= \left(\mathbf{u}_{ji} P_i^{\mathbf{u}_{ji}-1} Q_i^{1-\mathbf{u}_{ji}} - (1 - \mathbf{u}_{ji}) Q_i^{-\mathbf{u}_{ji}} P_i^{\mathbf{u}_{ji}} \right) \frac{\partial P_i}{\partial \zeta_i}.
\end{aligned}$$

Nota-se agora que o termo entre parênteses vale 1 quando $\mathbf{u}_{ji} = 1$ e vale -1 quando $\mathbf{u}_{ji} = 0$, portanto, pode-se reescrevê-lo como $(-1)^{\mathbf{u}_{ji}+1}$. Portanto,

$$\frac{\partial}{\partial \zeta_i} P(\mathbf{u}_{ji}|\theta, \zeta_i) = (-1)^{\mathbf{u}_{ji}+1} \left(\frac{\partial P_i}{\partial \zeta_i} \right). \quad (4.54)$$

Nota-se agora que

$$\frac{(-1)^{\mathbf{u}_{ji}+1} P_i Q_i}{P_i^{\mathbf{u}_{ji}} Q_i^{1-\mathbf{u}_{ji}}} \begin{cases} Q_i & \text{se } \mathbf{u}_{ji} = 1 \\ -P_i & \text{se } \mathbf{u}_{ji} = 0, \end{cases} \quad (4.55)$$

pode-se reescrever este termo como $\mathbf{u}_{ji} - P_i$. Daí, segue que (4.53) pode ser escrita como

$$\frac{\partial P(\mathbf{u}_j|\zeta, \eta)}{\partial \zeta_i} = \int_{\mathbb{R}} \left[\frac{(\mathbf{u}_{ji} - P_i)}{P_i Q_i} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right] P(\mathbf{u}_j|\theta, \zeta) g(\theta|\eta) d\theta. \quad (4.56)$$

Por conveniência, considera-se a seguinte ponderação

$$W_i = \frac{P_i^* Q_i^*}{P_i Q_i}, \quad (4.57)$$

em que

$$P_i^* = \{1 + \exp\{-Da_i(\theta - b_i)\}\}^{-1} \quad \text{e} \quad Q_i^* = 1 - P_i^*. \quad (4.58)$$

Desta feita, pode-se escrever a expressão (4.56) da seguinte maneira

$$\frac{\partial P(\mathbf{u}_j|\zeta, \eta)}{\partial \zeta_i} = \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] P(\mathbf{u}_j|\theta, \zeta) g(\theta|\eta) d\theta. \quad (4.59)$$

Usando a notação

$$g_j^*(\theta) \equiv g(\theta|\mathbf{u}_j, \zeta, \eta) = \frac{P(\mathbf{u}_j|\theta, \zeta) g(\theta|\eta)}{P(\mathbf{u}_j|\zeta, \eta)}, \quad (4.60)$$

tem-se que a função de verossimilhança (4.51) pode ser escrita como

$$\frac{\partial \ln L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\zeta}_i} = \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) \left(\frac{\partial P_i}{\partial \boldsymbol{\zeta}_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta. \quad (4.61)$$

Resta agora a obtenção das equações específicas para cada parâmetro do vetor $\boldsymbol{\zeta}_i = (a_i, b_i, c_i)'$. As expressões para as derivadas de P_i são dadas por (4.10) a (4.12) com P_{ji}, Q_{ji}, P_{ji}^* e Q_{ji}^* substituídas por P_i, Q_i, P_i^* e Q_i^* , respectivamente.

Para se obter a equação de estimação para o parâmetro de discriminação, a_i da expressão (4.61) tem-se que

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial a_i} &= \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) \left(\frac{\partial P_i}{\partial a_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ &= \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) D(1 - c_i) (\theta - b_i) P_i^* Q_i^* \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ &= D(1 - c_i) \sum_{j=1}^s r_j \int_{\mathbb{R}} [(\mathbf{u}_{ji} - P_i) (\theta - b_i) W_i] g_j^*(\theta) d\theta. \end{aligned} \quad (4.62)$$

Para o parâmetro de dificuldade, b_i , tem-se

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial b_i} &= \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) \left(\frac{\partial P_i}{\partial b_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ &= \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) (-1) D a_i (1 - c_i) P_i^* Q_i^* \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ &= -D a_i (1 - c_i) \sum_{j=1}^s r_j \int_{\mathbb{R}} [(\mathbf{u}_{ji} - P_i) W_i] g_j^*(\theta) d\theta. \end{aligned} \quad (4.63)$$

Para o parâmetro de acerto ao acaso, c_i , tem-se

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial c_i} &= \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) \left(\frac{\partial P_i}{\partial c_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ &= \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) Q_i^* \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ &= \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) \frac{W_i}{P_i^*} \right] g_j^*(\theta) d\theta. \end{aligned} \quad (4.64)$$

Finalmente, as equações de estimação para os parâmetros a_i, b_i e c_i são, respectivamente

$$a_i : D(1 - c_i) \sum_{j=1}^s r_j \int_{\mathbb{R}} [(\mathbf{u}_{ji} - P_i) (\theta - b_i) W_i] g_j^*(\theta) d\theta = 0, \quad (4.65)$$

$$b_i : -Da_i(1 - c_i) \sum_{j=1}^s r_j \int_{\mathbb{R}} [(\mathbf{u}_{ji} - P_i)W_i] g_j^*(\theta) d\theta = 0, \quad (4.66)$$

$$c_i : \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[(\mathbf{u}_{ji} - P_i) \frac{W_i}{P_i^*} \right] g_j^*(\theta) d\theta = 0, \quad (4.67)$$

as quais não possuem solução explícita.

No que concerne os métodos iterativos utilizados para a estimação na máxima verossimilhança marginal, pode-se citar: *i)* Newton-Raphson; *ii)* Métodos da quadratura; *iii)* Abordagem de Bock e Aitkin; *iv)* Algoritmo EM.

4.3 Estimação: aplicação do algoritmo EM

Como já mencionado, anteriormente, devido as expressões (4.65), (4.66) e (4.67) não apresentarem soluções explícitas, necessita-se de um método iterativo que possibilite a estimação dos parâmetros de interesse.

O algoritmo EM é um método iterativo de estimação por máxima verossimilhança dos parâmetros de modelos de probabilidade na presença de variáveis latentes. Cada iteração do algoritmo consiste em dois passos: Esperança (E) e Maximização (M).

No caso da Teoria de Resposta ao Item, o objetivo é estimar o vetor de parâmetros ζ na presença das variáveis latentes θ . Assumindo que $\mathbf{u}_{..}$ representa o vetor de dados incompletos e $(\mathbf{u}_{..}, \theta)$ o vetor de dados completos. Sendo $f(\mathbf{u}_{..}, \theta | \zeta)$ a densidade conjunta dos dados completos e $\hat{\zeta}^{(k)}$ uma estimativa de θ na iteração t , então para a obtenção da estimativa $\hat{\zeta}^{(k+1)}$, faz-se uso dos seguintes passos:

Passo E: $E \ln\{f(\mathbf{u}_{..}, \theta | \zeta) | \mathbf{u}_{..}, \hat{\zeta}^{(k)}\}$, utilizando o método de quadratura. Para mais detalhes sobre o método de quadratura, ver [Andrade et al. \(2000\)](#).

Passo M: Calcular o valor de $\hat{\zeta}^{(k+1)}$ que maximiza a função do **Passo E**.

No Passo M a maximização pode ser feita via algoritmo Newton-Raphson ou *Scoring* de Fisher, a qual é feita, em separado, para cada vetor de parâmetros de cada item.

O método de máxima verossimilhança gera estimadores eficientes e não viciados e o erro-padrão de cada é dado pela raiz quadrada do inverso da matriz de informação, expressa por:

$$-E \left(\frac{\partial \ln L(\zeta, \eta)}{\partial \zeta_i \partial \zeta_i'} \right). \quad (4.68)$$

Nesta mesma linha, vale ressaltar que para aplicar o algoritmo EM, é necessário atribuir valores iniciais aos parâmetros. [Andrade et al. \(2000\)](#) enfatizam que para o parâmetro de acerto casual, c_i , pode-se considerar que o indivíduo pode escolher a qualquer item com a mesma probabilidade, sendo a estimativa inicial, dada por $c_i = 1/m_i$, em que m_i corresponde ao número de alternativas que o item apresenta. Para o parâmetro de discriminação, a_i , pode-se utilizar o coeficiente de correlação linear de Pearson entre a proporção de itens marcados corretamente e as respostas ao item i . O parâmetro de dificuldade pode ter como valor inicial a proporção de respostas corretas dadas ao item i .

Por fim, para um aprofundamento das discussões referentes ao método apresentado nesta seção, bem como a estrutura dos processos iterativos citados, ver [Hambleton e Swaminathan \(1985\)](#), [Baker e Kim \(2004\)](#) e [Embretson e Reise \(2000\)](#).

Na Tabela 4.1, a seguir, apresenta-se as principais vantagens e desvantagens da utilização dos métodos de máxima verossimilhança discutidos na estimação dos parâmetros do modelo de TRI apresentados neste capítulo.

Em se tratando dos métodos de máxima verossimilhança apresentados, [Embretson e Reise \(2000\)](#) destacam as principais vantagens e desvantagens dos mesmos, que estão resumidas na Tabela 4.1.

Tabela 4.1: Principais vantagens e desvantagens dos métodos de máxima verossimilhança para a TRI.

Modelo	Vantagens	Desvantagens
Máxima verossimilhança conjunta (MVC)	O algoritmo é facilmente programável; é apropriado para muitos modelos da TRI; é computacionalmente eficiente.	A estimativa dos parâmetros dos itens não possui propriedades desejáveis: a estimativa é viesada (tendenciosa) e inconsistente; o significado do erro padrão é questionável; Estimativas não são válidas para itens ou pessoas com escores perfeitos.
Máxima verossimilhança marginal (MVM)	Possui mais características desejáveis; é facilmente aplicável para muitos modelos da TRI; é eficiente para testes curtos e longos; prevê estimativas para valores extremos (escores perfeitos).	Um algoritmo efetivo para a estimação é difícil programar; uma distribuição deve ser assumida para o desempenho (traço), e esta distribuição pode ser especificada de forma inapropriada.
Máxima verossimilhança condicional (MVC)	Nenhuma distribuição é requerida para o traço; as estimativas dos parâmetros dos itens refletem o princípio da invariância dos parâmetros dos itens; os estimadores possuem propriedades desejáveis sob várias condições (ex.: consistência);	Possui aplicabilidade restrita; é aplicado somente para o modelo de Rasch; há perda de informação com a sua aplicação; problemas numéricos desenvolvem-se frequentemente para testes longos.

Fonte: Elaborado pelos autores (2016). Adaptada de [Embretson e Reise \(2000\)](#), segundo [Mafra \(2010\)](#).

Fundamentando-se na Tabela 4.1, o método mais aplicado para os modelos da TRI é a máxima verossimilhança marginal. Neste contexto, é importante citar [Andrade et al. \(2000\)](#), ao ressaltarem que o método da máxima verossimilhança marginal, desenvolvido por [Bock e Lieberman \(1970\)](#), possui algumas vantagens em comparação ao método da máxima verossimilhança conjunta. O método MVM objetiva realizar a estimação em duas fases: a primeira, que estima o parâmetro dos itens e, a segunda, que estima as habilidades, destacam os autores.

Em se tratando do processo de estimação dos parâmetros na TRI, a aplicação dos métodos de estimação apresentados no escopo deste capítulo, requer a utilização de ferramentas matemáticas bastante complexas, que necessitam de recursos computacionais. Neste contexto, na Seção 4.4, apresenta-se alguns dos programas computacionais existentes.

No que concerne à estimação bayesiana, [Andrade et al. \(2000\)](#) destacam que esta consiste em estabelecer distribuições *a priori* para os parâmetros de interesse, construir uma nova

função denominada distribuição *a posteriori* e estimar os parâmetros de interesse com base em alguma característica dessa distribuição. Ressalta-se, também, que a estimação bayesiana propõe alternativas para solucionar os problemas existentes nos métodos de máxima verossimilhança, a saber: erros na estimação dos itens que são respondidos da mesma forma por todos indivíduos (acertadamente ou erroneamente) e na estimação do seu desempenho.

4.4 Programas Computacionais

Existem vários programas computacionais, disponíveis, que facilitam as estimativas de parâmetros em modelos de TRI. Estes programas podem ser agrupados nas seguintes categorias (Hambleton e Swaminathan, 1985, p. 147):

1. Método de estimação da máxima verossimilhança condicional;
2. Método de estimação da máxima verossimilhança conjunta;
3. Método de estimação da máxima verossimilhança marginal;
4. Estimativa aproximada no modelo de três parâmetros (ML3).

Os programas computacionais e suas características estão resumidas na Tabela 4.2 (ver também Wright e Mead (1976) e Wingersky (1983)). Vale ressaltar que, dentre os programas computacionais existentes, além dos apresentados na Tabela 4.2, destacam-se, também: BILOG MG, PARSCALE, MULTILOG, RUMMFOLD, MUDFOLD, GGUM2004, Mirt e Itm do R.

Tabela 4.2: *Programas computacionais com breves descrições.*

Programa	Características	Referência
LOGIST5	<ul style="list-style-type: none"> • Adequado para o método de estimação da máxima verossimilhança conjunta com um, dois e três parâmetros. • Capacidade de lidar com respostas omitidas e itens não-alcançados (neste caso, a função de verossimilhança é modificada). 	Wingersky <i>et al.</i> (1982)
BICAL	<ul style="list-style-type: none"> • Adequado para métodos de estimação com um parâmetro. • Fornece testes para o ajuste de itens. 	Wright e Mead (1976)
ANCILLES; OGIVA	<ul style="list-style-type: none"> • Adequado para estimativa conjunta dos parâmetros dos itens e da habilidade em modelo de três parâmetros. • Baseado, parcialmente, no critério do qui-quadrado mínimo. • Fornece boas estimativas para um grande número de examinandos e itens. 	Urry (1976)
PML	<ul style="list-style-type: none"> • Adequado para o método de estimação da máxima verossimilhança conjunta com um parâmetro • Fornece boas estimativas para modelos de multi categorias. 	Gustafsson (1980)
BilogMG de 1997	<ul style="list-style-type: none"> • Adequado para o método de estimação da máxima verossimilhança marginal com um e dois parâmetros. 	Mislevy e Bock (1982)

Fonte: Elaborado pelos autores (2016). Adaptada de (Hambleton e Swaminathan, 1985, p. 148)

A implementação dos programas computacionais descritos na Tabela 4.2, podem ser obtidos em (Hambleton e Swaminathan, 1985).

Apresenta-se, na sequência, o processo de estimação dos parâmetros na TRI para modelos politômicos.

Capítulo 5

Processo de estimação dos parâmetros na TRI - modelos politômicos

5.1 Introdução

Ao iniciar a discussão sobre o processo de estimação dos parâmetros na TRI para os modelos politômicos, vale ressaltar que uma vez definidos os parâmetros do modelo, é necessário estimá-los. Em Teoria de Resposta ao Item, esse processo é conhecido como calibração dos parâmetros. Também é necessário estimar a habilidade dos indivíduos.

Van Der Linden e Hambleton (1997) relatam que o início do desenvolvimento dos Modelos para itens Policotômicos se deu no final dos anos 60, mas só começou a se tornar realmente assunto de pesquisas no início dos anos 80.

A utilização da Teoria da Resposta ao Item - TRI pode ser expandida para situações onde os itens possuem mais de duas opções de resposta, ou seja, pode-se trabalhar com itens cujas respostas sejam politômicas, por exemplo: “Bastante insatisfeito”, “Insatisfeito”, “Levemente insatisfeito”, “Neutro” e “Levemente satisfeito” ou “Feliz”, “Muito feliz” e “Triste”. Esses itens, chamados de politômicos, permitem ao pesquisador oferecer aos indivíduos que serão entrevistados uma variedade maior de respostas.

Segundo a abordagem de Bock e Lieberman (1970), tem-se que há diversos métodos com as finalidades supra citadas. Além disso, várias situações podem ser consideradas: habilidades conhecidas; parâmetros dos itens conhecidos; estimação conjunta dos parâmetros e das habilidades, em uma única ou em duas etapas. Neste trabalho, considerar-se-á a situação em que os parâmetros dos itens e das habilidades são estimados conjuntamente, fato que ocasiona uma grande quantidade de parâmetros a serem estimados. Por isso, a forma mais adequada para realizar a estimação é dividir esse processo em duas etapas, por meio do Método de Máxima Verossimilhança Marginal (vide Seção 4.2.3).

Em Ostini e Nering (2006) encontram-se as razões para usar modelos politômicos da TRI, visto que esses podem ser mais proveitosos que os dicotômicos apresentados no Capítulo 4 desta dissertação. Os autores enfatizam que uma das razões é que os itens politômicos, em virtude do seu maior número de categorias de respostas, fornecem mais informação sobre a faixa de extensão do traço latente que os itens dicotômicos. A escolha entre os modelos de TRI disponíveis para examinar dados politômicos deve estar embasada nas considerações teóricas e empíricas, isto é, no ajuste do modelo aos dados.

Em se tratando da conceituação da pesquisa, é relevante mencionar que nesta, será dado maior ênfase aos modelos politômicos com categorias ordenadas.

5.1.1 Notações e definições

As notações e definições utilizadas neste Capítulo coincidem com as mesmas apresentadas na Seção 4.1 do Capítulo 4.

Apresentam-se, na Seção 5.2, as definições e os critérios para o estabelecimento dos itens e níveis âncoras, a partir do desenvolvimento da escala de medida.

5.2 Especificação de itens âncora

Como mencionado na Seção 3.3, na Teoria de Resposta ao Item, a habilidade de interesse é mensurada em uma escala arbitrária, com variação teórica no intervalo $(-\infty, +\infty)$ em que a escala utilizada para medir os níveis de habilidades é irrelevante. Assim, o fato da escala ser arbitrária não prejudica, de modo algum, a mensuração da habilidade. O importante é a relação entre o nível de habilidade de um indivíduo e sua posição na escala elaborada.

Complementarmente ao exposto, tem-se que a escolha arbitrária da escala das habilidades faz com que suas estimativas tenham pouca interpretação prática, a menos que as relacione com os conteúdos presentes no construto. Assumindo a escala comum de média e desvio padrão iguais a 0 e 1, respectivamente, não se pode afirmar o quanto um indivíduo de habilidade 1,50 sabe a mais que aquele com habilidade -0,60.

Desta forma, verifica-se que os níveis de habilidade não possuem uma interpretação que contextualize o objetivo da TRI. Com essa finalidade, surge a ideia de níveis âncora, que são caracterizados por conjuntos de itens denominados **itens âncora**.

No que concerne à definição dos níveis âncoras, vale citar a abordagem de [Andrade et al. \(2000\)](#), na qual estes consistem em “pontos selecionados pelo analista na escala de habilidade para serem interpretados pedagogicamente”. Conseqüentemente, os itens âncoras, destacam os autores, são aqueles definidos para cada um dos níveis âncoras estabelecidos.

Definição de item âncora ([Andrade et al., 2000](#)) ressaltam que um item é caracterizado como *âncora* se três critérios forem atendidos. Para a notação descrita, considere dois níveis âncoras consecutivos A e B com $A < B$. Diz-se que determinado item é âncora para o nível B se, e somente se, as 3 condições abaixo forem atendidas simultaneamente:

1. $P(U = 1|\theta = B) \geq 0,65$;
2. $P(U = 1|\theta = A) < 0,50$;
3. $P(U = 1|\theta = B) - P(U = 1|\theta = A) \geq 0,30$.

Concomitantemente aos dois critérios apresentados, para o item ser caracterizado como âncora no nível B , conforme o exemplo apresentado, o terceiro critério indica que a diferença de probabilidade os níveis de desempenho A e B deve ser igual ou maior que 30%, de acordo com [Andrade et al. \(2000\)](#).

“[...] Em outras palavras, para um item ser âncora em um determinado nível âncora da escala, ele precisa ser respondido corretamente por uma grande proporção de indivíduos (pelo menos 65%) com este nível de habilidade e por uma proporção menor de indivíduos (no máximo 50%) com o nível de habilidade imediatamente anterior. Além disso, a diferença entre a proporção de indivíduos com esses níveis de habilidade que acertam a esse item deve ser de pelo menos 30%. Assim, para um item ser âncora ele deve ser um item “típico” daquele nível, ou seja, bastante acertado por indivíduos com aquele nível de habilidade e pouco acertado por indivíduos com um nível de habilidade imediatamente inferior” ([Andrade et al., 2000](#)).

A Tabela 5.1 apresenta alguns exemplos de níveis e itens âncoras.

Tabela 5.1: Exemplo de itens âncoras: Escala de satisfação e definição de itens âncoras e quase-âncoras.

Itens	A	B	10	20	30	40	50	60	70	80	90	Classificação
Item 1	0,081	43,46	0,01	0,04	0,13	0,38	0,71	0,91	0,98	0,99	1,00	Item âncora
Item 2	0,094	41,11	0,01	0,03	0,14	0,46	0,81	0,95	0,99	1,00	1,00	Item âncora
Item 3	0,102	39,74	0,01	0,03	0,16	0,51	0,86	0,97	0,99	1,00	1,00	Item âncora
Item 4	0,096	47,93	0,00	0,01	0,05	0,21	0,58	0,88	0,97	0,99	1,00	Quase-âncora
Item 5	0,119	39,08	0,00	0,02	0,14	0,55	0,90	0,99	1,00	1,00	1,00	Quase-âncora
Item 6	0,109	43,15	0,00	0,01	0,08	0,36	0,78	0,96	0,99	1,00	1,00	Item âncora
Item 7	0,111	40,51	0,00	0,02	0,12	0,48	0,86	0,97	1,00	1,00	1,00	Item âncora
Item 8	0,112	55,07	0,00	0,00	0,01	0,05	0,28	0,72	0,94	0,99	1,00	Item âncora
Item 9	0,163	47,51	0,00	0,00	0,01	0,11	0,67	0,97	1,00	1,00	1,00	Item âncora
Item 10	0,118	50,64	0,00	0,00	0,02	0,11	0,47	0,87	0,98	1,00	1,00	Item âncora
Item 11	0,238	51,51	0,00	0,00	0,00	0,01	0,35	0,97	1,00	1,00	1,00	Item âncora
Item 12	0,153	50,09	0,00	0,00	0,01	0,07	0,49	0,93	0,99	1,00	1,00	Item âncora

Fonte: Elaborado pelos autores com base em (Borriá et al., 2009).

Os níveis âncoras apresentados na Tabela 5.1 encontram-se em negrito.

Considerando o exposto, de acordo com o exemplo da Tabela 5.1 e a abordagem de *Bornia et al. (2009)*, os itens 1, 2, 3, 6, 7, 8, 9, 10, 11, 12 são itens âncoras, pois satisfazem às três condições apresentadas. Neste contexto, conforme salientam os autores, os itens 1, 2, 3, 6, 7 e 9 são itens âncoras no nível 50 e os itens 8, 10, 11 e 12 caracterizam o nível âncora 60 da escala de satisfação.

Pode-se observar também, conforme salientam os autores, que os itens 4 e 5 não foram caracterizados como itens âncoras pois não atendem aos três critérios apresentados, mas somente a dois. Assim, os itens 4 e 5 são considerados como itens quase-âncoras. Dessa forma, no presente trabalho, define-se o item quase-âncora como aquele item que atende a somente dois critérios definidos.

Como último comentário, pode-se dizer que é bastante comum fazer uma transformação linear em todos os parâmetros envolvidos antes da construção das escalas. Tal procedimento tem como único objetivo facilitar a construção e utilização da escala, uma vez que procura transformar valores negativos ou decimais em números positivos e inteiros (*Andrade et al., 2000*).

Portanto, após a discussão da escala de desempenho, dos métodos de estimação e dos níveis âncoras, apresenta-se na sequência o conceito da equalização, com o intuito de possibilitar a compreensão referente à comparação entre grupos distintos.

5.3 Equalização

Segundo *Andrade et al. (2000)*, “equalizar significa equiparar, tornar comparável, o que no caso da TRI significa colocar parâmetros de itens vindos de provas distintas ou habilidades de respondentes de diferentes grupos, na mesma métrica, isto é, numa escala comum [...]”

Considerando o exposto pelos autores, pode-se ressaltar que o processo de equalização tem como objetivo a comparação de desempenhos medidos em diferentes grupos, ou por meio de instrumentos que não possuem todos os itens em comum.

Nesta mesma linha, *De Ayala (2009)* destaca que a equalização consiste numa série de procedimentos para o ajuste de estimativas de habilidade obtidas em métricas distintas para uma métrica comum. O autor ressalva que o propósito da equalização é possibilitar a comparação dos indivíduos ou empresas.

Em se tratando dos tipos de equalização, *Andrade et al. (2000)* citam dois tipos: *i*) a equalização *via população*, *ii*) a equalização *via itens comuns*. Isto significa que há duas maneiras de colocar parâmetros, tanto de itens quanto de habilidades, numa mesma métrica.

Na primeira usa-se o fato de que se um único grupo de respondentes é submetido a provas distintas, basta que todos os itens sejam calibrados conjuntamente para termos a garantia de que todos estarão na mesma métrica. Já na equalização via itens comuns, a garantia de que as populações envolvidas terão seus parâmetros em uma única escala será dada pelos itens comuns entre as populações, que servirão de ligação entre elas.

Vargas et al. (2007), com base na abordagem de *Andrade et al. (2000)*, apresenta um resumo (vide Tabela 5.2) dos diferentes tipos de equalização, considerando a situação na qual os itens não foram calibrados.

Tabela 5.2: Diferentes tipos de equalização.

Tipos de equalização	Nº. de populações e modos de verificação	Descrição
Não é necessário nenhum tipo de equalização	Um único grupo verificado por um único conjunto de itens	É o caso trivial em que se aplicam diretamente os modelos matemáticos e os métodos de estimação para uma população.
Equalização via população	Um único grupo verificado por dois conjuntos de itens totalmente distintos	Todos os itens de ambos os conjuntos são calibrados simultaneamente. A garantia de que todos os parâmetros estão na mesma escala é assegurada pelo fato que os respondentes representam uma amostra aleatória da população.
	Um único grupo verificado por dois conjuntos de itens parcialmente distintos	Semelhante ao caso anterior, porém os dois conjuntos de itens têm uma percentagem de itens em comum. Ao final dos processos de estimação todos os resultados obtidos serão comparáveis, não importando qual conjunto de itens cada grupo tenha respondido.
Equalização via itens comuns	Dois grupos verificados por um único conjunto de itens	Todos os itens respondidos por ambas populações são calibrados simultaneamente utilizando-se os modelos matemáticos e métodos de estimação para duas ou mais populações.
	Dois grupos verificados por dois conjuntos de itens totalmente distintos	Esse caso não pode ser resolvido pela TRI. Os conjuntos de itens podem ser calibrados separadamente, mas não se pode comparar os resultados dos itens nem das performances estimadas para os dois grupos.
	Dois grupos verificados por dois conjuntos de itens parcialmente distintos	O uso de itens comuns entre os conjuntos diferentes para verificar populações distintas possibilita que todos os parâmetros estejam na mesma escala ao final dos processos de estimação. Com isso é possível fazer comparações e construir escalas interpretáveis para o contexto verificado. Os itens são calibrados simultaneamente empregando-se os modelos e métodos para duas ou mais populações. Os itens comuns fazem a ligação entre as duas populações envolvidas.

Fonte: Elaborado pelos autores (2016). Adaptada de Vargas *et al.* (2007).

Conforme pode ser observado na Tabela 5.2 e de acordo com a abordagem apresentada por Andrade *et al.* (2000), tanto na equalização via população quanto na equalização via itens comuns - excetuando-se a situação na qual dois grupos realizam duas provas totalmente distintas, que não pode ser solucionada pela Teoria da Resposta ao Item - a equalização é realizada durante o processo de estimação dos parâmetros.

Um outro tipo de equalização apresentada pelos autores refere-se à equalização *a posteriori*, ou seja, depois de finalizado o processo de calibração dos itens.

Nesta mesma linha, tem-se que para a realização da equalização *a posteriori*, destacam os autores, é imprescindível o atendimento ao seguinte pressuposto: a presença de itens comuns nos dois conjuntos.

A equalização *a posteriori*, complementam os autores, engloba um processo no qual os dois conjuntos de itens são calibrados separadamente, resultado em dois conjuntos de estimativas para os itens comuns. Posteriormente, determina-se alguma relação que possibilite situar os parâmetros de um dos conjuntos de itens na escala do outro conjunto. Desta forma,

estando os itens na mesma métrica, é possível realizar a estimação do desempenho de todos os pesquisados, sendo que estes estarão na mesma escala de medida.

Na sequência, apresenta-se a estimação dos parâmetros para o Modelo de Resposta Gradual (MRG) por máxima verossimilhança.

5.4 Modelo de Resposta Gradual: Apresentação e estimação por Máxima Verossimilhança

Na Seção 4.5.2.2 descreveu-se o modelo de Resposta Gradual (MRG). No entanto, o processo de calibração, até então, não havia sido apresentado. Assim, nesta seção, apresenta-se a estimação de seus parâmetros utilizando o método de máxima verossimilhança.

Inicialmente, vale ressaltar que os modelos policotômicos possuem, em geral, 3 suposições que são imprescindíveis para a construção do processo de estimação. Além das suposições (S1) e (S2), descritas na Seção 4.1, considera-se uma terceira:

(S3) a probabilidade de escolha da categoria k do item i pelo indivíduo j , pode ser representada pelo modelo **multivariado de Bernoulli** expresso por:

$$P(\mathbf{U}_{ij} = \mathbf{u}_{ij} | \boldsymbol{\theta}, \zeta_i) \equiv (U_{ij} = \mathbf{u}_{ij}) = \prod_{k=1}^{m_i} P_{ijk}^{\mathbf{u}_{ijk}}, \quad (5.1)$$

onde $P_{ijk}^{\mathbf{u}_{ijk}}$ é uma **função de resposta ao item** relacionada a algum modelo politômico.

Nas subseções seguintes descreve-se o processo de estimação do **Modelo de Resposta Gradual**. Ressalta-se que as notações a serem utilizadas são as mesmas dos Capítulos anteriores.

5.4.1 Habilidades conhecidas

Os desenvolvimentos apresentados nesta subseção, são baseados nos trabalhos de Baker e Kim (2004), Hambleton e Swaminathan (1985, Cap. 7) e Van Der Linden e Hambleton (1997), por exemplo.

Portanto, considerando as suposições (S1), (S2) e (S3), tem-se que a verossimilhança associada ao Modelo de Resposta Gradual de Samejima (1969) é dado por:

$$L(\zeta) = \prod_{j=1}^n \prod_{i=1}^I \prod_{k=0}^{m_i} P_{ijk}^{\mathbf{u}_{ijk}}, \quad (5.2)$$

com P_{ijk} como definido em (3.11).

Desta feita, segue-se que a função de log-verossimilhança é dada por

$$\ln L(\zeta) = \sum_{j=1}^n \sum_{i=1}^I \sum_{k=0}^{m_i} \mathbf{u}_{ijk} \ln P_{ijk}. \quad (5.3)$$

O vetor de escore, o qual fornece os estimadores de máxima verossimilhança (EMV) dos parâmetros de interesse, é dado por

$$\begin{aligned}
 S(\zeta_i) = & \sum_{j=1}^n \frac{\partial}{\partial \zeta_i} \left\{ \mathbf{u}_{ij0} \ln(1 - P_{ij1}^+) + \mathbf{u}_{ij1} \ln(P_{ij1}^+ - P_{ij2}^+) + \dots + \right. \\
 & \mathbf{u}_{ij(h-1)} \ln(P_{ij(h-1)}^+ - P_{ijh}^+) + \mathbf{u}_{ijh} \ln(P_{ijh}^+ - P_{ij(h+1)}^+) + \dots + \\
 & \left. \mathbf{u}_{ijm_i} \ln(P_{ijm_i}^+) \right\}. \tag{5.4}
 \end{aligned}$$

Antes de descrever o vetor de escore para um determinado item e categoria (vide expressão 5.8) necessita-se dos seguintes resultados auxiliares (expressões (5.5), (5.6) e (5.7)):

$$\begin{aligned}
 \frac{\partial P_{ij(k-1)}}{\partial b_{ik}} &= \frac{\partial}{\partial b_{ik}} \left\{ \frac{1}{1 + \exp\{-Da_i(\theta_j - b_{i(k-1)})\}} - \frac{1}{1 + \exp\{-Da_i(\theta_j - b_{ik})\}} \right\} \\
 &= Da_i P_{ijk}^+ Q_{ijk}^+, \tag{5.5}
 \end{aligned}$$

$$\frac{\partial P_{ijk}}{\partial b_{ik}} = -Da_i P_{ijk}^+ Q_{ijk}^+ \quad e \tag{5.6}$$

$$\frac{\partial P_{ijk}}{\partial a_i} = D \left\{ (\theta_j - b_{ik}) W_{ijk}^+ - (\theta_j - b_{i(k+1)}) W_{ij(k+1)}^+ \right\}. \tag{5.7}$$

Logo,

$$S(\zeta_{ik}) = (S(a_i), S(b_{ik}))'. \tag{5.8}$$

Utilizando as expressões (5.5) a (5.7) substituída em (5.4), obtem-se o vetor escore, cujos elementos são dados pelas expressões (5.9) e (5.10):

$$\begin{aligned}
 S(a_i) &= \sum_{j=1}^n \sum_{k=0}^{m_i} \left\{ \frac{\mathbf{u}_{ijk}}{P_{ijk}} \left(\frac{\partial P_{ijk}}{\partial a_i} \right) \right\} \\
 &= D \sum_{j=1}^n \sum_{k=0}^{m_i} \left\{ \frac{\mathbf{u}_{ijk}}{P_{ijk}} \left[(\theta_j - b_{ik}) W_{ijk}^+ - (\theta_j - b_{i(k+1)}) W_{ij(k+1)}^+ \right] \right\} \tag{5.9}
 \end{aligned}$$

e

$$\begin{aligned}
 S(b_{ik}) &= \sum_{j=1}^n \left\{ \frac{\mathbf{u}_{ij(k-1)}}{P_{ij(k-1)}} \left(\frac{\partial P_{ij(k-1)}}{\partial b_{ik}} \right) + \frac{\mathbf{u}_{ijk}}{P_{ijk}} \left(\frac{\partial P_{ijk}}{\partial b_{ik}} \right) \right\} \\
 &= Da_i \sum_{j=1}^n \left\{ W_{ijk}^+ \left[\frac{\mathbf{u}_{ij(k-1)}}{P_{ij(k-1)}} - \frac{\mathbf{u}_{ijk}}{P_{ijk}} \right] \right\}, \tag{5.10}
 \end{aligned}$$

que são componentes de (5.8). Pela não-linearidade do sistema de equações descrito em (5.9) e (5.10), tem-se que utilizar algum método iterativo. Sendo assim, a partir de (5.8), tem-se que as componentes da matriz Hessiana são dadas por

$$\begin{aligned}
H(a_i, a_i) &= D^2 \sum_{j=1}^n \sum_{k=0}^{m_i} \left\{ \frac{\mathbf{u}_{ijk}}{P_{ijk}^2} \left[\left\{ (\theta_j - b_{ik})^2 \left\{ 1 - 2P_{ijk}^+ \right\} W_{ijk}^+ \right. \right. \right. \\
&\quad \left. \left. \left. - (\theta_j - b_{i(k+1)})^2 \left\{ 1 - 2P_{ij(k+1)} \right\} W_{ij(k+1)}^+ \right\} P_{ijk} \right. \right. \\
&\quad \left. \left. - \left[(\theta_j - b_{ik}) W_{ijk}^+ - (\theta_j - b_{i(k+1)}) W_{ij(k+1)}^+ \right]^2 \right] \right\}, \tag{5.11}
\end{aligned}$$

$$\begin{aligned}
H(b_{ik}, b_{ik}) &= D^2 a_i^2 \sum_{j=1}^n W_{ijk}^+ \left\{ \left[\frac{\mathbf{u}_{ij(k-1)}}{P_{ij(k-1)}} - \frac{\mathbf{u}_{ijk}}{P_{ijk}} \right] (1 - 2P_{ijk}^+) \right. \\
&\quad \left. - \left[\frac{\mathbf{u}_{ij(k-1)}}{P_{ij(k-1)}^2} + \frac{\mathbf{u}_{ijk}}{P_{ijk}^2} \right] W_{ijk}^+ \right\}, \tag{5.12}
\end{aligned}$$

$$H(b_{ik}, b_{i(k-1)}) = D^2 a_i^2 \sum_{j=1}^n \left\{ \frac{\mathbf{u}_{ij(k-1)} W_{ijk}^+ W_{ij(k-1)}^+}{P_{ij(k-1)}^2} \right\}, \tag{5.13}$$

$$H(b_{ik}, b_{i(k+1)}) = D^2 a_i^2 \sum_{j=1}^n \left\{ \frac{\mathbf{u}_{ij(k+1)} W_{ijk}^+ W_{ij(k+1)}^+}{P_{ij(k+1)}^2} \right\}, \tag{5.14}$$

$$\begin{aligned}
H(a_i, b_{ik}) &= \sum_{j=1}^n \left\{ D \left[\frac{\mathbf{u}_{ij(k-1)}}{P_{ij(k-1)}} - \frac{\mathbf{u}_{ijk}}{P_{ijk}} \right] W_{ijk}^+ \left\{ 1 + Da_i (\theta_j - b_{ik}) (1 - 2P_{ijk}^+) \right\} \right. \\
&\quad \left. - D^2 a_i W_{ijk}^+ \left[\frac{\mathbf{u}_{ij(k-1)}}{P_{ij(k-1)}^2} \left[(\theta_j - b_{i(k-1)}) W_{ij(k-1)}^+ - (\theta_j - b_{ik}) W_{ijk}^+ \right] \right. \right. \\
&\quad \left. \left. - \frac{\mathbf{u}_{ijk}}{P_{ijk}^2} \left[(\theta_j - b_{ik}) W_{ijk}^+ - (\theta_j - b_{i(k+1)}) W_{ij(k+1)}^+ \right] \right] \right\}, \tag{5.15}
\end{aligned}$$

com $W_{ijk}^+ = P_{ijk}^+ Q_{ijk}^+$ e $H(b_{ik}, b_{i(k+d)}) = 0$ e $H(b_{ik}, b_{i(k-d)}) = 0$, $d \geq 2$. Dessa feita, tem-se que a matriz Hessiana, para um determinado item, pode ser escrita como

$$\mathbf{H}(\zeta)_i = \begin{bmatrix} H(b_{i1}, b_{i1}) & H(b_{i1}, b_{i2}) & 0 & \cdots & \cdots & H(b_{i1}, a_i) \\ H(b_{i2}, b_{i1}) & H(b_{i2}, b_{i2}) & H(b_{i2}, b_{i3}) & 0 & \cdots & H(b_{i2}, a_i) \\ 0 & H(b_{i3}, b_{i2}) & H(b_{i3}, b_{i3}) & H(b_{i3}, b_{i4}) & \cdots & H(b_{i3}, a_i) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & H(b_{im_i}, b_{im_i}) & H(b_{im_i}, a_i) \\ H(a_i, b_{i1}) & H(a_i, b_{i2}) & H(a_i, b_{i3}) & \cdots & H(a_i, b_{im_i}) & H(a_i, a_i) \end{bmatrix}. \tag{5.16}$$

Agora, tomando o valor esperado do simétrico de (5.16), tem-se a Informação de Fischer:

$$I(a_i, a_i) = D^2 \sum_{j=1}^n \sum_{k=0}^{m_i} \left\{ \frac{[(\theta_j - b_{ik})W_{ijk}^+ - (\theta_j - b_{i(k+1)})W_{ij(k+1)}^+]^2}{P_{ijk}} \right\}, \quad (5.17)$$

$$I(b_{ik}, b_{ik}) = D^2 a_i^2 \sum_{j=1}^n \left\{ \left[\frac{1}{P_{ij(k-1)}} + \frac{1}{P_{ijk}} \right] (W_{ijk}^+)^2 \right\}, \quad (5.18)$$

$$I(b_{ik}, b_{i(k-1)}) = D^2 a_i^2 \sum_{j=1}^n \left\{ \frac{W_{ijk}^+ W_{ij(k-1)}^+}{P_{ij(k-1)}} \right\}, \quad (5.19)$$

$$I(b_{ik}, b_{i(k+1)}) = D^2 a_i^2 \sum_{j=1}^n \left\{ \frac{W_{ijk}^+ W_{ij(k+1)}^+}{P_{ijk}} \right\} \quad (5.20)$$

e

$$\begin{aligned} I(a_i, b_{ik}) &= D^2 a_i^2 \sum_{j=1}^n W_{ijk}^+ \left\{ \frac{1}{P_{ij(k-1)}} [(\theta_j - b_{i(k-1)})W_{ij(k-1)}^+ - (\theta_j - b_{ik})W_{ijk}^+] \right. \\ &\quad \left. - \frac{1}{P_{ijk}} [(\theta_j - b_{ik})W_{ijk}^+ - (\theta_j - b_{i(k+1)})W_{ij(k+1)}^+] \right\}. \end{aligned} \quad (5.21)$$

Portanto, os algoritmos iterativos de Newton-Raphson e Escore de Fisher podem ser empregados tais como descritos em (4.23) e (4.30).

5.4.1.1 Habilidades conhecidas e agrupadas

Segundo Baker (1992), devido ao fato de $z_{ijk} = a_{ik}(\theta_j - b_{ik})$ aparecer tanto no numerador quanto no denominador (equação (3.9)), existem problemas de indeterminação adicionais àqueles descritos para os modelos dicotômicos (Bock, 1972).

Nesta mesma linha, Bock (1972) sugere que deve-se impor restrições lineares arbitrárias, tais como $\sum_{k=1}^{m_i} z_{ijk} = 0$ de modo a tornar o problema em estudo estável. Para tanto, considera-se, inicialmente, os seguintes vetores paramétricos: $\zeta_i = (\mathbf{a}'_i, \mathbf{b}'_i)'$, $\zeta_i^* = (\mathbf{a}'_i, \mathbf{d}'_i)'$ e $\Gamma_i = (\boldsymbol{\alpha}'_i, \boldsymbol{\delta}'_i)'$, que denotam os vetores de parâmetros originais, transformados e ir-restritos, respectivamente (Bock, 1972), com $\mathbf{a}_i = (a_{i1}, \dots, a_{im_i})'$, $\mathbf{b}_i = (b_{i1}, \dots, b_{im_i})'$, $\mathbf{d}_i = (d_{i1}, \dots, d_{im_i})'$, $\boldsymbol{\alpha}_i = (\alpha_{i1}, \dots, \alpha_{i(m_i-1)})'$, $\boldsymbol{\delta}_i = (\delta_{i1}, \dots, \delta_{i(m_i-1)})'$, $d_{ik} = -a_{ik}b_{ik}$, $\alpha_{ih} = a_{i1} - a_{i(h+1)}$ e $\delta_{ih} = d_{i1} - d_{i(h+1)}$, $i = 1, \dots, I$, $k = 1, \dots, m_i$ e $h = 1, \dots, m_i - 1$.

No que tange às restrições sugeridas anteriormente, tem-se que $\sum_{k=1}^{m_i} z_{ijk} = 0$ implica que

$$\sum_{k=1}^{m_i} a_{ik} = 0 \quad \text{e} \quad \sum_{k=1}^{m_i} d_{ik} = 0. \quad (5.22)$$

Bock (1972) ressalta que o uso de uma transformação nos parâmetros ζ^* , de modo a contemplar a restrição descrita em (5.22) faz com que o processo iterativo usando o novo vetor de parâmetros Γ , seja computacionalmente mais eficiente do que, por exemplo, a aplicação de técnicas como a dos **Multiplicadores de Lagrange**, usando-se o vetor ζ^* .

Desta feita, na sequência, descreve-se o processo de implementação de tais restrições.

Para tanto, considere-se:

$$\mathbf{Z}_i = \mathbf{B}\boldsymbol{\zeta}^* \mathbf{A}_i, \quad (5.23)$$

em que $\mathbf{B} = (1, \theta)'$, $\mathbf{A}_i = \mathbf{S}_i \mathbf{T}_i$ e

$$\{\mathbf{A}_i\}_{rk} = \Delta_{rk} - \frac{1}{m_i}, \quad (5.24)$$

onde

$$\Delta_{rk} = \begin{cases} 1, & \text{se } r = k \\ 0, & \text{caso contrário} \end{cases}$$

é conhecido como **Delta de Kronecker** (Bock, 1972).

Nesta mesma linha, vale ressaltar que a matriz \mathbf{A}_i é não-inversível ($\text{posto}(\mathbf{A}_i) = m_i - 1$), o que impossibilita a definição dos parâmetros irrestritos através da mesma. A partir de algumas manipulações algébricas, uma reparametrização de (5.23) pode ser considerada fazendo-se

$$\mathbf{Z}_i = \mathbf{B}\boldsymbol{\zeta}^* \mathbf{S}_i \mathbf{T}_i, \quad (5.25)$$

com

$$\mathbf{S}_i = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{bmatrix}$$

e

$$\mathbf{T}_i = \left(\mathbf{S}_i' \mathbf{S}_i \right)^{-1} \mathbf{S}_i' \mathbf{A}_i = \begin{bmatrix} \frac{1}{m_i} & \frac{1}{m_i} - 1 & \cdots & \cdots & \frac{1}{m_i} \\ \frac{1}{m_i} & \cdots & \frac{1}{m_i} - 1 & \cdots & \frac{1}{m_i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{m_i} & \frac{1}{m_i} & \cdots & \cdots & \frac{1}{m_i} - 1 \end{bmatrix}, \quad (5.26)$$

sendo $\boldsymbol{\Gamma}_i = \boldsymbol{\zeta}_i^* \mathbf{S}_i$, os **parâmetros irrestritos**. A matriz \mathbf{S}_i representa um conjunto de contrastes simples (na presente pesquisa, considera-se o caso da categoria 1 contra as demais). Para mais detalhes, ver Azevedo (2003).

Desta feita, de acordo com os autores supra citados e considerando a abordagem das habilidades descrito na Subseção 5.2.1.2.1, tem-se que a log-verossimilhança é dada por

$$\ln(\boldsymbol{\Gamma}_i) = \text{const} + \sum_{l=1}^q \sum_{k=0}^{m_i} \mathbf{u}_{lki} \ln P_{lki}. \quad (5.27)$$

Agora, diferenciando a equação (5.27) com relação aos parâmetros dos itens, seguindo o mesmo desenvolvimento da abordagem de não-agrupamento¹, obtêm-se as componentes do vetor escore dadas por:

$$S(a_i) = D \sum_{l=1}^q \sum_{k=0}^{m_i} \left\{ \frac{r_{ilk}}{P_{ilk}} \left[(\bar{\theta}_l - b_{ik}) W_{ilk}^+ - (\bar{\theta}_l - b_{i(k+1)}) W_{il(k+1)}^+ \right] \right\} \quad (5.28)$$

e

$$S(b_{ik}) = D a_i \sum_{l=1}^q \left\{ W_{ilk}^+ \left[\frac{r_{il(k-1)}}{P_{il(k-1)}} - \frac{r_{ilk}}{P_{ilk}} \right] \right\}, \quad (5.29)$$

em que $P_{ilk} \equiv P_{ijk}$, com θ_j substituído por $\bar{\theta}_l$ e $W_{ilk}^+ = P_{ilk}^+ Q_{ilk}^+$. E, sem perda de generalidade, vale ressaltar que f_{il} e r_{il} correspondem ao número de indivíduos com nível de habilidade em torno de $\bar{\theta}_l$ que respondem ao item i e, dentre esses indivíduos, o número daqueles que respondem corretamente a esse mesmo item, respectivamente.

Desta forma, as componentes da matriz Hessiana são dados pelas seguintes expressões

$$\begin{aligned} H(a_i, a_i) &= D^2 \sum_{l=1}^q \sum_{k=0}^{m_i} \left\{ \frac{r_{ilk}}{P_{ilk}^2} \left[\left[(\bar{\theta}_l - b_{ik})^2 \{1 - 2P_{ilk}^+\} W_{ilk}^+ \right. \right. \right. \\ &\quad \left. \left. - (\bar{\theta}_l - b_{i(k+1)})^2 \{1 - 2P_{il(k+1)}^+\} W_{il(k+1)}^+ \right] P_{ilk} \right. \\ &\quad \left. \left. - \left[(\bar{\theta}_l - b_{ik}) W_{ilk}^+ - (\bar{\theta}_l - b_{i(k+1)}) W_{il(k+1)}^+ \right]^2 \right] \right\}, \end{aligned} \quad (5.30)$$

$$\begin{aligned} H(b_{ik}, b_{ik}) &= D^2 a_i^2 \sum_{l=1}^q W_{ilk}^+ \left\{ \left[\frac{r_{il(k-1)}}{P_{il(k-1)}} - \frac{r_{ilk}}{P_{ilk}} \right] (1 - 2P_{ilk}^+) \right. \\ &\quad \left. - \left[\frac{r_{il(k-1)}}{P_{il(k-1)}^2} + \frac{r_{ilk}}{P_{ilk}^2} \right] W_{ilk}^+ \right\}, \end{aligned} \quad (5.31)$$

$$H(b_{ik}, b_{i(k-1)}) = D^2 a_i^2 \sum_{l=1}^q \left\{ \frac{r_{il(k-1)} W_{ilk}^+ W_{il(k-1)}^+}{P_{il(k-1)}^2} \right\}, \quad (5.32)$$

¹Um procedimento alternativo para a estimação dos parâmetros dos itens, quando as habilidades dos indivíduos são conhecidas, é considerar um agrupamento dessas em categorias. Isso pode acarretar em algumas vantagens computacionais, haja vista que o número de *laços* nos algoritmos pode diminuir sensivelmente, sobretudo quando estamos lidando com um grande número de indivíduos. Dito isto, podemos considerar o agrupamento das habilidades, por exemplo, definindo um conjunto de q intervalos cujos valores médios (ou outra medida central), sejam denotados por $\bar{\theta}_l$, $l = 1, \dots, q$. Podemos considerar, para fins de desenvolvimento, que todos os indivíduos pertencentes à categoria l , têm habilidade igual (ou muito próxima) a $\bar{\theta}_l$, isso, como já foi dito, pode reduzir bastante o esforço computacional, (Azevedo, 2003).

$$H(b_{ik}, b_{i(k+1)}) = D^2 a_i^2 \sum_{l=1}^q \left\{ \frac{r_{il(k+1)} W_{ilk}^+ W_{il(k+1)}^+}{P_{il(k+1)}^2} \right\} \quad (5.33)$$

e

$$\begin{aligned} H(a_i, b_{ik}) &= \sum_{j=1}^n \left\{ D \left[\frac{r_{il(k-1)}}{P_{il(k-1)}} - \frac{r_{ilk}}{P_{ilk}} \right] W_{ilk}^+ \left\{ 1 + D a_i (\bar{\theta}_l - b_{ik}) (1 - 2P_{ilk}^+) \right\} \right. \\ &- D^2 a_i W_{ilk}^+ \left[\frac{r_{il(k-1)}}{P_{il(k-1)}^2} \left[(\bar{\theta}_l - b_{i(k-1)}) W_{il(k-1)}^+ - (\bar{\theta}_l - b_{ik}) W_{ilk}^+ \right] \right. \\ &\left. \left. - \frac{r_{ilk}}{P_{ilk}^2} \left[(\bar{\theta}_l - b_{ik}) W_{ilk}^+ - (\bar{\theta}_l - b_{i(k+1)}) W_{il(k+1)}^+ \right] \right] \right\}. \end{aligned} \quad (5.34)$$

Logo, as componentes da informação de Fisher são dadas por

$$I(a_i, a_i) = D^2 \sum_{l=1}^q \sum_{k=0}^{m_i} \left\{ \frac{f_{il}}{P_{ilk}} \left[(\bar{\theta}_l - b_{ik}) W_{ilk}^+ - (\bar{\theta}_l - b_{i(k+1)}) W_{il(k+1)}^+ \right]^2 \right\}, \quad (5.35)$$

$$I(a_i, a_i) = D^2 a_i^2 \sum_{l=1}^q \left\{ f_{il} \left[\frac{1}{P_{il(h-1)}} + \frac{1}{P_{ilh}} \right] (W_{ilh}^+)^2 \right\}, \quad (5.36)$$

$$I(b_{ih}, b_{i(h-1)}) = D^2 a_i^2 \sum_{l=1}^q \left\{ \frac{f_{il} W_{ilh}^+ W_{il(h-1)}^+}{P_{il(h-1)}} \right\}, \quad (5.37)$$

$$I(b_{ih}, b_{i(h+1)}) = D^2 a_i^2 \sum_{l=1}^q \left\{ \frac{f_{il} W_{ilh}^+ W_{il(h+1)}^+}{P_{ilh}} \right\} \quad (5.38)$$

e

$$\begin{aligned} I(a_i, b_{ih}) &= \sum_{l=1}^q f_{il} W_{ilh}^+ \left\{ \frac{1}{P_{il(h-1)}} \left[(\bar{\theta}_l - b_{ik}) W_{ilk}^+ - (\bar{\theta}_l - b_{i(k+1)}) W_{il(k+1)}^+ \right] \right. \\ &\left. - \frac{1}{P_{ilh}} \left[(\bar{\theta}_l - b_{ih}) W_{ilh}^+ - (\bar{\theta}_l - b_{i(h+1)}) W_{il(h+1)}^+ \right] \right\}. \end{aligned} \quad (5.39)$$

Na sequência, apresenta-se a estimação dos parâmetros dos itens quando a habilidade é desconhecida.

5.4.2 Habilidades desconhecidas

Na Subseção 4.2.3.1 deste estudo de dissertação, tem-se a probabilidade marginal é dada por Bock e Lieberman (1970):

$$\begin{aligned} P(\mathbf{u}_{.j}|\boldsymbol{\zeta}, \boldsymbol{\eta}) &= \int_{\mathbb{R}} P(\mathbf{u}_{.j}|\boldsymbol{\theta}, \boldsymbol{\zeta}, \boldsymbol{\eta})g(\boldsymbol{\theta}|\boldsymbol{\eta})d\boldsymbol{\theta} \\ &= \int_{\mathbb{R}} P(\mathbf{u}_{.j}|\boldsymbol{\theta}, \boldsymbol{\zeta})g(\boldsymbol{\theta}|\boldsymbol{\eta})d\boldsymbol{\theta}, \end{aligned} \quad (5.40)$$

e a verossimilhança e a log-verossimilhança dadas por

$$L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \prod_{j=1}^n P(\mathbf{u}_{.j}|\boldsymbol{\zeta}, \boldsymbol{\eta}) \quad (5.41)$$

e

$$\ln L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \sum_{j=1}^n \ln P(\mathbf{u}_{.j}|\boldsymbol{\zeta}, \boldsymbol{\eta}). \quad (5.42)$$

As componentes do vetor escore, ver a Subseção 4.2.3.1 e as equações (5.9) e (5.10), são dadas por

$$S(a_i) = D \sum_{l=1}^q \sum_{k=0}^{m_i} \left\{ \frac{\bar{r}_{ilk}}{P_{ilk}} \left[(\bar{\theta}_l - b_{ik})W_{ilk}^+ - (\bar{\theta}_l - b_{i(k+1)})W_{il(k+1)}^+ \right] \right\} \quad (5.43)$$

e

$$S(b_{ih}) = Da_i \sum_{l=1}^q \left\{ W_{ilh}^+ \left[\frac{\bar{r}_{il(h-1)}}{P_{il(h-1)}} - \frac{\bar{r}_{ilh}}{P_{ilh}} \right] \right\}, \quad (5.44)$$

com $\bar{r}_{ilk} = \sum_{j=1}^n \mathbf{u}_{ijk} g_j^*(\bar{\theta}_l)$ e $g_j^*(\bar{\theta}_l) = \frac{P(\mathbf{u}_{.j}|\boldsymbol{\zeta}, \bar{\boldsymbol{\theta}}_l)g(\bar{\boldsymbol{\theta}}_l|\boldsymbol{\eta})}{\sum_{i=1}^q P(\mathbf{u}_{.j}|\boldsymbol{\zeta}, \bar{\boldsymbol{\theta}}_l)g(\bar{\boldsymbol{\theta}}_l|\boldsymbol{\eta})}$.

Devido as expressões (5.43) e (5.44) não apresentarem soluções explícitas, Samejima (1997) sugere uma adaptação do algoritmo EM para a resolução destas. Na sequência, apresenta-se esta adaptação.

5.4.2.1 Adaptação do algoritmo EM

A seguir, apresenta-se uma extensão, proposta por Samejima (1997), do uso do algoritmo EM nos modelos dicotômicos, a saber:

Passo E

Usar os pontos de quadratura $\bar{\boldsymbol{\theta}}_l$, os passos $\mathbf{A}_l^{(t)}$, $l = 1, \dots, q$ e estimativas provisórias dos parâmetros dos itens, $\hat{\boldsymbol{\zeta}}_i^{(t)}$, $i = 1, \dots, I$, para $g_j^*(\bar{\boldsymbol{\theta}}_l)^{(t)}$ e, posteriormente, $\bar{\mathbf{r}}_{ilk}^{(t)}$ e $\bar{\mathbf{f}}_{il}^{(t)}$, $i = 1, \dots, I$, $l = 1, \dots, q$ e $k = 0, 1, \dots, m_i$.

Passo M

com $\bar{\mathbf{r}}^{(t)}$ e $\bar{\mathbf{f}}^{(t)}$ obtidos no passo **Passo E**, resolver as equações de estimação para $\boldsymbol{\zeta}_i$, $i = 1, \dots, I$, usando o algoritmo de Newton-Raphason ou Escore de Fisher.

Azevedo (2003) em sua dissertação de mestrado, ressalta que como as quantidades \mathbf{r} e \mathbf{f} são substituídas pelos seus respectivos valores esperados $\bar{\mathbf{r}}$ e $\bar{\mathbf{f}}$, os algoritmos iterativos podem ser implementados usando-se as expressões (5.30) e (5.38), com as convenientes substituições $(\bar{\mathbf{r}}, \bar{\mathbf{f}})'$.

Samejima (1997) e Bock e Zimowski (1997) ressaltam que este procedimento gera estimativas consistentes dos parâmetros dos itens. Os autores ainda enfatizam que a superfície de verossimilhança é convexa em quase toda parte. Logo, o vetor de pontos pontos estacionários será único e, portanto, o máximo global da função de verossimilhança.

Nesta mesma linha, Bock e Zimowski (1997) enfatizam que o processo de estimação convergirá a menos de padrões de respostas aberrantes. A informação de Fisher fornecerá (desde que esta seja positiva definida) os erros-padrão assintóticos das estimativas dos parâmetros dos itens.

5.4.3 Estimação das habilidades

Levando em consideração o mesmo argumento dos modelos anteriores bem como as notações matemáticas, tem-se que a verossimilhança para a habilidade de um determinado indivíduo é dada por

$$L(\boldsymbol{\theta}_j) = \prod_{i=1}^I \prod_{k=0}^{m_i} P_{ijk}^{\mathbf{u}_{ijk}}, \quad (5.45)$$

e a log-verossimilhança dada por

$$\ln L(\boldsymbol{\theta}_j) = \sum_{i=1}^I \sum_{k=0}^{m_i} \mathbf{u}_{ijk} \ln P_{ijk}, \quad (5.46)$$

De acordo com o explicitado na Seção 5.4.2, o processo descrito se aplica a qualquer modelo politômico, sendo necessário, portanto, algumas poucas adaptações devido a mudanças do modelo. Estas são tão somente, usar as expressões definidas anteriormente.

Nesta mesma linha, a partir de algumas manipulações algébricas, obtém-se o seguinte vetor escore

$$\begin{aligned} S(\boldsymbol{\theta}_j) &= \sum_{i=1}^I \sum_{k=0}^{m_i} \left\{ \frac{\mathbf{u}_{ijk}}{P_{ijk}} \left(\frac{\partial P_{ijk}}{\partial \boldsymbol{\theta}_j} \right) \right\} \\ &= D \sum_{i=1}^I \sum_{k=0}^{m_i} \left\{ \frac{a_i \mathbf{u}_{ijk} (W_{ijk}^+ - W_{ij(k+1)})}{P_{ijk}} \right\}. \end{aligned} \quad (5.47)$$

Desta feita, analogamente ao apresentado na Subseção 4.2.1.2.1, ressalta-se que a equação (5.47) não apresenta solução explícita para $\boldsymbol{\theta}_j$ e, por isso, precisa-se de algum método iterativo para obter as estimativas desejadas.

Para tanto, tem-se que a matriz Hessiana é dada por

$$H(\boldsymbol{\theta}_j) = D^2 \sum_{i=1}^I \sum_{k=0}^{m_i} \left\{ a_i^2 \mathbf{u}_{ijk} \left[\frac{(1 - 2P_{ijk}^+)W_{ijk}^+ - (1 - 2P_{ij(k+1)}^+)W_{ij(k+1)}^+}{P_{ijk}} - \frac{(W_{ijk}^+ - W_{ij(k+1)}^+)^2}{P_{ijk}^2} \right] \right\}, \quad (5.48)$$

e a Informação de Fisher dada por

$$I(\boldsymbol{\theta}_j) = D^2 a_i^2 \sum_{i=1}^I \sum_{k=0}^{m_i} \left\{ - \frac{(W_{ijk}^+ - W_{ij(k+1)}^+)^2}{P_{ijk}} \right\}. \quad (5.49)$$

Assim, de maneira análoga, os métodos iterativos, descritos acima, podem ser empregados como descritos nas equações (4.2.1.2.1.1) (4.2.1.2.1.2) desta pesquisa.

Por fim, ressalta-se que neste trabalho de dissertação não será apresentada a estimação Bayesiana para o *Modelo de Resposta Gradual* de Samejima (1969). Esta será uma das contribuições futuras da preponente pesquisa. No entanto, para mais detalhes sobre a abordagem Baysiana ver os seguintes autores: Samejima (1969), Samejima (1997), Bock e Lieberman (1970), Bock e Zimowski (1997) e Azevedo (2003).

Os parâmetros utilizados neste trabalho, assim como a escala de habilidade, foram estimados com auxílio dos *softwares* Multilog e R, versão 3.3.1.

Apresenta-se, na sequência, a metodologia que caracteriza o presente estudo.

Capítulo 6

Pré-processamento dos dados

Neste capítulo apresenta-se alguns procedimentos utilizados no pré-processamento dos dados, de forma a eliminar inconsistências, tratar informações faltantes e padronizar os dados, de modo a criar-se escalas que não sejam influenciadas pelas sub-áreas.

Posteriormente, apresenta-se as formas nas quais os dados foram categorizados. Este processo é necessário para permitir a utilização de modelos de TRI politômicos. Ao final apresenta-se um esboço dos procedimentos a serem utilizados na ordenação dos periódicos após a criação de uma escala pela TRI.

Como foi mencionado ao final do Capítulo 2, trabalhou-se, prioritamente, com uma base de dados do ano de 2012, por esta conter a informação sobre a área ou grupo (Matemática, Matemática Aplicada, Estatística ou Outras) de cada periódico.

A informação sobre a área prioritária à qual o periódico pertence é imprescindível para que se possa padronizar os dados em cada um destes grupos, na tentativa de eliminar efeitos sistemáticos nos indicadores, causados pela área de concentração dos periódicos. É importante frisar que objetiva-se construir uma escala ou métrica única, para que se possa ordenar os periódicos e, posteriormente, classificá-los.

6.1 Os dados do ano-base de 2012

Na fase inicial da presente pesquisa considerou-se os dados fornecidos pela CAPES (<http://qualis.capes.gov.br/webqualis>, vide Capítulo 2), no formato de um arquivo PDF, uma vez que este documento é o único que proporciona a informação sobre a área ou grupo a qual cada periódico pertence. Este arquivo foi convertido para o formato XLSM, utilizando-se a ferramenta disponível no site <http://bit.ly/2lybrha>. Feito isso, alguns problemas de formatação foram corrigidos, utilizando-se scripts de programação de Macros em *Visual Basic for Applications* (VBA), gerando assim arquivos texto, com posterior refinamento na linguagem de processamento de textos **AWK** (<http://bit.ly/1HTaJ66>).

As linhas de comandos desenvolvidas em VBA, na presente pesquisa, encontram-se no **Apêndice H**.

Extraídas as informações para uma base de dados, a Tabela 6.1 apresenta os resultados apurados na base proveniente do arquivo PDF (dados originais da CAPES):

Tabela 6.1: *Resultados apurados a partir do arquivo PDF de 2012.*

Informações	Totais
Registros no arquivo	1.197
Registros sem ISSN	5
Registros considerados <i>outliers</i>	3
Registros sem informações nos indicadores (NA)	307
Registros para análise	885
Indicadores TODOS preenchidos	318
Registros do indicador (AIS) preenchidos	679
Registros do indicador (JCR) preenchidos	782
Registros do indicador (HL) preenchidos	726
Registros do indicador (SciNet) preenchidos	460

Fonte: Elaborado pelos autores (2017).

Com base na Tabela 6.1 faz-se as seguintes observações:

- i)* os registros considerados *outliers* foram retirados nas análises;
- ii)* por não se obter informações, seja na literatura ou na própria plataforma da CAPES, sobre o indicador “SciNet”, este não foi incluído nos estudos desta dissertação;
- iii)* apenas 664 registros possuem informações para os indicadores Article Influence Score (AIS), fator de impacto (FI) e meia vida (HL) conjuntamente, sem levar em conta o SciNet.

Inicialmente, no processo de imputação, o SciNet foi considerado, sendo este retirado, apenas, nas análises que envolvem a TRI. Assim, verificou-se o quantitativo de 307 *missings* a serem imputados nos 4 indicadores.

Na Tabela 6.2 apresenta-se algumas estatísticas descritivas sobre os registros que não tinham informações (*missings*) no PDF, conforme visto na Tabela 6.1

Tabela 6.2: *Quantitativo dos registros no arquivo pdf com informações ausentes nos indicadores segundo suas classificações por estrato.*

Classificação 2012	Frequência
A1	3
A2	9
B1	7
B2	15
B3	28
B4	72
B5	152
C	21
Total Geral	307

Fonte: Elaborado pelos autores (2017).

Considerando os mesmos dados descritos na Tabela 6.2, apresenta-se na Tabela 6.3 esses dados levando-se em consideração a classificação por área de interesse.

Observa-se na Tabela 6.2 que 73% dos dados faltantes referem-se aos periódicos que foram classificados como B4 e B5 pela CAPES.

Tabela 6.3: *Quantitativo dos registros com informações ausentes da Tabela 6.2, segundo suas classificações nos estratos e grupos.*

Grupo	Classificação por estrato							
	A1	A2	B1	B2	B3	B4	B5	C
Engenharia	0	0	0	0	1	0	0	0
Estatística	0	1	0	3	4	8	15	1
Matemática	2	4	3	3	2	25	52	4
Matemática Aplicada	1	3	4	2	3	8	6	0
Outra área	0	1	0	7	18	31	79	16
Total	3	9	7	15	28	72	152	21

Fonte: Elaborado pelos autores (2017).

No que tange aos registros **com** informação, tem-se na Seção 6.1.1 algumas estatísticas descritivas por área.

Na Seção 6.2 será discutido o processo de imputações dos dados utilizados nesta dissertação. Este processo é necessário para que seja possível trabalhar com toda a base de dados de 2012, envolvendo os 1189 periódicos. A seguir apresentam-se uma breve análise descritiva dos dados realizada nos dados sem imputação.

6.1.1 Análise descritiva para a base de dados *antes* do processo de imputação

Na Figura 6.1 apresenta-se algumas representações gráficas para a base de dados **antes** do processo de imputação. Analisando os histogramas, verifica-se que os indicadores AIS e Fator de Impacto (FI) apresentam comportamento assimétrico à direita, os boxplots confirmam tais evidências. Já o fator de Meia Vida (HL) descreve um comportamento assimétrico à esquerda e uma grande concentração de informações acima da mediana como observa-se no boxplot referente a este indicador. Uma observação importante sobre a Meia Vida (HL) é que este tipo de dado é censurado em 10 anos, de modo que os periódicos cujo $HL > 10$ não trazem informação exata sobre esse indicador bibliométrico. Nesses casos, adotou-se o valor $HL = 11$ anos.

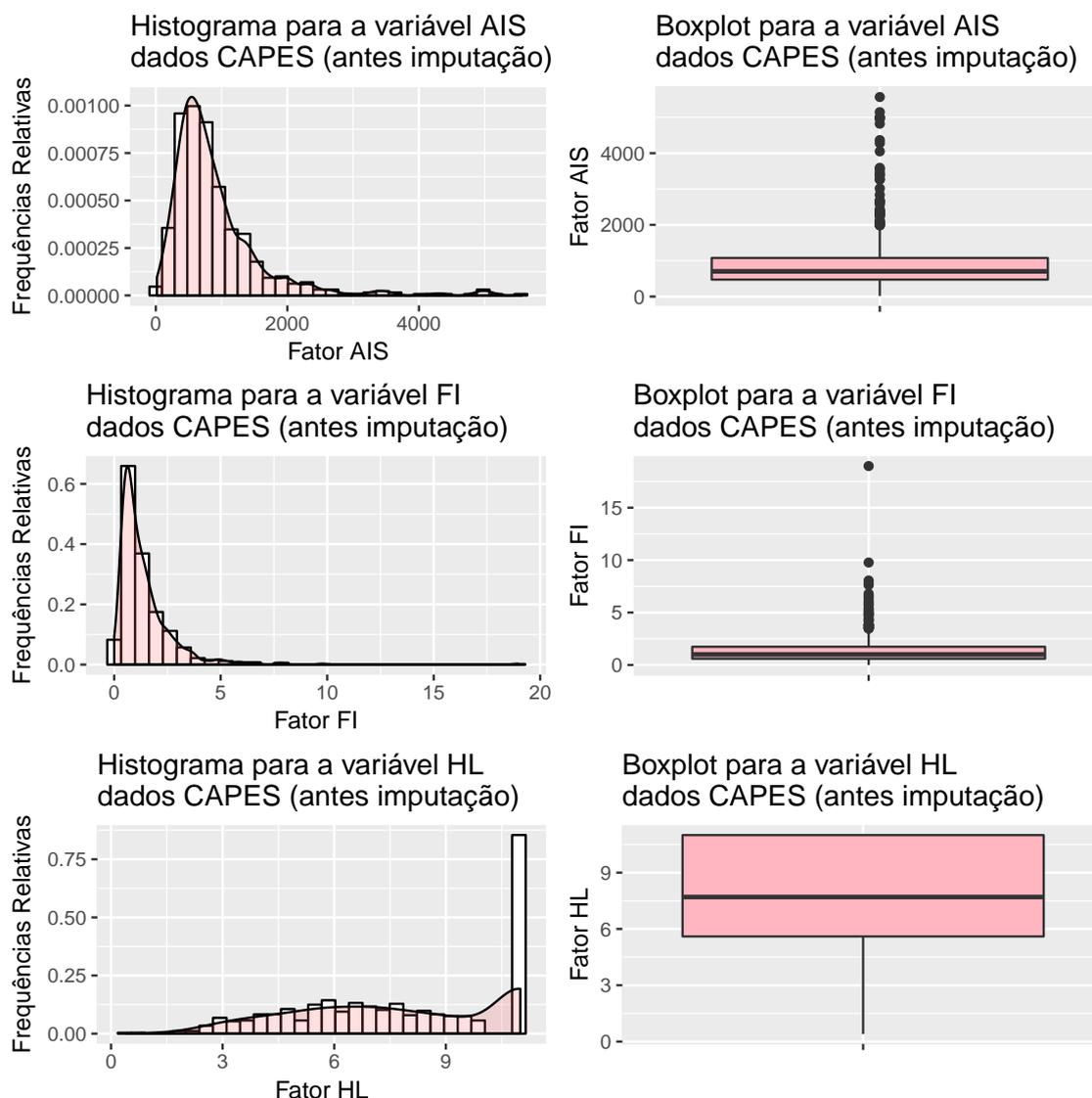


Figura 6.1: Representações gráficas para os indicadores bibliométricos: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) antes do processo de imputação.

No que tange às áreas de interesse desta pesquisa, isto é, as áreas nos quais os periódicos estão classificados, na Tabela 6.4 apresenta-se antes do processo de imputação algumas informações sobre os registros, por área de interesse, que formam a base de dados antes do processo composta por $N = 882$ periódicos.

Tabela 6.4: Representação tabular dos periódicos científicos por área de interesse antes do processo de imputação.

Área de interesse	Frequências
Matemática	314 (35.60%)
Matemática Aplicada	136 (15.41%)
Estatística	102 (11.56%)
Outra área	330 (37.41%)

Fonte: Elaborado pelos autores (2017).

Baseado na Tabela 6.4, verifica-se que o número de periódicos classificados em Outra área representa uma parte substancial dos dados, contendo 37.41% do total dos registros, ao passo

que a área de Estatística é a que comporta o menor percentual de registros, representando pouco mais de 10% do total.

A título de informação, na Tabela 6.5 apresenta-se algumas análises descritivas dos indicadores bibliométricos **antes** da imputação (dados originais).

Tabela 6.5: Análise descritiva dos indicadores bibliométricos antes da imputação (dados originais).

Indicador	Core	Medidas resumo						
		Min.	1st. Qu.	Mediana	Média	3rd. Qu.	Max.	NA's
AIS	Est.	229	625	961	1205	1535	4822	25
	Mat.	153	476	718	941	1070	5568	111
	Mat. Apl.	129	513	706	890.3	1001	4998	15
	Out. Área	17	436	660	814.1	1014	4995	55
FI	Est.	0.0560	0.6098	0.8400	1.0520	1.4400	3.5000	10
	Mat.	0.0000	0.4580	0.6300	0.8414	0.9820	9.7710	76
	Mat. Apl.	0.2920	0.7340	1.0660	1.2390	1.5402	6.5800	11
	Out. Área	0.0210	0.9058	1.5840	1.9580	2.5200	18.9700	6
HL	Est.	1.853	6.700	8.900	8.595	11.000	11.000	19
	Mat.	2.300	6.200	9.700	8.615	11.000	11.000	88
	Mat. Apl.	1.800	5.600	7.600	7.722	11.000	11.000	11
	Out. Área	0.402	5.125	6.700	6.917	8.800	11.000	40

Fonte: Elaborado pelos autores (2017).

Observa-se na Tabela 6.5 que os valores médios/mediano dos indicadores em cada área são bem distintos, em especial o AIS e o FI. Tal constatação justifica a necessidade de padronização, dentro de cada área ou grupo, previamente às análises via TRI. O processo de padronização dos dados é descrito na Seção 6.3. A seguir apresenta-se o processo de imputação realizado na base de dados. Denota-se este procedimento por Etapa I desta pesquisa.

6.2 Processo de imputação na base de dados

Nesta seção discute-se o processo de imputação dos 307 dados faltantes informados na Tabela 6.2.

As razões pelas quais a imputação dos dados faz-se necessária são devidas às seguintes duas questões:

- (1) Objetiva-se ordenar e classificar todos os 1189 periódicos incluídos na base de dados de 2012. Caso os 307 registros com dados faltantes (vide Tabela 6.2) fossem sumariamente removidos, os artigos que compõem a “base da pirâmide” dos periódicos (periódicos nos estratos inferiores B4 ou B5) - seriam quase todos removidos. Desta forma, a nova base da pirâmide seria composta por periódicos classificados pela CAPES como B3 ou mesmo B2, o que, claramente, é incorreto pois a classificação resultante dos periódicos ficaria viesada.
- (2) No caso do tratamento dos dados utilizando modelos politômicos de TRI, pode-se pensar em discretizar os dados já padronizados e tratar os dados faltantes como sendo *itens não apresentados*. No entanto, verificou-se que a adoção de tal procedimento conduz a estimativas de escores de TRI incorretos, sendo calculados como zero caso onde os “itens” de um dado “indivíduo” fossem faltantes. Desta forma, observou-se que para todos os periódicos cujas medidas de todos os indicadores bibliométricos eram faltantes, os respectivos “escores de proficiência” foram estimados como sendo zero,

indicando que o escore do determinado jornal foi “alocado” no centro da escala, o que acarreta em classificações nos estratos incorretos.

Desta maneira, considerando o exposto nos itens 1) e 2) acima, não restou outra alternativa a não ser a de se trabalhar com dados imputados.

O mais simples dos métodos e que também está disponível na maioria dos pacotes estatísticos, é a substituição de valores faltantes por alguma medida resumo, no qual os “missings” são substituídos pela média (ou mediana) dos dados válidos (não “missings”) da variável em estudo. Todavia, pelo fato de o mesmo valor estar sendo atribuído a cada um dos faltantes, este método diminui artificialmente a variância da variável em questão, além de influenciar na relação com as demais variáveis. Mais informações sobre imputação de dados podem ser encontradas no **Apêndice L**.

No processo de imputação de dados, considerou-se as $5 \times 8 = 40$ possibilidades visualizadas na Tabela 6.3 de classificação cruzada, onde se tem 5 áreas de interesse (Grupos) e 8 estratos de classificação. Este processo realizou-se da seguinte maneira: observado o cruzamento citado entre os registros com informação preenchida em *pelo menos um* dos 4 indicadores considerados, totalizando 882 ($885-3$ (*outliers*)=882) registros, o mesmo cruzamento foi observado entre os 307 registros sem informações *em todos os 4 indicadores*. Desta forma, as Tabelas 6.3 e 6.6, as quais serviram como referência para o processo de imputação. Observa-se que, em relação à Tabela 6.3, o máximo de $4 \times 8 = 32$ possibilidades foi considerado, em função da realocação da área “Engenharia” para *Outra área*.

Tabela 6.6: *Quantitativo de registros (entre os 882) utilizados como base para o processo de imputação segundo suas classificações por Área.*

Grupo	Classificação por estrato							
	A1	A2	B1	B2	B3	B4	B5	C
Estatística	16	27	26	17	5	6	5	0
Matemática	42	52	63	61	38	28	30	0
Matemática Aplicada	17	38	41	16	15	5	4	0
Outra área	19	49	111	54	41	30	25	1
Total	94	166	241	148	99	69	64	1

Fonte: Elaborado pelos autores (2017).

Observa-se na Tabela 6.3 que a grande maioria dos dados a serem imputados (73%) são devidos aos periódicos classificados como sendo B4 ou B5 pela CAPES. Tal fato evidencia a grande dificuldade que este tipo de periódico representa aos representantes de área da CAPES.

O processo de imputação foi avaliado tanto para imputações feitas com base na média, quanto com base na mediana e, embora muito simples, não foi possível imputar-se valor para todos os indicadores em todos os registros faltantes por uma questão estrutural, tal como se observa no estrato de classificação *C* da Tabela 6.6, onde há periódicos disponíveis apenas para “Outra Área”. Assim, para cada periódico informado em dada célula da Tabela 6.3, em cada um dos indicadores, foi atribuído a média ou mediana calculada sobre o total de periódicos informados pela respectiva célula na Tabela 6.6.

Portanto, para fins de construção da base de dados com vistas ao desenvolvimento de análises via TRI imputou-se cada observação faltante pela média/mediana das informações disponíveis. Não se pode deixar de mencionar que a imputação é uma poderosa ferramenta na construção de bancos de dados. No entanto, é importante que se tenha cautela, assim como em qualquer procedimento estatístico. É claro que se o modelo de imputação não consegue capturar o mecanismo de não resposta, as análises feitas com dados imputados podem ficar

comprometidas. Estudos para avaliar a escolha do processo de imputação são discutidos na Seção 7.2.

Na Tabela 6.7 apresenta-se informações sobre os registros, por área de interesse, que formam a base de dados de $N = 1189$ periódicos, após o processo de imputação (pela média ou mediana).

Tabela 6.7: *Representação tabular dos periódicos científicos, por área de interesse, após o processo de imputação.*

Área de interesse	Frequências
Matemática	409 (34.39%)
Matemática Aplicada	163 (13.70%)
Estatística	134 (11.26%)
Outra área	483 (40.62%)

Fonte: Elaborado pelos autores (2017).

A Tabela 6.7 mostra o número de periódicos classificados como Outra área representa 40.62% do total de registros. Analogamente à Tabela 6.4, a área de Estatística representa pouco mais de 10% do total de periódicos de base de dados com imputação.

Nas Tabelas 6.8 e 6.9 apresentam-se algumas estatísticas descritivas para os dados completos ($N = 1189$), com imputação pela média e pela mediana, respectivamente. Uma análise tentativa para estabelecer a melhor escolha, se a imputação pela média ou pela mediana, é apresentada na Seção 7.2.

Tabela 6.8: *Análise descritiva dos indicadores bibliométricos após a imputação via média.*

Indicador	Core	Medidas resumo						
		Mín.	1st. Qu.	Mediana	Média	3rd. Qu.	Max.	NA's
AIS	Est.	229.0	521.1	801.6	1059.0	1361.0	4822.0	35
	Mat.	153.0	305.0	506.0	745.7	844.0	5568.0	86
	Mat. Apl.	129.0	439.0	670.0	811.3	946.9	4998.0	10
	Out. Área	17.0	264.3	413.5	612.8	736.0	4995.0	17
FI	Est.	0.0560	0.4188	0.6780	0.8358	1.0360	3.5000	1
	Mat.	0.0000	0.3410	0.4600	0.6025	0.7370	9.7710	4
	Mat. Apl.	0.2920	0.7130	1.0010	1.1800	1.4920	6.5800	10
	Out. Área	0.0210	0.4470	1.0140	1.5260	2.0140	18.9700	-
HL	Est.	1.853	6.200	8.200	8.055	11.000	11.000	21
	Mat.	2.300	6.800	7.700	8.185	11.000	11.000	86
	Mat. Apl.	1.800	5.500	7.400	7.384	11.000	11.000	11
	Out. Área	0.402	4.087	6.032	6.254	7.475	11.000	17

Fonte: Elaborado pelos autores (2017).

Tabela 6.9: *Análise descritiva dos indicadores bibliométricos após a imputação via mediana.*

Indicador	Core	Medidas resumo						
		Min.	1st. Qu.	Mediana	Média	3rd. Qu.	Max.	NA's
AIS	Est.	229.0	476.0	767.0	1047.0	1361.0	4822.0	35
	Mat.	153.0	290.0	506.0	739.5	818.5	5568.0	86
	Mat. Apl.	129.0	439.0	622.0	804.7	933.0	4998.0	10
	Out. Área	17.0	201.5	413.5	587.0	729.0	4995.0	17
FI	Est.	0.0560	0.3460	0.6780	0.8223	1.0150	3.5000	1
	Mat.	0.0000	0.2990	0.4600	0.5812	0.7235	9.7710	4
	Mat. Apl.	0.2920	0.6880	0.8960	1.1540	1.4550	6.5800	10
	Out. Área	0.0210	0.4470	1.0000	1.5010	2.0050	18.9700	-
HL	Est.	1.853	6.200	8.200	8.083	11.000	11.000	21
	Mat.	2.300	5.800	7.700	8.070	11.000	11.000	86
	Mat. Apl.	1.800	5.350	7.200	7.387	11.000	11.000	10
	Out. Área	0.402	3.650	5.900	6.124	7.475	11.000	17

Fonte: Elaborado pelos autores (2017).

Apesar dos procedimentos e esforços envidados para se obter uma base livre de informações faltantes, 15 registros não podem ser imputados.

Portanto, a base imputada apresenta alguns NaN's (missings) porque o processo de imputação tinha como referência a combinação das categorias *versus* áreas e por este motivo, haviam células deste cruzamento que não possuía informações, de maneira a permitir imputação para todos os elementos.

A Tabela 6.10 lista-se os 15 periódicos que apresentam NaN's na base de dados após o processo de imputação via média, $N = 1189$ periódicos.

Tabela 6.10: *Representação tabular dos NaN's para base de dados após o processo de imputação via média, $N = 1189$ periódicos.*

ISSN	Periódico	CAPES/2012	Core
0930-0325	Lecture Notes in Statistics	C	Estatística
0075-8469	Lecture Notes in Pure and App. Math.	C	Matemática
0076-0552	London Mathematical Soc. Lect.	C	Matemática
1270-900X	ESAIM. Proceedings	C	Matemática
1807-0299	Alea : Estudos Neolatinos	C	Matemática
0972-0960	Far East Journal of Applied Math.	B5	Mat. Apl.
1311-8080	Inter. Journal of Pure and App. Math.	B5	Mat. Apl.
1607-2510	Applied Mathematics E-Notes	B5	Mat. Apl.
1787-5021	Annales Mathematicae et Informaticae	B5	Mat. Apl.
0973-1377	Inter. Journal of App. Math. & Stat.	B5	Mat. Apl.
0973-5607	Current Development in Th. and App.	B5	Mat. Apl.
0973-7545	Inter. Journal of Applied Math. and Stat.	B5	Mat. Apl.
1083-0464	Southwest journal of pure and App. Math.	B5	Mat. Apl.
1679-365X	Biomatemática (UNICAMP)	B5	Mat. Apl.
1687-2177	Journal of Applied Math. and Stoch.	B5	Mat. Apl.

Fonte: Elaborado pelos autores (2017).

Mesmo com o processo de imputação realizado, de acordo com a Tabela 6.10, tem-se que 15 registros continuaram com NaN's (missings) em seus indicadores. Esses missings foram tratados como código "1" (indica a categoria ao qual o item não tem classificação).

O processo de imputação verificou para cada indicador, e segundo a Tabela 6.3, a qual observam-se estatísticas descritivas sobre os registros com informações ausentes segundo suas classificações nos estratos e grupos, o quantitativo informado pela Tabela 6.6 (882 periódicos), valendo-se da comparação entre estes para o cálculo dos valores a serem imputados.

Como exemplo, considere na Tabela 6.6 os 4 periódicos da Matemática Aplicada com classificação B5 que possuem informação em algum dos indicadores analisados. A informação contida em cada um dos indicadores destes 4 periódicos foi utilizada no cálculo da função de imputação (média/mediana) a ser alocado aos 6 periódicos citados na Tabela 6.3, de forma que o processo efetuou o cálculo do valor a ser imputado para cada indicador, substituindo o cálculo nos respectivos indicadores dos periódicos sem informação. Porém, em função da regra de negócio para seleção dos periódicos utilizada, houveram situações em que mesmo havendo quantitativo suficiente de periódicos com informações a “emprestar” (Tabela 6.6), não havia, em nenhum destes periódicos, informações disponíveis para um dado indicador, impossibilitando assim a imputação para aquele indicador. Estes casos totalizaram 15 periódicos, os quais mesmo após o processo de imputação, permaneceram com missings em algum dos indicadores considerados na análise.

Um caso particular deste ocorrido foi exatamente os periódicos da Matemática Aplicada classificados como B5. Para estes, observou-se um total de 10 periódicos com missings no processo pós-imputação. A explicação para tal ocorrência deu-se inicialmente em função da regra de negócio ter filtrado apenas 6 dos 10 periódicos existentes para tal classificação, para serem imputados, isolando os 4 periódicos restantes como fonte para imputação. Porém, estes 4 periódicos possuíam informação apenas no indicador SciNet, de forma que não havia como imputar dados para os demais indicadores nos 6 periódicos separados anteriormente. Em geral, os casos mais semelhantes deste tipo de comportamento ocorreram para periódicos que possuíam informação apenas no indicador SciNet, estabelecendo assim um padrão de comportamento que impossibilitava a correta atuação do processo de imputação. Assim, mesmo optando por retirar esse indicador, o cálculo final da imputação registrou os missings resultantes deste processo.

Na Tabela 6.10 está em destaque os 4 registros da Matemática Aplicada, do total dos 10 registros dessa área, que levou em consideração o SciNet, os demais registros, a saber: 0973-1377, 0973-5607, 0973-7545, 1083-0464, 1679-365X e 1687-2177, correspondem aos 6 periódicos B5 identificados na tabela 6.3 que não levaram em consideração o SciNet.

6.3 Padronização dos dados

Como foi mencionado ao final da Seção 6.1.1, após a fase de imputação dos dados, faz-se necessário padronizar os indicadores bibliométricos dentro de cada área ou grupo: Matemática, Matemática Aplicada, Estatística e Outra área.

Tal procedimento faz-se necessário devido às peculiaridades dos padrões de publicação em cada área. Por exemplo, observa-se na Tabela 6.5 que o Fator de Impacto máximo observado em “Outra área” é de 18.97. No entanto, na área de Estatística tal valor não ultrapassa a 3.5. Tais diferenças são devidas, provavelmente, às especificidades próprias de cada área e valores máximos de Fator de Impacto, assim tão dispares, podem ser devidos a “fatores de escala” de modo que um periódico de “Outra área” com Fator de Impacto 18.97 não necessariamente é 5.42 vezes mais importante do que um periódico da área de Estatística cujo Fator de Impacto é 3.5. Em conclusão, a padronização dos dados de cada indicador dentro de cada área ou grupo em estudo visa atenuar os problemas aqui apontados.

Ao considerar os escores padronizados de cada indicador bibliométrico, dentro de cada área, tem-se que os dados resultantes passam a ser comparáveis, isto é, estes passam a ser

descritos em uma mesma escala, ou *escala universal*.

O processo de imputação, seguido da padronização dos indicadores bibliométricos, causa modificações nas distribuições dos indicadores bibliométricos.

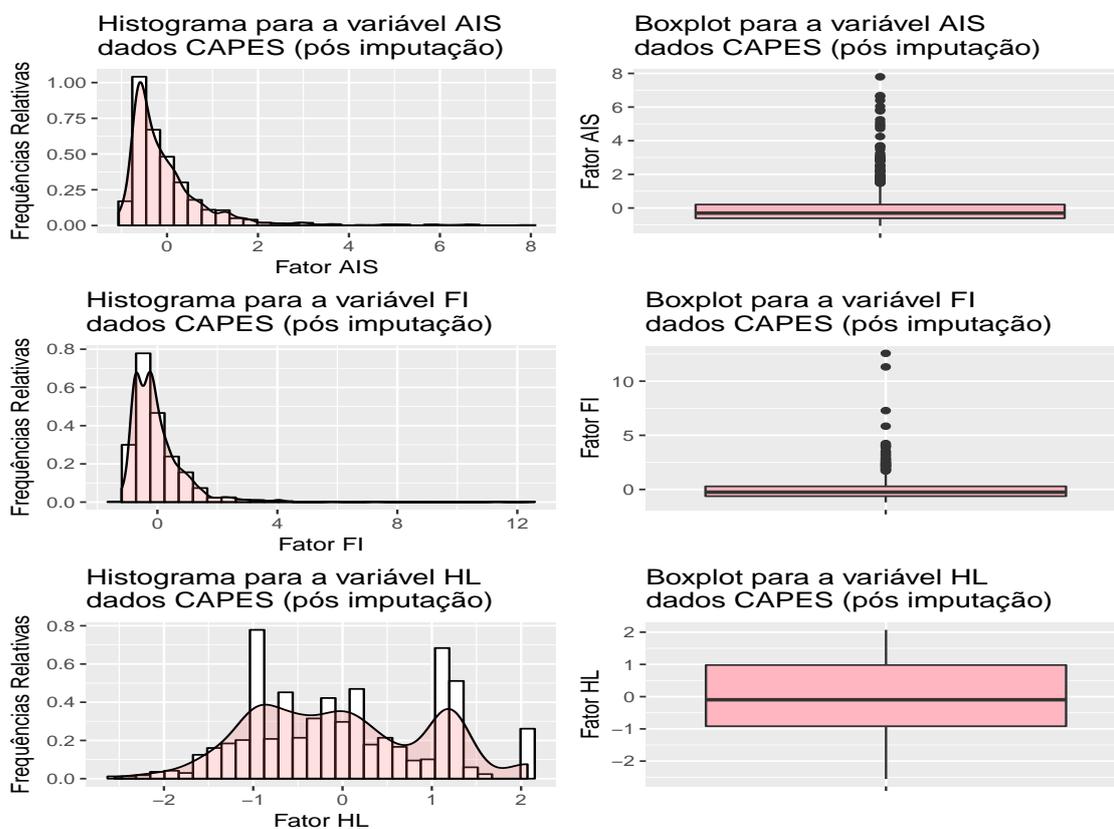


Figura 6.2: Representações gráficas para os indicadores bibliométricos: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) após o processo de imputação.

Conforme pode ser visualizado na Figura 6.2 e levando em consideração o processo de imputação apresentado na Seção 6.2 desta pesquisa, nota-se diferenças no comportamento dos indicadores bibliométricos, sendo que, a mais notória é visualizada no Fator de Meia Vida (HL), no qual o histograma apresenta blocos com comportamentos distintos.

Vale ressaltar que a análise supra citada foi realizado para a base imputada via mediana. No entanto, resultados similares foram observados quando utilizou-se imputação via média.

Apresenta-se a seguir algumas análises e descrição da padronização dos dados. Inicialmente, na Tabela 6.11 descreve-se algumas medidas sumarização dos indicadores bibliométricos padronizados (z) após a imputação via média. Na Tabela 6.12 apresenta-se análises similares considerando a imputação via mediana.

Tabela 6.11: *Análise descritiva dos indicadores bibliométricos padronizados (z) após a imputação via média.*

Indicador(z)	Core	Medidas resumo						
		Min.	1st. Qu.	Mediana	Média	3rd. Qu.	Max.	NA's
AIS	Est.	-1.043e+00	-6.760e-01	-3.236e-01	-7.034e-17	3.790e-01	4.726e+00	35
	Mat.	-8.142e-01	-6.054e-01	-3.292e-01	4.822e-17	1.351e-01	6.624e+00	86
	Mat. Apl.	-1.043e+00	-5.690e-01	-2.160e-01	4.800e-17	2.072e-01	6.398e+00	10
	Out. Área	-1.061e+00	-6.203e-01	-3.548e-01	9.065e-17	2.192e-01	7.800e+00	17
JCR	Est.	-1.169e+00	-6.251e-01	-2.365e-01	7.743e-17	3.002e-01	3.994e+00	1
	Mat.	-8.261e-01	-3.585e-01	-1.954e-01	-4.652e-17	1.845e-01	1.257e+01	4
	Mat. Apl.	-1.197e+00	-6.294e-01	-2.410e-01	-2.122e-17	4.217e-01	7.283e+00	10
	Out. Área	-9.769e-01	-7.005e-01	-3.324e-01	5.138e-17	3.161e-01	1.132e+01	NA
HL	Est.	-2.543e+00	-7.604e-01	5.959e-02	-6.008e-17	1.208e+00	1.208e+00	21
	Mat.	-2.431e+00	-5.722e-01	-2.005e-01	-8.173e-17	1.162e+00	1.162e+00	86
	Mat. Apl.	-2.020e+00	-6.817e-01	5.629e-03	1.559e-16	1.308e+00	1.308e+00	11
	Out. Área	-2.554e+00	-9.455e-01	-9.687e-02	1.916e-16	5.329e-01	2.071e+00	17

Fonte: Elaborado pelos autores (2017).

Tabela 6.12: *Análise descritiva dos indicadores bibliométricos padronizados (z) após a imputação via mediana.*

Indicador(z)	Core	Medidas resumo						
		Min.	1st. Qu.	Mediana	Média	3rd. Qu.	Max.	NA's
AIS	Est.	-1.022e+00	-7.132e-01	-3.495e-01	1.178e-16	3.930e-01	4.719e+00	35
	Mat.	-8.081e-01	-6.193e-01	-3.217e-01	-8.939e-18	1.088e-01	6.652e+00	86
	Mat. Apl.	-1.037e+00	-5.610e-01	-2.803e-01	4.125e-17	1.968e-01	6.433e+00	10
	Out. Área	-9.908e-01	-6.700e-01	-3.015e-01	-9.522e-17	2.469e-01	7.662e+00	17
FI	Est.	-1.144e+00	-7.108e-01	-2.154e-01	-5.092e-17	2.875e-01	3.996e+00	1
	Mat.	-7.958e-01	-3.864e-01	-1.659e-01	7.405e-17	1.949e-01	1.258e+01	4
	Mat. Apl.	-1.154e+00	-6.240e-01	-3.455e-01	-3.806e-17	4.028e-01	7.264e+00	10
	Out. Área	-9.519e-01	-6.779e-01	-3.223e-01	-1.755e-17	3.238e-01	1.124e+01	NA
HL	Est.	-2.548e+00	-7.703e-01	4.777e-02	5.822e-17	1.193e+00	1.193e+00	21
	Mat.	-2.244e+00	-8.827e-01	-1.439e-01	1.709e-16	1.139e+00	1.139e+00	86
	Mat. Apl.	-2.032e+00	-7.410e-01	-6.800e-02	-1.157e-16	1.314e+00	1.314e+00	10
	Out. Área	-2.408e+00	-1.041e+00	-9.410e-02	-8.988e-18	5.686e-01	2.052e+00	17

Fonte: Elaborado pelos autores (2017).

Observando as Tabelas 6.11 e 6.12, verifica-se que os dados dos indicadores bibliométricos, que possuem escala contínua, foram previamente separados por grupo: Matemática Pura, Matemática Aplicada, Estatística e Outras e, posteriormente foram calculados os escores padronizados de cada indicador, dentro de cada grupo. Com isto retira-se o efeito de média e escala de cada grupo, obtendo-se uma escala universal por indicador que, posteriormente, será discretizado.

6.4 Discretização dos dados

Bartolucci *et al.* (2015) observam que, ao discretizar os indicadores bibliométricos, em vez de se trabalhar diretamente com os valores em suas escalas originais, implica certa perda de informações, mas permite-nos evitar fortes pressupostos paramétricos. Também oferece algumas robustez aos erros de medição. No entanto, uma vez que os indicadores disponíveis são discretizados é essencialmente avaliar a sensibilidade dos resultados para a discretização assumida.

Inicialmente a discretização dos dados para uso nos modelos de TRI se fez pelos quintis, isto é, considerando-se 5 categorias equiespaçadas, cada uma contendo 20% dos dados. Em análises preliminares de TRI observou-se um número excessivo de empates dos escores de proficiência dos periódicos. Além disso observou-se um mal ajustamento dos modelos. Para tentar atenuar o problema, considerou-se 8 categorias (escala discretizada considerando 8 categorias, cada uma representando 12.5% da distribuição). Em ambos os casos criou-se

uma categoria adicional para acomodar os dados dos periódicos para os quais não foi possível efetuar a imputação em dado indicador bibliométrico. Dessa forma, a esta categoria adicional atribuiu-se o valor 1, a categoria de valores mais baixo. As 8 categorias resultantes da discretização atribuiu-se valores de 2 a 9.

Na Tabela 6.13, apresenta-se os limites das categorias resultantes da discretização e que foram utilizados neste trabalho. Os dados foram previamente imputados através da média e padronizados (vide Seções 6.2 e 6.3). Cada categoria de 2 a 9 representa 12.5% da distribuição da respectiva variável contínua, sendo o valor 1 atribuído à categoria onde não foi possível a imputação.

Tabela 6.13: *Representação tabular dos indicadores bibliométricos disponíveis na base de dados de 2012. Os dados foram previamente imputados (via média) e padronizados. Cada categoria de 2 a 9, encerra 12.5% da distribuição da variável contínua (AIS, FI ou HL).*

Variáveis	Categorias	Limites assumidos pelos itens
AIS	1	item sem classificação
	2	item 1 < -0.67597009
	3	-0.67597009 ≤ item 1 < -0.61039840
	4	-0.61039840 ≤ item 1 < -0.44645884
	5	-0.44645884 ≤ item 1 < -0.30228292
	6	-0.30228292 ≤ item 1 < -0.09434683
	7	-0.09434683 ≤ item 1 < +0.20792003
	8	+0.20792003 ≤ item 1 < +0.82442518
	9	item 1 > +0.82442518
FI	1	item sem classificação
	2	item 2 < -0.77238042
	3	-0.77238042 ≤ item 2 < -0.62531238
	4	-0.62531238 ≤ item 2 < -0.36125013
	5	-0.36125013 ≤ item 2 < -0.25271397
	6	-0.25271397 ≤ item 2 < -0.02907665
	7	-0.02907665 ≤ item 2 < +0.28376928
	8	+0.28376928 ≤ item 2 < +0.85913971
	9	item 2 > +0.85913971
HL	1	item sem classificação
	2	item 3 < -1.07082086
	3	-1.07082086 ≤ item 3 < -0.91808079
	4	-0.91808079 ≤ item 3 < -0.53088022
	5	-0.53088022 ≤ item 3 < -0.09687468
	6	-0.09687468 ≤ item 3 < +0.23507291
	7	+0.23507291 ≤ item 3 < +0.98021855
	8	+0.98021855 ≤ item 3 < +1.20755049
	9	item 3 > +1.20755049

Fonte: Elaborado pelos autores (2017).

Analogamente, agora considerando-se a imputação pela mediana, apresenta-se os limites das categorias resultantes da discretização.

Tabela 6.14: Representação tabular dos indicadores bibliométricos disponíveis na base de dados de 2012. Os dados foram previamente imputados (via mediana) e padronizados. Cada categoria de 2 a 9, encerra 12.5% da distribuição da variável contínua (AIS, FI ou HL).

Variáveis	Categorias	Limites assumidos pelos itens
AIS	1	item sem classificação
	2	item 1 < -0.71349621
	3	-0.71349621 ≤ item 1 < -0.63102902
	4	-0.63102902 ≤ item 1 < -0.43980874
	5	-0.43980874 ≤ item 1 < -0.29560554
	6	-0.29560554 ≤ item 1 < -0.07786211
	7	-0.07786211 ≤ item 1 < +0.20866298
	8	+0.20866298 ≤ item 1 < +0.84313632
	9	item 1 > +0.84313632
FI	1	item sem classificação
	2	item 2 < -0.76052832
	3	-0.76052832 ≤ item 2 < -0.65140113
	4	-0.65140113 ≤ item 2 < -0.38636849
	5	-0.38636849 ≤ item 2 < -0.24120662
	6	-0.24120662 ≤ item 2 < -0.01389995
	7	-0.01389995 ≤ item 2 < +0.28160796
	8	+0.28160796 ≤ item 2 < +0.86878380
	9	item 2 > +0.86878380
HL	1	item sem classificação
	2	item 3 < -1.04088576
	3	-1.04088576 ≤ item 3 < -0.88270752
	4	-0.88270752 ≤ item 3 < -0.47440112
	5	-0.47440112 ≤ item 3 < -0.09410307
	6	-0.09410307 ≤ item 3 < +0.29104294
	7	+0.29104294 ≤ item 3 < +1.10515502
	8	+1.10515502 ≤ item 3 < +1.19305206
	9	item 3 > +1.19305206

Fonte: Elaborado pelos autores (2017).

Ao inspecionar as Tabelas 6.13 e 6.14, aparentemente as diferenças ao considerar a imputação pela média ou mediana são mínimas.

Na Seção 6.5 algumas análises adicionais serão apresentadas para entendermos melhor os dados e decidirmos qual é o melhor tipo de imputação mais adequado, se média ou mediana.

6.5 Algumas análises descritivas adicionais para os dados pré-processados

No intuito de investigar possíveis relacionamentos entre os indicadores bibliométricos, algumas análises descritivas são apresentadas.

A Figura 6.3 apresenta-se os diagramas de dispersão dos indicadores após a fase de pré-processamento (imputação pela média ou mediana e padronização). Observa-se que o AIS e FI descrevem um relacionamento aparentemente linear, ao passo que o HL não parece estar relacionado com os demais indicadores.

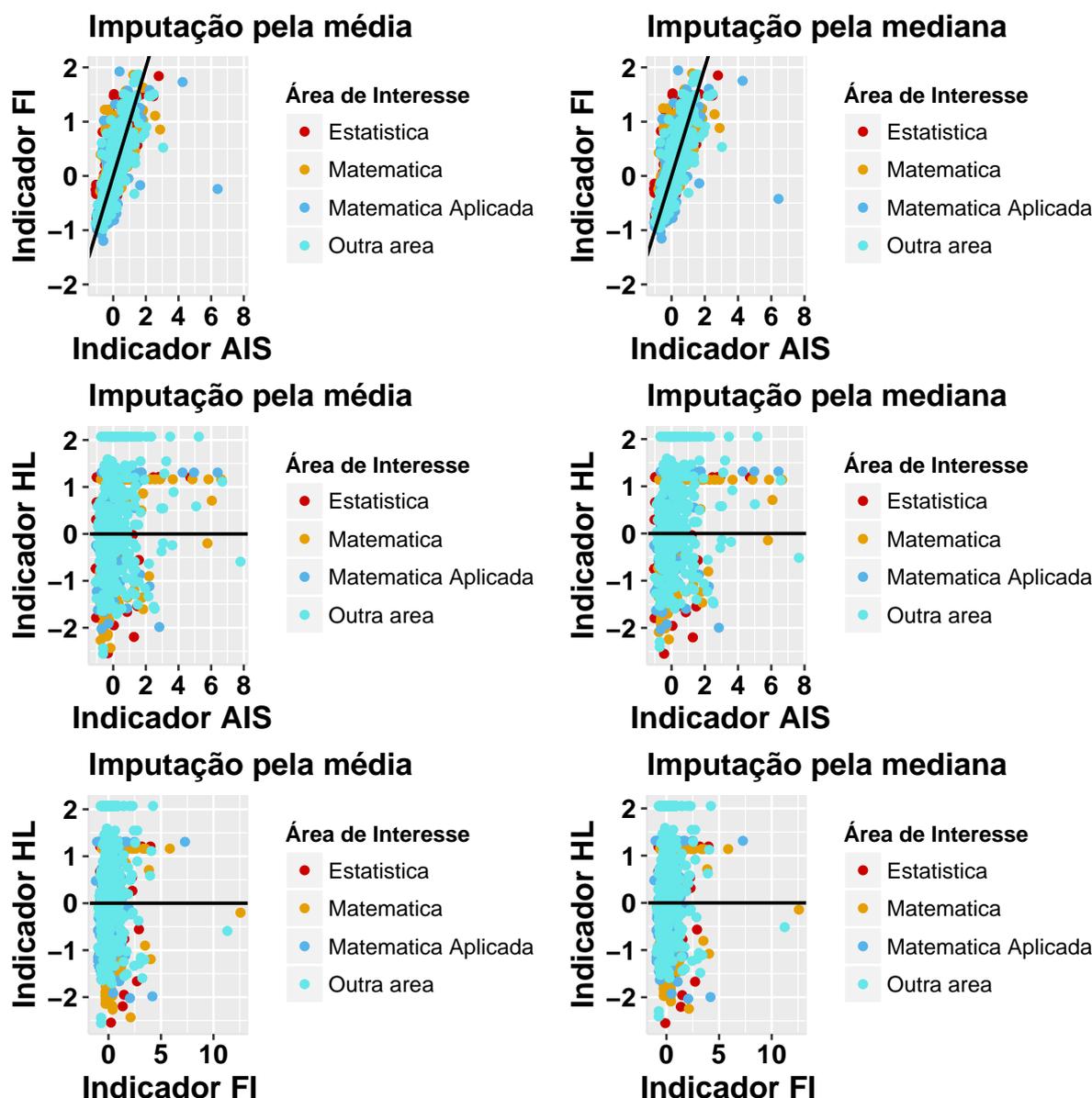


Figura 6.3: Diagramas de dispersão para os indicadores bibliométricos observados: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) após pré-processamento.

A matriz de correlações, \mathbf{R} , entre o AIS, FI e HL é dada a seguir, e reforça as observações feitas acima, uma vez que AIS e FI apresentam forte correlação (0.812), ao passo que as correlações entre o HL e os demais indicadores estão abaixo de 0.3.

$$R = \begin{bmatrix} & \text{AIS} & \text{FI} & \text{HL} \\ \text{AIS} & 1.0000000 & 0.8128488 & 0.2867314 \\ \text{FI} & 0.8128488 & 1.0000000 & 0.1438292 \\ \text{HL} & 0.2867314 & 0.1438292 & 1.0000000 \end{bmatrix}.$$

Os gráficos tridimensionais apresentados nas Figuras 6.4 e 6.5 corroboram com as observações feitas até o momento sobre as relações entre os indicadores pré-processados, quer seja pela média ou mediana. Em qualquer um destes casos, observa-se que o HL assume uma ampla gama de valores sem, aparente, relação com os valores do AIS e FI. Estes últimos tendem a ter um comportamento similar: para pequenos (grandes) valores de FI observa-se pequenos (grandes) valores de AIS.

Gráfico 3D dos 3 indicadores
Base de dados após imputação via média

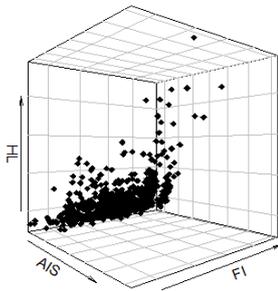


Figura 6.4: Gráfico 3D para os indicadores bibliométricos observados: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) após o processo de imputação via média e padronização.

Gráfico 3D dos 3 indicadores
Base de dados após imputação via mediana

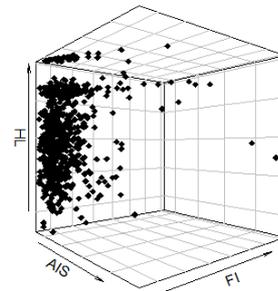


Figura 6.5: Gráfico 3D para os indicadores bibliométricos observados: Fator AIS, Fator de Impacto (FI) e o Fator de Meia Vida (HL) após o processo de imputação via mediana e padronização.

A análise de componentes principais envolvendo os indicadores AIS, FI e HL, após o pré-processamento, revela que os dois eixos principais explicam mais de 90% da variância nos dados, sendo que o primeiro eixo principal mais de 60% da mesma. Além disso, o primeiro eixo principal é descrito por uma combinação linear do AIS e do FI, enquanto o segundo eixo principal é descrito, majoritariamente, pelo HL.

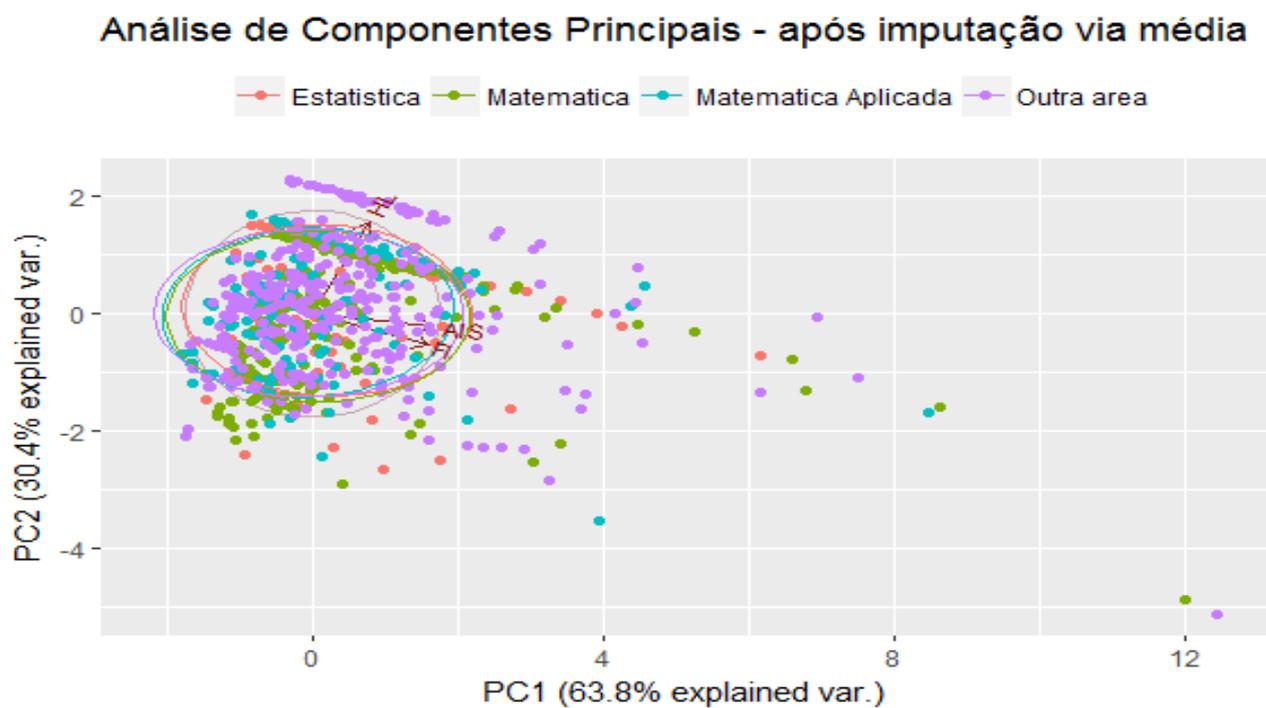


Figura 6.6: *Análise de Componentes Principais (PCA) após o processo de imputação via média.*

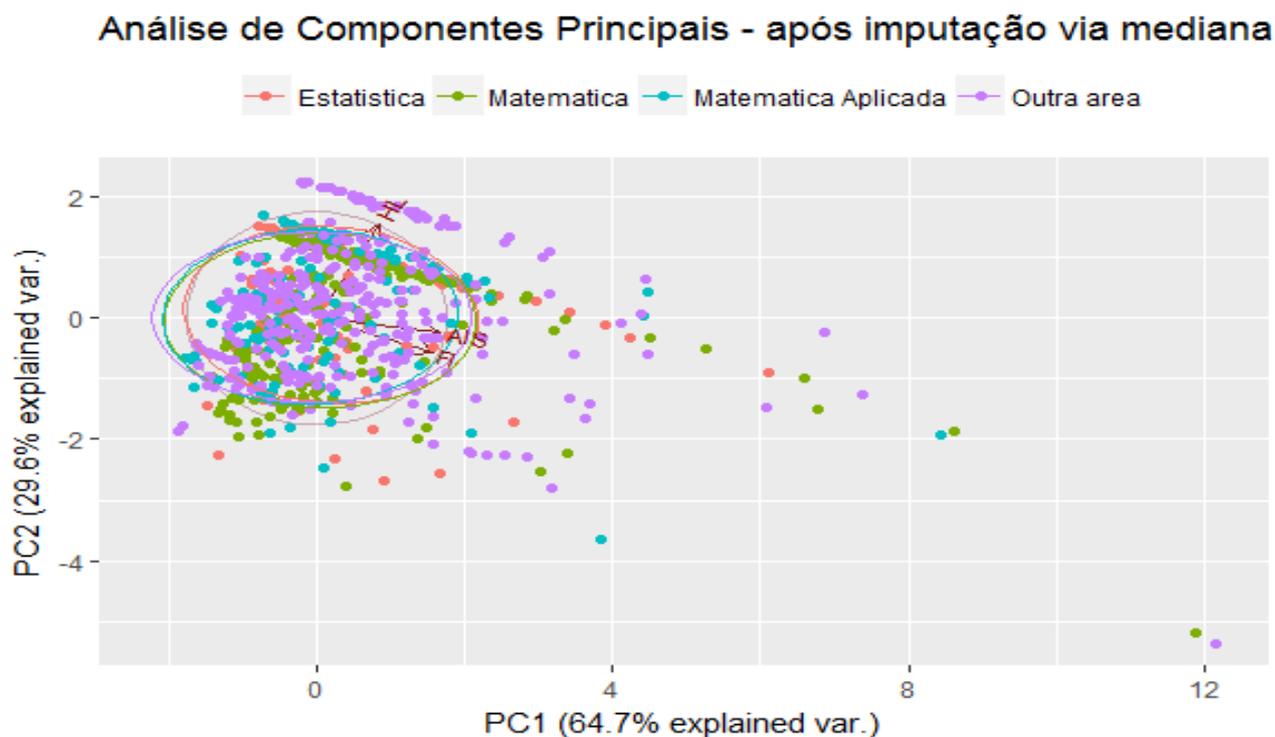


Figura 6.7: *Análise de Componentes Principais (PCA) após o processo de imputação via mediana.*

Na próxima seção descreve-se algumas análises via CART para especificações de pontos de cortes nos indicadores bibliométricos.

6.6 Classificação dos periódicos via CART

Nesta seção apresenta-se uma análise de CART tentativa para a classificação dos periódicos. Objetiva-se estimar “pontos de corte” adequados nos indicadores bibliométricos pré-processados (fases de imputação e padronização) que permitam discriminar os periódicos de uma forma bastante prática e de uso simplificado por parte dos agentes da CAPES.

No que tange a apresentação de uma árvore de decisão, segundo [Onoda e Ebecken \(2001\)](#), o critério utilizado para realizar as partições é o da utilidade do atributo para a classificação. Aplica-se, por este critério, um determinado ganho de informação a cada atributo. O atributo escolhido como atributo teste para o corrente nó é aquele que possui o maior ganho de informação. A partir desta aplicação, inicia-se um novo processo de partição.

Os autores supra citados enfatizam que nos casos em que a árvore é usada para classificação, os critérios de partição mais conhecidos são baseados na entropia e no índice Gini ([Gini e Lexis, 1912](#)). Para mais detalhes sobre o índice de Gini vide [Ceriani e Verme \(2012\)](#).

Para as análises desta seção considerou-se o índice de Gini que está implementado no pacote `tree: Fit a Classification or Regression Tree` do *software* R.

Neste trabalho utilizou-se o algoritmo CART - do inglês Classification and Regression Tree ([Breiman *et al.*, 1984](#)), que utiliza a partição recursiva binária. [Onoda e Ebecken \(2001\)](#) enfatiza que as principais características deste algoritmo são: definir o conjunto de regras para dividir cada nó da árvore; decidir quando a árvore está completa; associar cada nó terminal a uma classe ou a um valor preditivo no caso de regressão. A partir deste método de classificação, objetiva-se propor pontos de corte nos indicadores bibliométricos que sejam definidos de acordo com critérios estatísticos. Os dados utilizados foram pré-processados, com imputação pela média.

Neste trabalho utilizou-se o modelo de regressão a seguir, em que o AIS figura como variável resposta, ao passo que o FI e o HL, como covariáveis.

$$AIS \sim FI + HL + \varepsilon. \quad (6.1)$$

Além deste modelo, citado acima, outros foram testados: $FI \sim AIS + HL + \varepsilon$ e $AIS \sim FI + \varepsilon$. Estes dois últimos modelos estão apresentados no **Apêndice J** desta pesquisa. Neste apêndice apresenta-se uma tabela comparativa entre os 3 modelos supra citados.

A Figura 6.8 descreve a representação da árvore de decisão associada ao modelo (6.1). Cade cada nó de decisão contém um teste para algum atributo, cada ramo descendente corresponde a um possível valor deste atributo, o conjunto de ramos são distintos, cada folha está associada a uma classe e, cada percurso da árvore, da raiz à folha corresponde uma regra de classificação.

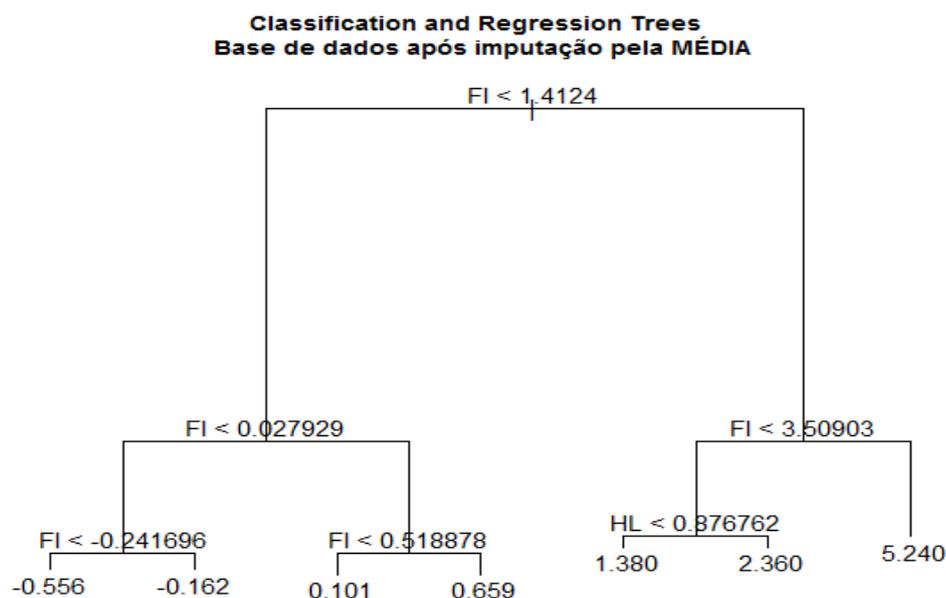


Figura 6.8: Representação gráfica do CART para a base de dados após imputação pela média.

Conforme pode ser visualizado na Figura 6.8, esta árvore de decisão apresenta 6 folhas, indicando as classes que se tem e 7 nós de decisão. O número em cada folha é a média da resposta para as observações que lá estão, neste caso, o modelo tem como variável resposta o Fator AIS e como covariáveis, o Fator de Impacto (FI) e o Fator de Meia Vida (HL).

A partir de uma árvore de decisão é possível derivar regras. As regras são escritas considerando o caminho do nó raiz até uma folha da árvore. A derivação de regras e a construção da árvore de decisão são geralmente utilizados em conjunto. Uma das principais razões para esta utilização é o fato das árvores de decisão tenderem a crescer muito, daí muitas vezes serem substituídas por regras.

Na Tabela 6.15 apresenta-se as regras que foram derivadas a partir da árvore de decisão/classificação representada na Figura 6.8.

Para cada uma das áreas (Matemática, Matemática Aplicada, Estatística e Outra área) utilizou-se as mesmas regras de classificação, descritas pela coluna 2 da Tabela 6.15, uma vez que os dados pré-processados estão descritos em uma escala universal.

A Figura 6.9 proporciona uma outra representação das regras derivadas através da árvore de decisão/classificação apresentada em 6.8.

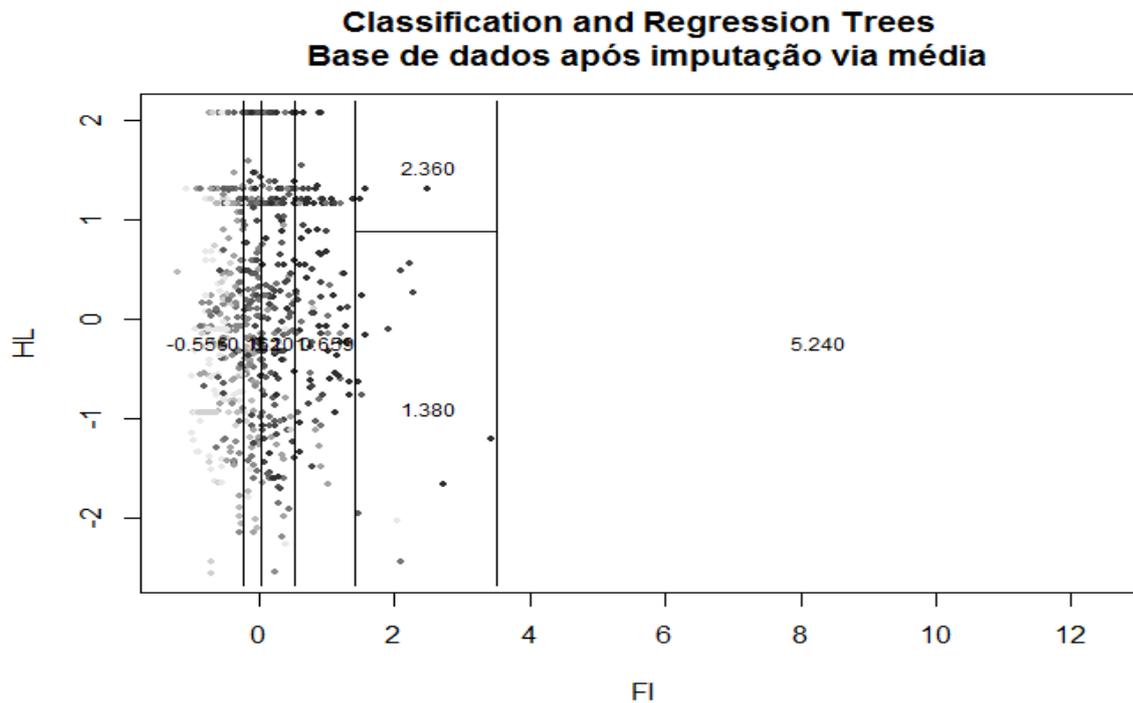


Figura 6.9: Representação gráfica das partições do CART após imputação pela média.

Na Figura 6.9 observa-se que poucas observações determinam pontos de corte na escala, a exemplo dos pontos de corte $1.412 \leq FI < 3.509$ (valor médio 2.36, para $HL > 0.876$ e valor médio 1.380, para $HL < 0.876$) e $FI \geq 3.509$. Também observa-se que para $FI < -0.242$ também não há muitas observações. Dessa forma, provavelmente as caudas da distribuição de FI não estão bem representadas, o que pode afetar, negativamente, os resultados destas análises.

Tabela 6.15: Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo o CART.

Categoria	Critério	AIS médio	Número de periódicos por área			
			Mat.	Mat. Apl.	Est.	Out. área
1	$FI \geq 3.509$	5.240	5	2	1	4
2	$1.412 \leq FI < 3.509$ e $HL \geq 0.876$	2.360	5	3	6	9
3	$1.412 \leq FI < 3.509$ e $HL < 0.876$	1.380	5	3	6	21
4	$0.519 \leq FI < 1.412$	0.659	47	23	16	66
5	$0.027 \leq FI < 0.519$	0.101	71	27	20	66
6	$-0.242 \leq FI < 0.027$	-0.162	79	22	17	56
7	$FI < -0.242$	-0.556	111	73	33	244
Total			323	153	99	466

Fonte: Elaborado pelos autores (2017).

De acordo com a Tabela 6.15, as quatro áreas apresentam uma concentração maior de periódicos na categoria $FI < -0.242$, em que o AIS médio é igual a -0.556.

Deve-se salientar que a soma das colunas referente ao total na Tabela 6.15, totaliza 1041 periódicos, de um total de 1198 registros que forma a base de dados. A diferença, 157, corresponde aos dados que não puderam ser imputados.

O grande problema com as regras da Tabela 6.15 é o desenvolvimento das categorias. as três primeiras categorias, 1 a 3, são muito rarefeitas, sendo que as três juntas agregam apenas

70 periódicos, representando diminutos 6.7% do total. Quando as quatro primeiras categorias são agregadas, obtem-se 222 periódicos inclusos, que perfaz 21.3% do total. Dessa forma, com as quatro categorias juntas ter-se-ia periódicos suficientes, apenas, para as classificações A1 e A2 da CAPES, uma vez que estas devem perfazer no máximo 25% do total dos periódicos.

Quando às quatro primeiras categorias junta-se os dados da categoria 5, obtem-se, apenas, 39% dos dados. Somente quando, às cinco primeiras categorias junta-se os dados da sexta categoria é que se obtem 55,7% dos dados, percentual próximo ao especificado pela CAPES, para a inclusão de periódicos A1, A2, e B1, sendo este percentual de no máximo 50% dos periódicos.

Dessa forma, os periódicos que estão inclusos na categoria 7, que sozinho engloba 44.3% dos periódicos, teriam que ser distribuídos nos demais estratos especificados pela CAPES: B2, B3, B4 e B5. No entanto, não há pontos de corte adicionais para classificar os periódicos em tais estratos.

Em resumo, infelizmente o CART não proporcionou, com os dados disponíveis, resultados úteis para a classificação dos periódicos.

N próximo capítulo descreve-se os procedimentos via modelo TRI politômico que foram desenvolvidas nesta dissertação para a ordenação e classificação dos periódicos da base Qualis 2012.

Capítulo 7

Ordenação de periódicos científicos via Teoria da Resposta ao Item

O foco deste capítulo consiste na descrição dos métodos e resultados alcançados na pesquisa para ordenação e classificação dos periódicos científicos que fazem parte do Qualis 2012 do comitê de Matemática e Estatística.

Na seção que segue, apresenta-se, uma descrição detalhada das etapas desenvolvidas na presente pesquisa.

7.1 Etapas da pesquisa

Nesta seção apresenta-se uma descrição detalhada das etapas que foram adotadas nesta dissertação com vistas à ordenação e classificação dos periódicos científicos pela CAPES em 2012 na área de Matemática e Estatística.

7.1.1 Pré processamento dos dados entre as etapas 1 a 5

Etapa 1: Inicialmente, separou-se os periódicos por área, isto é, Matemática Pura, Matemática Aplicada, Estatística e Outra Área.

Etapa 2: Para a construção da base de dados de 2012 procedeu-se a imputação de dados. Uma discussão detalhada acerca deste tema pode ser encontrada na Seção 6.2.

Etapa 3: Para cada uma das áreas (separadamente) tomar-se os escores padronizados para cada um dos indicadores. No **Apêndice F** apresenta-se algumas das propostas que podem ser utilizadas na padronização. No Capítulo 6, Seção 6.3 são discutidas as razões pelas quais o uso de padronização dos dados é necessária.

Etapa 4: Após a **Etapa 3** procedeu-se a discretização dos dados. A discretização dos dados é necessária para que estes possam ser usados em análises de modelos politômicos da Teoria da Resposta ao Item. Mais informações sobre o processo de discretização utilizado nesta dissertação podem ser encontrados na Seção 6.4.

Etapa 5: Escolha do método de imputação dos dados faltantes. Este tópico está descrito na Seção 6.5.

7.1.2 Análise dos dados

Etapa 6: Uso de árvore de decisão na classificação dos periódicos. O intuito desse procedimento é averiguar a possibilidade de utilizar a referida metodologia na especificação de pontos de corte na escala contínua dos indicadores bibliométricos, já previamente pré-processados, de modo a fornecer aos agentes da CAPES um processo simples e familiar de classificação dos periódicos. Mais informações podem ser encontradas no Capítulo 6.6.

Etapa 7: Uso do Modelos de Escala Gradual de Samejima (1969) na estimação dos escores do *valor* de um periódico Nesta etapa objetiva-se construir uma escala ou métrica única que permita ordenar os periódicos de todos os grupos ou áreas em estudo (Matemática Pura, Matemática Aplicada, Estatística e Outras áreas). Posteriormente, objetiva-se classificar os periódicos nos estratos A1 a B5 da CAPES.

Etapa 8: Classificação dos periódicos. Os critérios de classificação dos periódicos a serem descritos nesta dissertação atendem aos critérios especificados pela CAPES, a saber.

Seja N o número total de periódicos nos quais pesquisadores brasileiros publicam artigos, em um determinado ano, em dado Comitê da CAPES.

Sejam n_{A_1} , n_{A_2} n_{B_1} o número de periódicos classificados pelo Comitê nos estratos A_1 , A_2 e B_1 . A CAPES destaca que

(a) A proporção de periódicos classificados como A_1 ou A_2 por um dado Comitê seja inferior a 25% do total de periódicos avaliados, isto é, $(n_{A_1} + n_{A_2})/N < 0.25$.

(b) A proporção de periódicos classificados como A_1 , A_2 ou B_1 não ultrapasse a 50% do total de periódicos avaliados, isto é, $(n_{A_1} + n_{A_2} + n_{B_1})/N \leq 0.5$.

Maiores informações sobre os processos de classificação dos periódicos utilizados nesta dissertação são descritos na Seção 7.3.

A seguir discute-se o procedimento utilizado na escolha do método de imputação dos dados faltantes, que foi mencionado na **Etapa 5**.

7.2 Escolha da estratégia de imputação dos dados

Para decidir acerca da melhor estatística de sumarização, se média ou mediana, a ser utilizada no processo de imputação dos dados, foram empreendidas algumas análises preliminares de modelos politômicos de TRI. Tais modelos foram ajustados tanto considerando imputações feitas utilizando a média quanto a mediana. Para tanto, considerou-se os dados de cada indicador bibliométrico e o cruzamento das informações referentes ao grupo ou área de dado periódico e o estrato ao qual o periódico foi classificado pela CAPES em 2012 (vide Tabela 6.3). Utilizando tais informações, colocou-se os referidos valores de média e mediana. Posteriormente os dados foram padronizados dentro de cada grupo (vide Seção 6.3) e discretizados para uso nos modelos politômicos.

Na discretização dos dados (vide Seção 6.4) utilizou-se 8 categorias, cada uma representando 12.5% dos dados. Adicionalmente, criou-se uma categoria artificial para acomodar 15 observações para as quais não foi possível efetuar a imputação. A esta categoria adicional atribuiu-se o valor 1. As 8 categorias resultantes da discretização atribuiu-se valores de 2 a 9.

Na tentativa de avaliar a qualidade/deficiências dos dois métodos de imputação dos dados (vide Seção 6.2), modelos de TRI politômicos foram ajustados aos dados discretizados e previamente imputados (pela média/mediana). Os dados foram, então padronizados (vide Seção 6.3). Na Figura 7.1 apresenta-se um diagrama de dispersão dos postos dos escores de proficiência, obtidos pela TRI, de cada um dos 1189 periódicos, considerando-se as imputações pela média e pela mediana. Apesar de boa concordância que se observa nos pontos situados na reta de 45 graus, há um número substancial de pontos para os quais tal concordância não é observada.

O gráfico da parte inferior da Figura 7.1 foi construído ao considerarmos a razão entre a diferença dos postos dos escores obtidos quando os dados foram imputados pela média e pela mediana e o total de periódicos na amostra ($N = 1189$). Quando há concordância entre os métodos, espera-se que a maioria dos pontos se situem na altura do zero. No entanto, observa-se que a grande maioria dos pontos se situam acima do zero, indicando que as imputações feitas pela mediana conduzem a postos, em geral, maiores do que aqueles observados com as imputações feitas utilizando a média.

Devido a restrições de tempo não foi possível empreender novas análises ou métodos alternativos de imputação de dados. Por ora optou-se em empreender as análises utilizando dados imputados pela média. No entanto, reconhecemos que as análises empreendidas até o momento são um tanto inconclusivas.

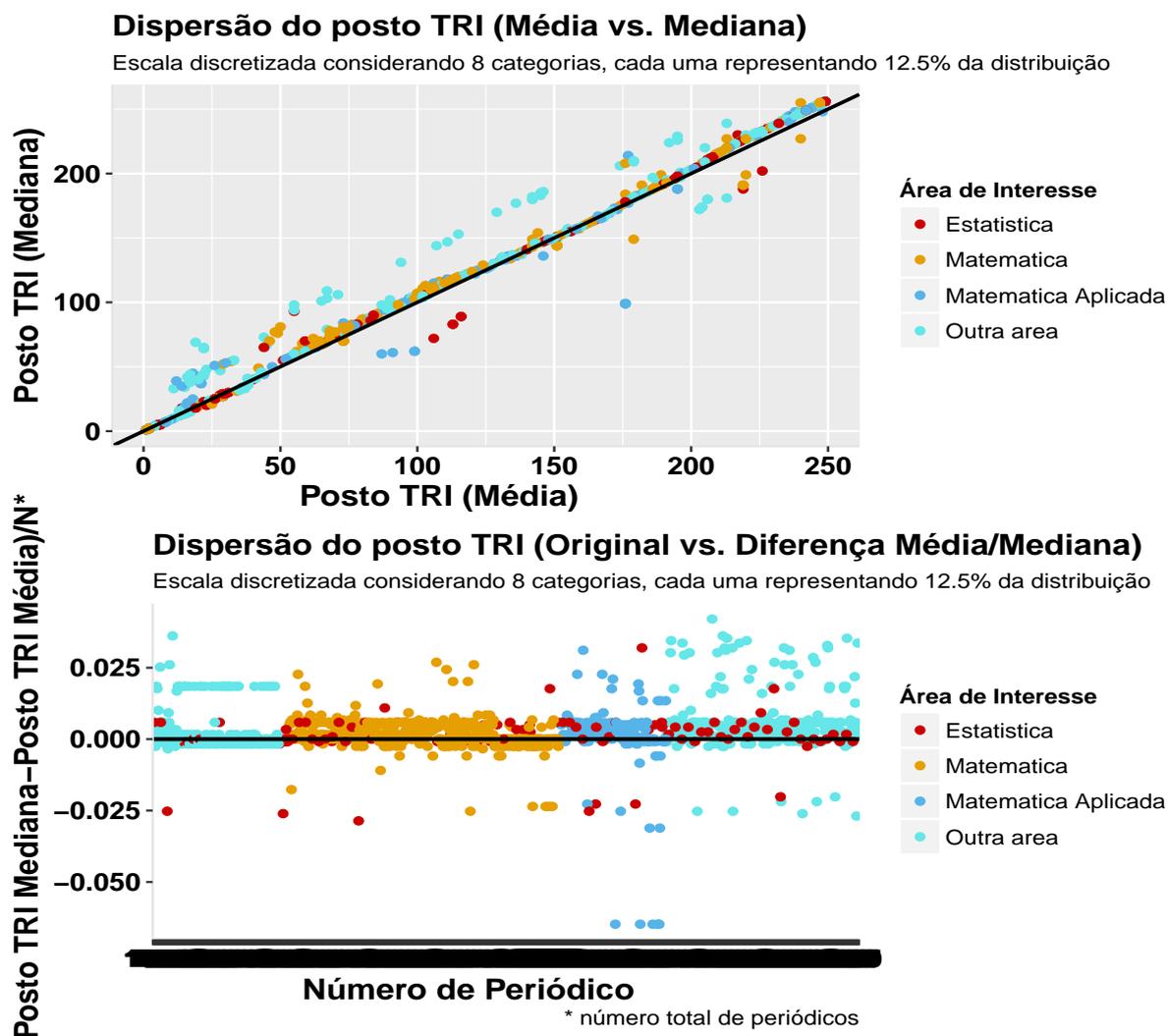


Figura 7.1: Representação gráfica da dispersão do posto TRI (Média versus Mediana)

Apresenta-se, na sequência, na Seção 7.3 os resultados e discussões via TRI com dados politômicos da preponente pesquisa.

7.3 Ordenação e classificação dos periódicos via TRI

Nesta seção, apresenta-se uma análise dos dados com as estimativas iniciais dos parâmetros dos itens (vide Tabela 7.2) obtidas com o auxílio dos *softwares* Multilog e R, versão 3.3.1.

7.3.1 Estimação dos parâmetros TRI

Nesta seção apresenta-se os resultados preliminares da estimação dos parâmetros da TRI politômica aplicada ao conjunto de 3 indicadores bibliométricos, AIS, FI e HL, dos 1189 periódicos que compõem a base de dados do ano-base de 2012, após as etapas de pré-processamento, e imputação via média (vide Capítulo 6).

Esta base dados contém 9 variáveis, no entanto, apenas 5 destas são úteis para este trabalho. Vale ressaltar que dessas 5 variáveis tem-se que uma variável representa o ISSN e uma outra variável representa a área ou grupo a qual está os periódicos. Por fim, as 3 variáveis, restantes, caracterizam os indicadores bibliométricos que compõem a base considerada.

Os dados dos 3 indicadores bibliométricos, que originalmente possuem escala contínua, foram previamente separados por grupo: Matemática, Matemática Aplicada, Estatística e Outras áreas e, posteriormente foram calculados os escores padronizados de cada indicador, dentro de cada grupo. Com isto retira-se o efeito de média e escala de cada grupo, obtendo-se uma escala universal por indicador que, posteriormente, será discretizado considerando 8 categorias, cada uma representando 12.5% da distribuição. Nesta mesma linha, enfatiza-se que uma categoria adicional foi criada para representar dados faltantes. Criou-se 9 categorias para representar os dados discretizados, sendo a categoria 1 relativa aos dados que não puderam ser imputados (vide Seção 6.3). Ao final do pré-processamento tem-se 3 itens politômicos (AIS, FI e HL), cada um com 9 categorias.

Inicialmente ajustou-se um modelo politômico de Resposta Gradual de Samejima (1969) considerando-se as 9 categorias (vide Seção 6.4). Após uma sucessão extensiva de análises, obteve-se que o modelo mais parsimonioso e que proporciona um melhor ajuste (melhor descrição dos dados), incorpora um número bem menor de categorias para cada um dos itens (AIS, FI e HL).

O modelo inicial era composto de um parâmetro de discriminação para cada um dos itens (a_i , $i = 1, 2, 3$) e 8 parâmetros de dificuldade para cada item, ($b_{i,j}$ $j = 1, \dots, 8$). Utilizou-se o *software* Multilog para as análises.

Com junções de categorias adjacentes para os itens, necessárias à obtenção de um ajuste adequado, o modelo final é composto de quatro categorias ordenadas para o item AIS, de modo que há três parâmetros de dificuldade para este item ($b_{1,1}, b_{1,2}, b_{1,3}$), três categorias ordenadas para o item FI, implicando em dois parâmetros de dificuldade para o item FI ($b_{2,1}$ e $b_{2,2}$), e similarmente ao FI, três categorias ordenadas para o item HL, implicando em dois parâmetros de dificuldade, ($b_{3,1}$ e $b_{3,2}$)

Na Tabela 7.1 estão descritas as novas categorias estabelecidas para cada um dos itens ao combinarmos categorias adjacentes considerando-se as categorias preliminares c_1 a c_9 .

Tabela 7.1: *Categorias estabelecidas para cada um dos itens ao combinarmos categorias adjacentes considerando-se as categorias preliminares C_1 a C_9 .*

Item	Novas categorias			
	Categoria 1	Categoria 2	Categoria 3	Categoria 4
AIS	C_1 e C_2	C_3, C_4 e C_5	C_6, C_7 e C_8	C_9
FI	C_1 e C_2	C_3, C_4 e C_5	C_6, C_7, C_8 e C_9	-
HL	C_1, C_2 e C_3	C_4, C_5 e C_6	C_7, C_8 e C_9	-

Fonte: Elaborado pelos autores (2017).

Na Tabela 7.2 apresenta-se os resultados da estimação dos parâmetros de cada item

Tabela 7.2: *Estimativas dos parâmetros dos itens AIS, FI e HL da base de dados imputada, CAPES/Qualis ano-base 2012, $N = 1189$ periódicos. EP (erro padrão da estimativa).*

Parâmetros	Item		
	AIS	FI	HL
a_i (EP)	4.45 (0.26)	4.54 (0.30)	1.34 (0.10)
$b_{i,1}$ (EP)	-0.77 (0.04)	-1.18 (0.04)	-0.67 (0.07)
$b_{i,2}$ (EP)	0.18 (0.03)	-0.08 (0.03)	0.66 (0.08)
$b_{i,3}$ (EP)	1.31 (0.04)	-	-

Fonte: Elaborado pelos autores (2017).

Considerando a Tabela 7.2, verifica-se que os itens com maior poder de discriminar os periódicos são o AIS e o FI. Esses itens, como pode ser visualizado na Figura 7.2, são os que possuem as Curvas Características do Item (CCI) mais acentuadas (Vide Seção 3.2).

Dessa forma, estes itens trazem informações importantes no cômputo do *valor* de um periódico.

Ainda de acordo com a Tabela 7.2, em comparação ao AIS e FI, verifica-se que o item HL é um item que pouco contribui para discriminar os periódicos.

Como mencionado anteriormente, as Curvas Características do Item (CCI) (Vide Seção 3.2) associadas ao ajuste em discussão estão apresentadas na Figura 7.2.

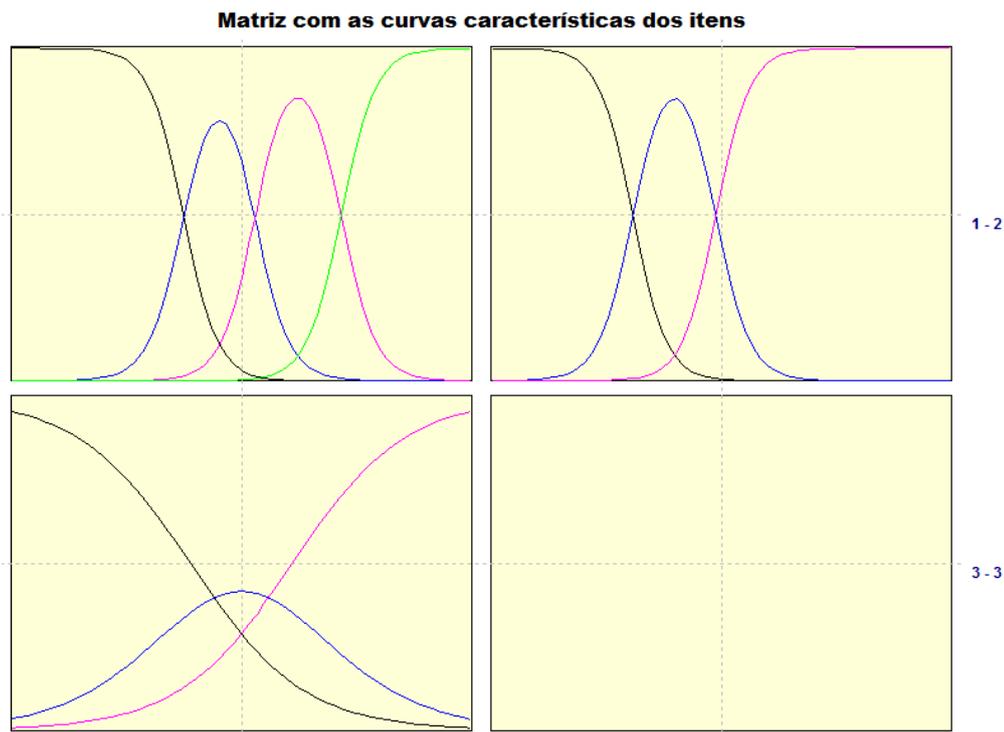


Figura 7.2: Representação gráfica da curvas características dos itens. Elaborado pelos autores (2017).

No canto superior esquerdo da Figura 7.2 encontra-se a CCI associada ao AIS, que contém 4 curvas, cada uma associada a uma categoria ordenada previamente recodificada. No eixo horizontal estão representados os valores do traço latente e no eixo vertical as probabilidades associadas ao modelo. No canto superior direito representa-se a CCI associada a FI e no canto inferior direito a CCI associada ao HL.

Um detalhamento da Figura 7.2 é apresentado nas Figuras 7.3 a 7.8.

De acordo com a Figura 7.3, a CCI associada ao AIS indica que os periódicos com nível de traço latente até -1 possuem uma maior probabilidade de pertencer à primeira categoria ordenada. Já os periódicos com habilidade entre -1 e 0.5 aproximadamente, possuem maior probabilidade de estarem situados na categoria 2. Jornais com desempenho entre 0.5 e 1.5, aproximadamente, apresentam maior probabilidade de estar situada na categoria 3. Já os jornais com desempenho superior a 1.5, apresentam maior probabilidade de estarem situados na categoria 4.

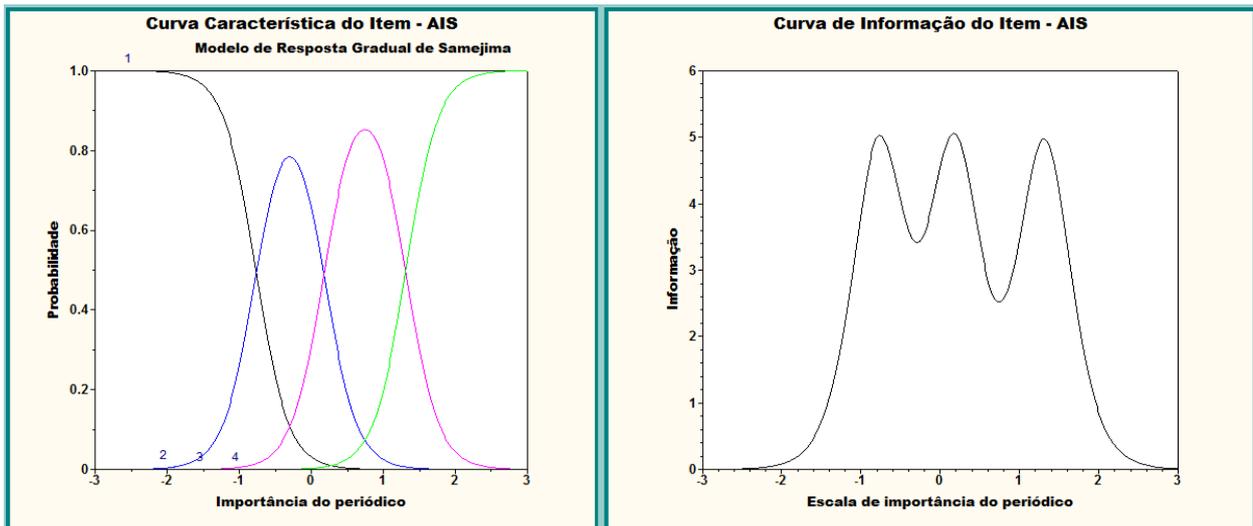


Figura 7.3: Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

Com base na Figura 7.3 a Curva de Informação do Item (CII) (vide Seção 3.5) o seu valor máximo atinge para os traços latentes no intervalo de -1 a 1,5, aproximadamente.

De acordo com a Figura 7.4, a CCI associada ao item FI indica que os periódicos com nível de traço latente até -1,5 têm maior probabilidade de pertencerem à primeira categoria ordenada. Já os periódicos com traço latente entre -1,5 e 0, têm mais chance de pertencerem à categoria 2. Finalmente, os periódicos com traço latente acima de zero têm mais chance de estarem associados à categoria 3. Na Figura 7.4 a curva à direita indica que a informação fornecida pelo item FI cai drasticamente para os escores menores do que -2 ou maiores do que 1.

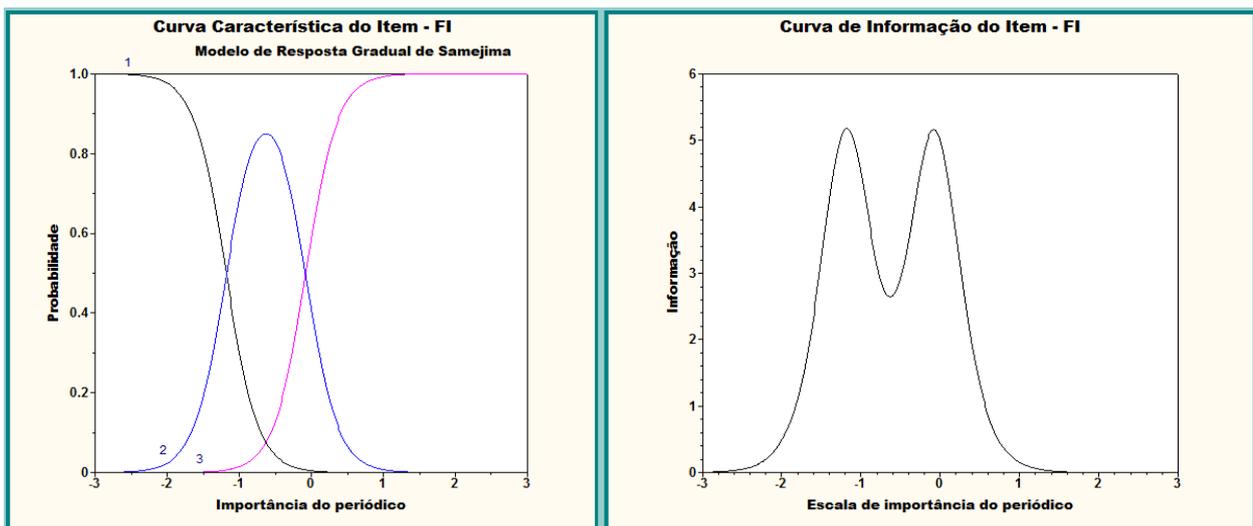


Figura 7.4: Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

Na Figura 7.8 descreve a CCI associada ao item HL, indicando que os periódicos com traço latente até -0.5 têm maior probabilidade de pertencerem à primeira categoria ordenada. Já os periódicos com traço latente entre -0.5 e 0.5 , têm mais chance de pertencerem à categoria 2. Finalmente, os periódicos com traço latente acima de 0.5 têm mais chance de estarem associados à categoria 3. Na Figura 7.8, a curva à direita indica que a informação fornecida pelo item HL é bastante modesta, descrevendo uma curva platicúrtica.

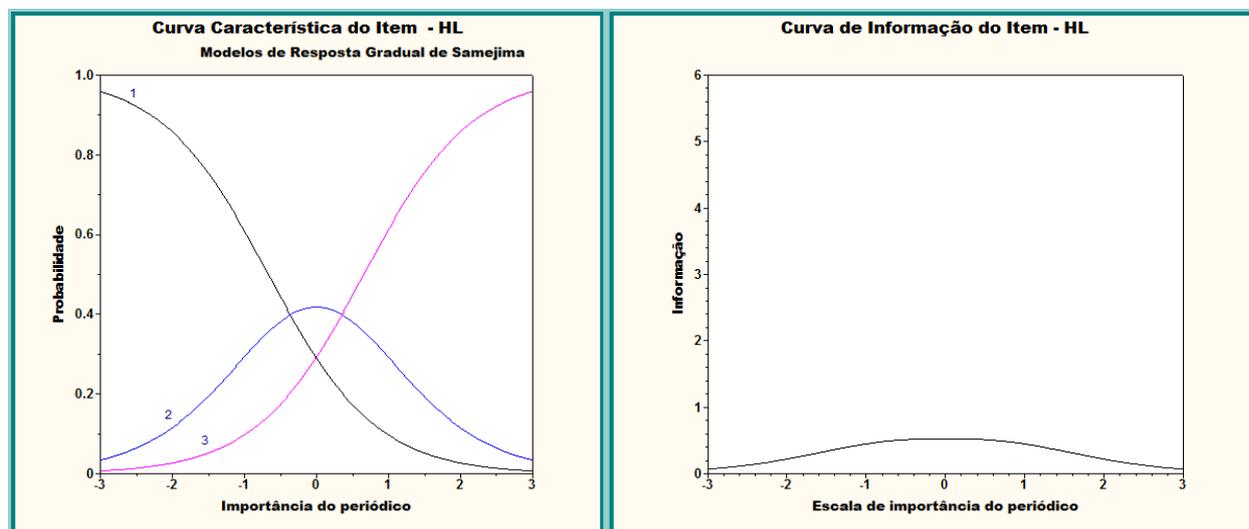


Figura 7.5: Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

A soma das funções de informação de todos os itens resulta na Função de Informação do Teste - FIT, que pode ser visualizada na Figura 7.6.

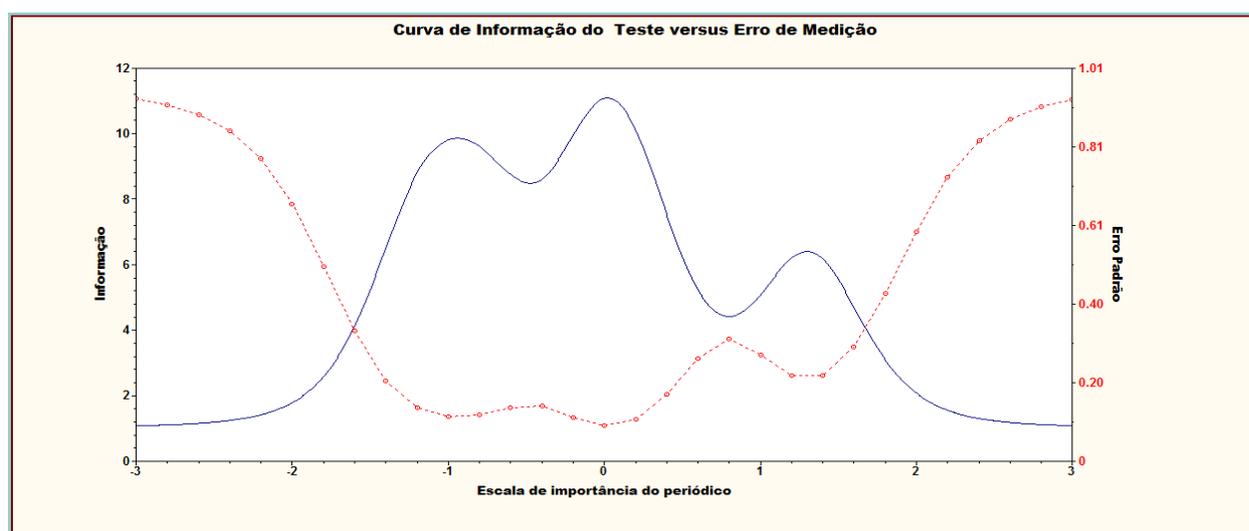


Figura 7.6: Curva de Informação do Teste (CIT) e o erro padrão de medição entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

De acordo com a Figura 7.6, a função de informação do teste apresenta maior quantidade de informação na faixa de desempenho entre -1 a 0 da escala. A linha pontilhada mostra o erro padrão, que caracteriza uma relação oposta à quantidade de informação.

De uma forma geral, observa-se através da Figura 7.6 que boa parte da cauda da distribuição (valores do traço latente menores do que -1.5 ou maiores do que 1.5) encerra grande

incerteza nos valores estimados. Tal resultado evidencia a falta de itens que sejam informativos para os periódicos cujo traço latente está situado nas referidas caudas. Claramente, o fato de, até o momento estaremos trabalhando com apenas com 3 itens não possibilita uma cobertura adequada da distribuição associada ao traço latente. No entanto, um estudo acerca do impacto da incorporação de novos indicadores a um subconjunto de periódicos será abordado no Capítulo 8 desta dissertação.

Uma observação importante a ser feita sobre os ajustes dos modelos de TRI politômicos desta dissertação, e que já foi mencionada no início da Seção 6.2, é que a tentativa de tratar os dados faltantes para os quais não foi possível fazer imputação (15 periódicos no total) como se fossem “itens não-apresentados” resulta em valores estimados de escore no centro da escala. Mais precisamente, os escores destas 15 observações foram estimados como 0.0. No entanto, estes valores estimados não estão corretos e causam distorções nas ordenações dos periódicos.

A seguir discute-se a ordenação dos periódicos com base nos escores de *valor do periódico* estimados com base no modelo apresentado nesta seção.

7.3.2 Ordenação dos periódicos

Na Tabela 7.3, apresenta-se apenas os 30 primeiros periódicos de maior traço latente entre os 1189 que compõem a base de dados. Para esta tabela considerou-se apenas os $(\theta_{i's})$ distintos, ou seja, dentre os periódicos que possuíam os mesmos escores (empates) escolheu-se, ao caso, apenas um periódico.

Tabela 7.3: *Representação tabular parcial da ordenação dos periódicos científicos.*

Ordenação	Periódico	Área	Escore (θ_i)
1	Human Heredity	Outra área	1.531
2	Applied Physics Letters	Outra área	1.320
3	Publications of the Astronomical Society	Outra área	1.256
4	The Journal of the Acoustical Society	Outra área	0.743
5	Games and Economic Behavior	Outra área	0.609
6	Computer Aided Design	Mat. Apl.	0.529
7	Information Sciences	Outra área	0.417
8	Anais da Academia Brasileira de Ciências	Matemática	0.144
9	Annalen der Physik (Leipzig)	Outra área	0.127
10	Rendiconti del Circolo Matematico di Palermo	Matemática	0.055
11	Applied Economics	Outra área	0.037
12	Applied Mathematics and Computation	Mat. Apl.	-0.027
13	International Journal of Geometric	Outra área	-0.05
14	Journal of Mathematical Economics	Mat. Apl.	-0.096
15	Ars Combinatoria	Matemática	-0.201
16	American Journal of Physics	Outra área	-0.274
17	Mathematical Physics. Analysis and Geometry	Mat. Apl.	-0.291
18	Demography (Chicago)	Outra área	-0.314
19	Indagationes Mathematicae (Print)	Matemática	-0.360
20	International Journal of Wavelets	Outra área	-0.458

Representação tabular parcial da ordenação dos periódicos científicos. Continuação da Tabela 7.3.

Ordenação	Periódico	Área	Escore (θ_i)
21	Commentationes Math. Univer. Carolinae	Matemática	-0.477
22	Optimization	Mat. Apl.	-0.750
23	Czechoslovak Mathematical Journal	Matemática	-0.755
24	Optimal Control Applications & Methods	Mat. Apl.	-0.818
25	Arquivos Brasileiros de Oftalmologia	Outra área	-0.819
26	Lecture Notes in Physics	Outra área	-0.937
27	Automation and Remote Control	Outra área	-0.940
28	Theory of Probability and its Applications	Estatística	-1.236
29	International Journal of Computer Mathematics	Mat. Apl.	-1.288
30	American Mathematical Society Notices	Matemática	-1.473

Fonte: Elaborado pelos autores (2017).

A título de informação, na Tabela 7.4 apresenta-se a frequência absoluta e relativa dos escores dos periódicos “proficiência dos periódicos” na ordenação dos periódicos científicos por área de interesse.

Tabela 7.4: *Representação tabular da frequência absoluta e relativa dos escores dos periódicos “proficiência dos periódicos” na ordenação dos periódicos científicos por área de interesse.*

Escore (θ_i)	Estatística	Matemática	Mat. Aplicada	Outra área	Total
1.531	9(0.11)	28(0.33)	15(0.18)	32(0.38)	84
1.32	2(0.08)	3(0.12)	1(0.04)	18(0.75)	24
1.256	3(0.14)	5(0.24)	3(0.14)	10(0.48)	21
0.743	19(0.10)	75(0.40)	17(0.09)	75(0.40)	186
0.609	0(0.00)	0(0.00)	0(0.00)	1(1.00)	1
0.529	11(0.10)	21(0.20)	24(0.23)	49(0.47)	105
0.417	2(0.04)	19(0.41)	4(0.09)	21(0.46)	46
0.144	3(0.12)	14(0.54)	3(0.12)	6(0.23)	26
0.127	2(0.06)	6(0.18)	11(0.33)	14(0.42)	33
0.055	11(0.11)	74(0.74)	6(0.06)	9(0.09)	100
0.037	1(0.07)	0(0.00)	6(0.43)	7(0.50)	14
-1.473	22(0.17)	86(0.65)	22(0.17)	2(0.02)	132
-1.288	0(0.00)	0(0.00)	4(0.57)	3(0.43)	7
-1.236	2(0.40)	0(0.00)	1(0.20)	2(0.40)	5
-0.94	0(0.00)	0(0.00)	0(0.00)	7(1.00)	7
-0.937	12(0.34)	1(0.03)	2(0.06)	20(0.57)	35
-0.819	9(0.14)	2(0.03)	5(0.08)	49(0.75)	65
-0.818	0(0.00)	0(0.00)	3(0.60)	2(0.40)	5
-0.755	13(0.57)	3(0.13)	2(0.09)	5(0.22)	23
-0.75	0(0.00)	0(0.00)	2(0.67)	1(0.33)	3
-0.477	1(0.01)	22(0.18)	3(0.02)	94(0.78)	120
-0.458	2(0.29)	3(0.43)	1(0.14)	1(0.14)	7
-0.36	3(0.05)	11(0.19)	10(0.17)	34(0.59)	58
-0.314	1(0.25)	0(0.00)	1(0.25)	2(0.50)	4
-0.291	0(0.00)	0(0.00)	2(1.00)	0(0.00)	2
-0.274	2(0.05)	16(0.39)	6(0.15)	17(0.41)	41

Representação tabular da frequência absoluta e relativa dos escores dos periódicos “proficiência dos periódicos” na ordenação dos periódicos científicos por área de interesse. Continuação da Tabela 7.4.

Escore (θ_i)	Estatística	Matemática	Mat. Aplicada	Outra área	Total
-0.201	0(0.00)	1(1.00)	0(0.00)	0(0.00)	1
-0.096	0(0.00)	0(0.00)	1(1.00)	0(0.00)	1
-0.05	0(0.00)	1(0.25)	2(0.50)	1(0.25)	4
-0.027	4(0.14)	18(0.62)	6(0.21)	1(0.03)	29
Total de periódicos					1189

Fonte: Elaborado pelos autores (2017).

Observando a Tabela 7.3 em que tem-se a representação tabular parcial da ordenação dos periódicos científicos e comparando-a com a Tabela 7.4 na qual enfatiza-se a representação tabular da frequência absoluta e relativa dos escores dos periódicos, isto é, a “proficiência dos periódicos” na ordenação dos periódicos científicos por área de interesse, verifica-se nesta última, por exemplo, que para $\theta = 1.531$, dos 1189 periódicos 9 (11%) são da Estatística, 28 (33%) da Matemática, 15 (18%) da Matemática Aplicada e 32 (38%) recebem classificação de Outra área.

Uma análise comparativa visual dos escores de valor dos periódicos em cada grupo é apresentada na Figura 7.7. Nestas análises são considerados os escores de valor dos 1189 periódicos da base de 2012.

As análises que seguem não são definitivas, uma vez que os dados são esparsos e sofreram processos de imputação, com consequências ainda não plenamente ou satisfatoriamente estudadas. Além disso, estamos cientes da fragilidade de um método TRI que incorpora apenas 3 itens. Deve-se considerar, no entanto, que os resultados aqui descritos são fruto de um procedimento piloto, ou seja, fruto de um protótipo, a ser aperfeiçoado em análises futuras e de posse de bases de dados mais informativas. Em resumo, os resultados aqui descritos precisam ser analisados com cautela.

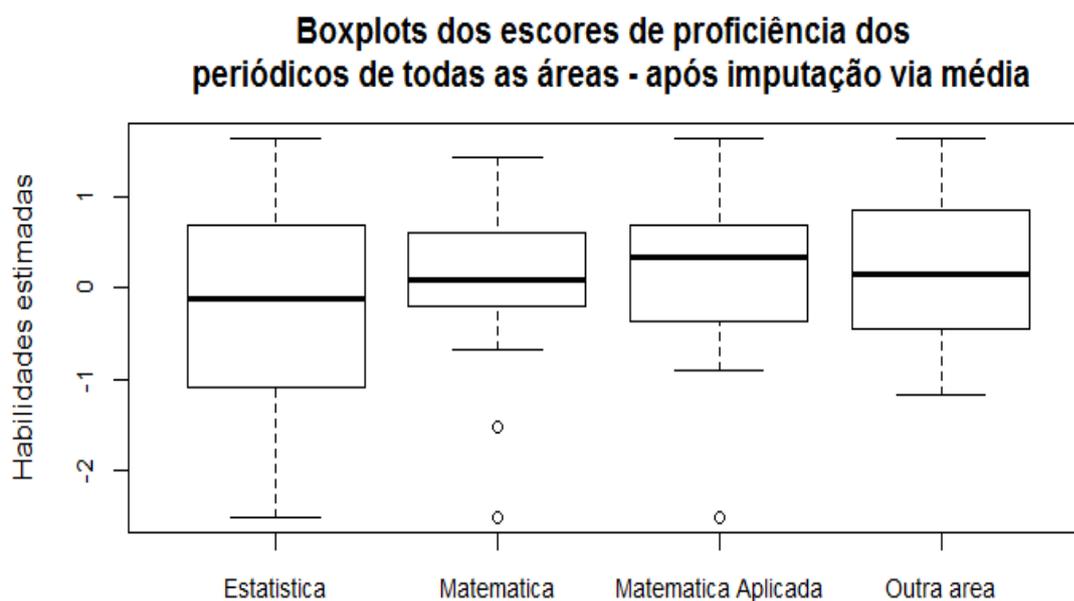


Figura 7.7: *Boxplots dos escores de proficiência dos periódicos em todas as áreas de interesse.*

De acordo com a Figura 7.7, observa-se que, de acordo com os boxplots comparativos as

medianas dos escores de valor dos periódicos por área apresentam algumas diferenças, sendo as medianas dos escores nas áreas de Matemática e em Outras áreas, aproximadamente iguais. O escore mediano dos periódicos na área de Matemática Aplicada é ligeiramente superior às demais áreas, enquanto o escore mediano na área de Estatística é ligeiramente inferior às demais áreas. A distribuição dos escores na área de Estatística possui enorme variabilidade, comparativamente às distribuições dos escores nas outras áreas. Isso pode ser devido ao pequeno número de periódicos disponíveis na área, bem como ao excesso de imputações que foram empreendidas: cerca de 24% dos dados dos 134 periódicos sofreram algum tipo de imputação, possivelmente causando um vício negativo nos escores. Estes resultados sinalizam a necessidade de estudos mais aprofundados para a imputação de dados.

7.4 Uma proposta para a classificação dos periódicos

Uma vez construída a métrica para a ordenação dos periódicos, necessita-se classificá-los segundo as normas da CAPES, nas categorias A1, A2, B1, B2, B3, B4, B5 e C. Na **Etapa 8** da Seção 7.1.2 estão descritas as duas regras básicas a serem seguidas por cada Comitê avaliador. No entanto, estas ainda não permitem a especificação dos pontos de corte necessários à separação dos periódicos nos estratos, A1 a B5. A nossa estratégia baseou-se no caso de dados históricos.

Utilizando-se a Plataforma Sucupira (sucupira.capes.gov.br), coletou-se informações sobre as classificações dos periódicos científicos na área de Estatística para os anos de 2012 (1471 periódicos), 2013 (563 periódicos), 2014 (563 periódicos) e 2015 (1126 periódicos). A Tabela 7.5 traz os valores percentuais dos periódicos classificados em cada estrato.

Tabela 7.5: Representação tabular da classificação dos periódicos científicos por ano em termos percentuais (%) segundo o método apresentado pelo Qualis/CAPES.

Estrato de classificação	Classificação (%) Qualis/CAPES				
	2012	2013	2014	2015	Média
A1	7.205982	11.90053	11.36767	9.413854	9.972011
A2	13.9361	12.25577	12.07815	11.54529	12.453828
B1	20.05439	24.51155	23.80107	19.98224	22.087309
B2	13.32427	19.36057	23.44583	16.0746	18.051316
B3	10.94494	9.413854	8.525755	11.98934	10.195721
B4	12.44052	6.216696	6.039076	9.058615	8.438726
B5	19.57852	15.63055	13.85435	8.525755	14.397294
C	2.515296	0.710480	0.88810	13.41030	4.381044

Fonte: Elaborado pelos autores (2017).

Para ter maior comparabilidade com as classificações do QUALIS 2012, uma possibilidade é utilizar os valores percentuais de 2012 dados pela Tabela 7.5. No entanto, verificou-se que tomar as proporções médias ao longo dos anos, em cada categoria, diminui a incerteza quanto aos pontos de corte mais adequados para o estabelecimento dos estratos. As análises que foram realizadas para o ano de 2012 encontram-se no **Apêndice G** desta pesquisa.

A seguir apresenta-se os pontos de corte calculados levando-se em conta a coluna mais à direita da Tabela 7.5.

Sejam p_1, p_2, \dots, p_7 as porcentagens médias, obtidas ao longo dos anos, com base na Tabela 7.5. Sejam a_1, a_2, \dots, a_7 os respectivos valores percentuais acumulados, de modo que

$$a_i = \sum_{j=1}^i p_j. \quad (7.1)$$

Defina-se como pc_i , $i = 1, \dots, 7$, os pontos dos escores ordenados em ordem decrescente, que servirão de base para a descrição dos pontos de corte. Seja

$$pc_i = \left[\frac{a_i}{100} \times N \right], \quad i = 1, \dots, 7 \quad (7.2)$$

onde $[\cdot]$ designa a parte inteira de um número e N designa o tamanho de nossa base de dados. No caso do ano de 2012, $N = 1189$.

A classificação A1-TRI incluirão todos os pc_1 periódicos de postos mais altos, isto é, $pc_1 = 118$ periódicos. Já as classificações A2-TRI incluirão todos os periódicos de postos $pc_1 + 1$ a pc_2 , isto é, todos os periódicos cujos postos, ordenados em ordem decrescente, estejam compreendidos entre 119 a 266. Procedimento análogo é seguido para os demais estratos. A Tabela 7.6 sumariza estas informações.

Tabela 7.6: *Classificação TRI nos estratos A1 a B5 considerando postos/periódicos ordenados em ordem decrescente.*

Estrato	Ponto de corte (posto em ordem decrescente)	Intervalos de postos dos escores (ordem decrescente)
A1-TRI	$pc_1 = 118$	Incluir os 118 periódicos de posts mais altos.
A2-TRI	$pc_2 = 266$	Incluir os periódicos de postos 119 a 266 na escala decrescente.
B1-TRI	$pc_3 = 529$	Incluir os periódicos de postos 267 a 529 na escala decrescente.
B2-TRI	$pc_4 = 743$	Incluir os periódicos de postos 530 a 743 na escala decrescente.
B3-TRI	$pc_5 = 865$	Incluir os periódicos de postos 744 a 865 na escala decrescente.
B4-TRI	$pc_6 = 965$	Incluir os periódicos de postos 866 a 965 na escala decrescente.
B5-TRI	$pc_7 = 1136$	Incluir os periódicos de postos 966 a 1136 na escala decrescente.
C-TRI	-	Incluir todos os registros de postos superiores a 1136.

Fonte: Elaborado pelos autores (2017).

7.5 Comparação entre as classificações CAPES e via TRI

Nesta seção apresenta-se uma comparação entre as classificações dos periódicos divulgadas pela CAPES em 2012, que foram baseados em métodos *Ad-Hoc*, e as classificações destes periódicos obtidos através dos procedimentos descritos na Seção 7.4, em que os escores de valor de periódicos foram estudados através da metodologia de TRI explanada nas seções anteriores.

Apresenta-se, na Tabela 7.7, a classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo o método de classificação via TRI proposto na Seção 7.4.

Tabela 7.7: *Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área por área - método de classificação via TRI, considerando-se a média das proporções de 2012 a 2015, $N = 1189$ periódicos, e o método proposto nesta dissertação*

Categoria	Número de periódicos por área de classificação				Total
	Estatística	Matemática	Mat. Apl.	Outra área	
A1 TRI	11	35	16	56	118
A2 TRI	16	57	11	64	148
B1 TRI	24	82	51	106	263
B2 TRI	19	112	34	49	214
B3 TRI	6	27	10	79	122
B4 TRI	19	8	7	66	100
B5 TRI	33	56	20	62	171
C TRI	6	32	14	1	53
				Total	1189

Fonte: Elaborado pelos autores (2017).

A título de comparação com os montantes obtidos pelo modelo da TRI (Tabela 7.7) com os obtidos pelo Comitê de Matemática e Estatística, apresentam-se na Tabela 7.8.

Tabela 7.8: *Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo os critérios considerados pelo Qualis/CAPES, ano base 2012.*

Categoria	Número de periódicos por área de classificação				Total
	Estatística	Matemática	Mat. Apl.	Outra área	
A1	16	44	18	19	97
A2	28	56	41	50	175
B1	26	66	45	111	248
B2	20	64	18	61	163
B3	9	40	18	60	127
B4	14	53	13	61	141
B5	20	82	10	104	216
C	1	4	0	17	22
				Total	1189

Fonte: Elaborado pelos autores (2017).

Uma breve comparação entre as Tabelas 7.7 e 7.8 evidencia que, sob a metodologia proposta nesta dissertação para a ordenação e classificação dos periódicos, os estratos superiores do Qualis passam a ser ocupados, majoritariamente, por periódicos da “Outra área”. Além disso, houve um aumento substancial de periódicos da área de Estatística classificados como B5, o que pode ser devido ao reduzidíssimo número de itens disponíveis, não permitindo que estimativas mais precisas dos escores sejam obtidas. Os métodos de imputação utilizados também podem ter contribuído para tais resultados. Adicionalmente, o processo de padronização dos dados no pré-processamento pode não ter sido ainda efetivo para remover efeitos próprios das áreas nos dados. Algumas análises iniciais na investigação deste tópico são abordadas na Seção 7.6.

Na tentativa de avaliar as concordâncias/discordâncias entre as classificações feitas nesta dissertação com as divulgadas pelo Qualis-2012, apresenta-se as Tabelas 7.9 e 7.10. No caso

de uma concordância perfeita entre as duas classificações, todos os dados não-nulos estariam concentrados na diagonal principal de tais tabelas. No entanto observa-se um elevado número de discordâncias. Devemos esclarecer que, no presente estágio desta pesquisa, não há como julgar precisamente qual o procedimento, se o Ad-Hoc utilizado pela CAPES ou o que se desenvolveu nesta dissertação, é o que produz classificações mais justas e coerentes dos periódicos, uma vez que ambos os procedimentos são falhos e “verdadeiro estado” de cada periódico é desconhecido.

Tabela 7.9: Representação tabular da classificação Qualis/CAPES versus classificação proposta pela TRI, media das proporções de 2012 a 2015, $N = 1189$ periódicos.

Classificação CAPES	Proposta de classificação via TRI							
	A1 TRI	A2 TRI	B1 TRI	B2 TRI	B3 TRI	B4 TRI	B5 TRI	C TRI
A1	72	20	5	0	0	0	0	0
A2	34	63	77	1	0	0	0	0
B1	12	55	117	51	10	3	0	0
B2	0	7	53	70	11	20	2	0
B3	0	3	7	35	43	22	17	0
B4	0	0	3	56	3	18	51	10
B5	0	0	1	1	55	37	80	42
C	0	0	0	0	0	0	21	1

Fonte: Elaborado pelos autores (2017).

Tabela 7.10: Representação tabular da distribuição percentual (%) por classificação Qualis/CAPES versus classificação proposta pela TRI, media das proporções de 2012 a 2015, $N = 1189$ periódicos.

Classificação CAPES	Proposta de classificação via TRI							
	A1 TRI	A2 TRI	B1 TRI	B2 TRI	B3 TRI	B4 TRI	B5 TRI	C TRI
A1	0.74	0.21	0.05	0.00	0.00	0.00	0.00	0.00
A2	0.19	0.36	0.44	0.01	0.00	0.00	0.00	0.00
B1	0.05	0.22	0.47	0.21	0.04	0.01	0.00	0.00
B2	0.00	0.04	0.33	0.43	0.07	0.12	0.01	0.00
B3	0.00	0.02	0.06	0.28	0.34	0.17	0.13	0.00
B4	0.00	0.00	0.02	0.40	0.02	0.13	0.36	0.07
B5	0.00	0.00	0.00	0.00	0.25	0.17	0.37	0.19
C	0.00	0.00	0.00	0.00	0.00	0.00	0.95	0.05

Fonte: Elaborado pelos autores (2017).

De acordo com a Tabela 7.10, cujas linhas representam perfis - linha das proporções de concordância entre os dois procedimentos, observa-se concordâncias de 74% nos periódicos classificados como A1, 36% nos periódicos classificados como A2, 47% nos periódicos classificados como B1 e, 43% nos periódicos classificados como B2.

A Figura 7.8 proporciona uma ilustração adicional dos dados informados na Tabela 7.10. As concordâncias são representadas, nesta figura, por bolas na cor vermelha, de tamanhos proporcionais à magnitude das concordâncias. Já as bolas na cor verde representam as dissimilaridades.

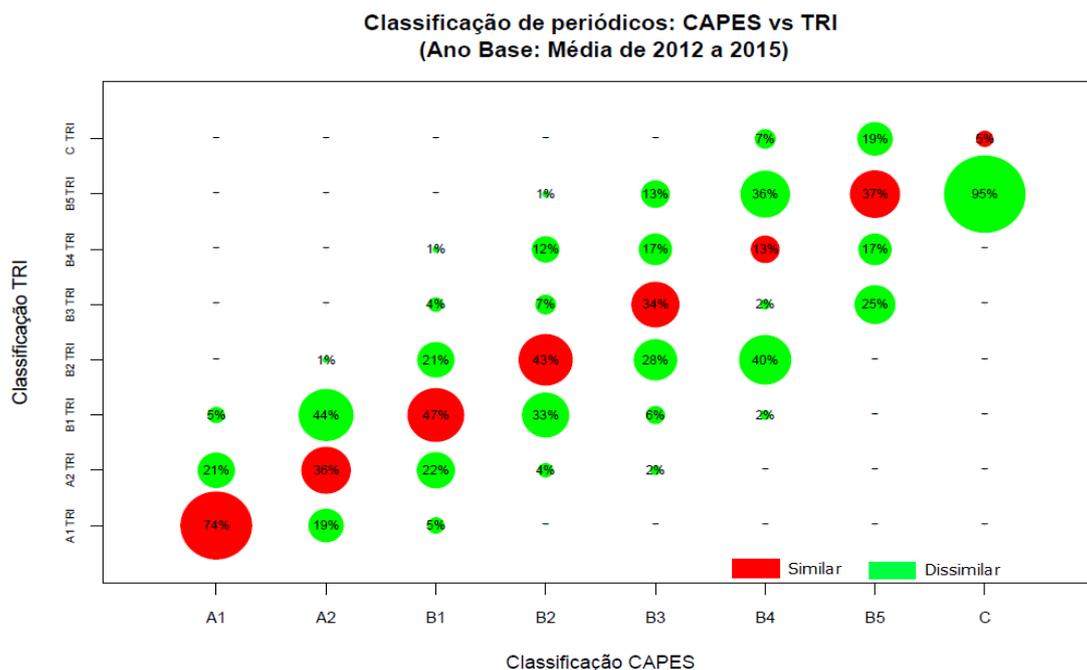


Figura 7.8: *Classificação de periódicos CAPES versus TRI, ano base média de 2012 a 2015, N = 1189 periódicos. Elaborado pelos autores (2017).*

7.6 Analisando a persistência de efeito de grupo

Na seção anterior mencionou-se a possibilidade de que a padronização dos dados dentro dos grupos não tenha sido suficiente para eliminar possíveis efeitos de grupo nos indicadores bibliométricos. Nesta seção faremos algumas análises para tentar elucidar essa questão. A hipótese de pesquisa, é a de não existência de efeito de grupo após a etapa de padronização.

Inicialmente apresenta-se as Figuras 7.9 e 7.10, que descrevem, respectivamente, os histogramas dos escores de proficiência (valor do periódico) globais, isto é, para os 1189 registros, e os histogramas do referidos escores dentro de cada área. Observa-se na Figura 7.9 a presença de certo agrupamento nos escores. Na Figura 7.10 observa-se que as distribuições dos escores em cada área são distintas entre si.

No intuito de realizar um teste de hipóteses para testar o efeito de diferença entre as distribuições dos escores nos grupos, realizou-se uma análise de DIF, Differential Item Functioning, (Embretson e Reise, 2000).

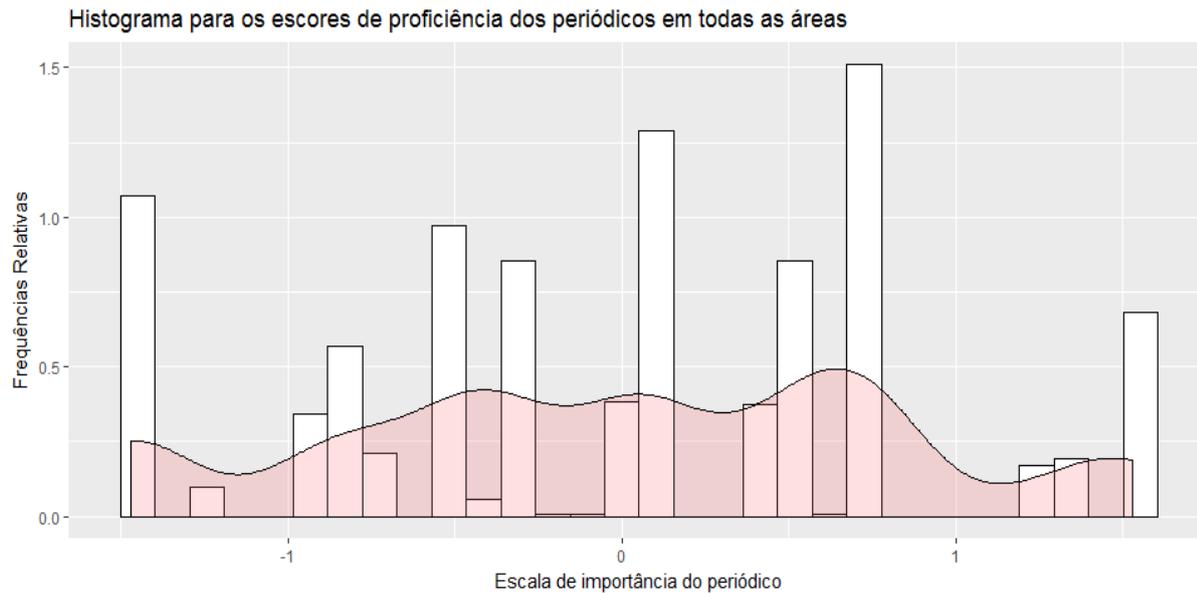


Figura 7.9: Histograma para os escores de proficiência dos periódicos em todas as áreas, $N = 1189$ periódicos. Elaborado pelos autores (2017).

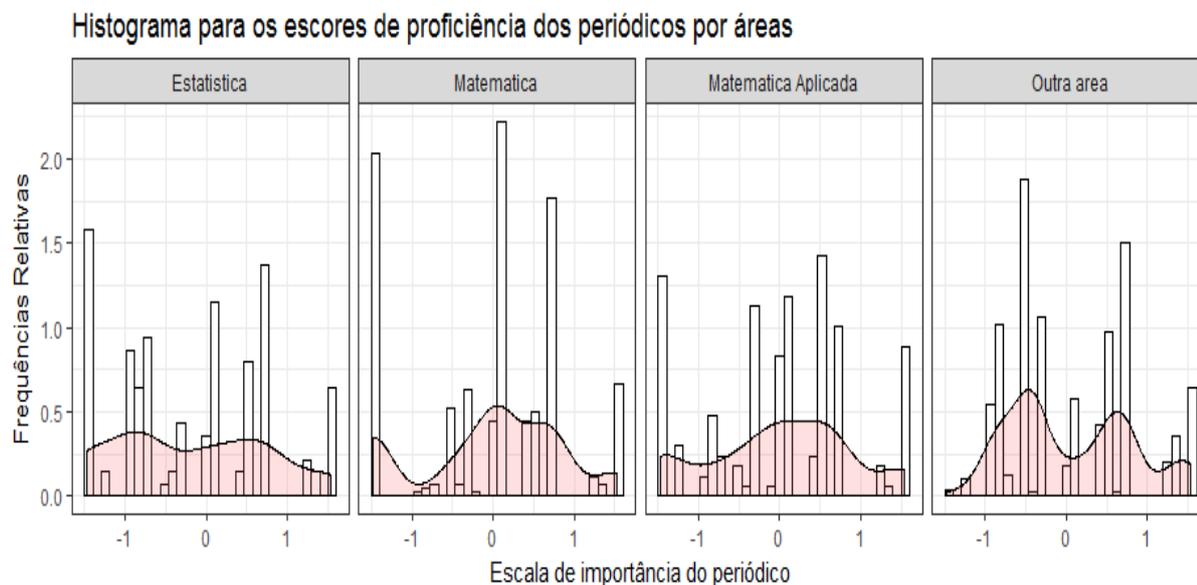


Figura 7.10: Histograma para os escores de proficiência dos periódicos por áreas, $N = 1189$ periódicos. Elaborado pelos autores (2017).

Segundo Embretson e Reise (2000), Cap. 10, o efeito de DIF ocorre quando os itens de um teste não possuem o mesmo relacionamento com a variável latente em estudo ao longo de dois ou mais grupos. Em termos concretos, um dado item apresenta DIF se sua CCI difere nos diferentes grupos, ou, equivalentemente, quando qualquer parâmetro de um item difere entre os grupos. Thissen *et al.* (1993) esclarece que o DIF trata-se de uma interação entre os parâmetros de um item e os grupos.

Neste trabalho os grupos são descritos pelas áreas: Matemática pura, Matemática Aplicada, Estatística e Outras áreas. Adotaremos a área de Estatística como o *grupo de referência* e as demais áreas como *grupos focais*.

As análises de DIF foram realizadas com o auxílio do *software* Parscale. Os programas (syntaxes) utilizados estão documentados no **Apêndice I**.

Inicialmente tentou-se utilizar nestas análises, o modelo de Resposta Gradual de Samejima (1969). No entanto, possivelmente devido ao excessivo número de parâmetros do modelo (agora incluindo efeito de grupo) aliado a um conjunto de dados muito esparsos, não obteve-se convergência. Optou-se, então, pelo uso do Modelo de Crédito Parcial Generalizado (Muraki, 1992) que prevê parâmetros de discriminação “ a_j ” iguais, em cada grupo, para um dado item. Na próxima seção descreve-se como o teste de DIF foi realizado.

7.7 Testando o efeito de DIF

Nesta seção, realiza-se uma análise de DIF para testar o efeito de grupo entre os escores dos periódicos. Utilizou-se o Modelo de Crédito Parcial Generalizado (Muraki, 1992), ora considerando a existência de grupos, ora desconsiderando os grupos. Os *softwares* utilizados para a análise foram o Parscale e o R.

As etapas desta análise estão descritas a seguir.

Etapa 1 Hipóteses:

$$\begin{cases} H_0: & \text{inexistência de DIF entre os parâmetros de posição (b}'_j\text{'s) nos diferentes itens e grupos;} \\ H_1: & \text{pelo menos um dos itens apresenta DIF.} \end{cases}$$

Os resultados da aplicação do modelo de DIF em conjunto com o modelo politômico de Crédito Parcial Generalizado estão resumidos na Tabela 7.11. Mais detalhes sobre o teste podem ser encontrados no **Apêndice I** da pesquisa.

Etapa 1.1

Analisar os dados sem considerar grupos, isto é, como se todos os respondentes tivessem vindo de uma única população. Anotar o último valor de - LOG LIKELIHOOD obtido na FASE 2 (definido no comando INPUT do *software*). Denotaremos este valor por LL .

Etapa 1.2

Analisar os dados considerando grupos separados usando o modelo de DIF. Anotar o último valor de - LOG LIKELIHOOD obtido na FASE 2. Denotaremos este valor por LL_g .

Etapa 1.3

Calcular a estatística Qui-quadrado

$$\chi^2 = \text{abs}(LL_g - LL), \quad (7.3)$$

que, sob H_0 , segue distribuição $\chi^2_{(n-1)(g-1)}$ g.l., onde n = número de itens, g = número de grupos e abs =valor absoluto.

Quando (7.3) é significativo, há evidências de que os efeitos diferenciais estão presentes.

Etapa 2

 Resultados:

Seguindo os passos descritos na **Etapa 1**, obteve-se $LL = 6855.310$ e $LL_g = 6778.134$. Logo, tem-se

$$\begin{aligned}\chi^2 &= 6855.310 - 6778.134 \\ &= 77.176.\end{aligned}$$

O valor- p do teste χ^2 , considerando 6 g.l, é obtido como a seguir:

```
1 p.value<-pchisq(77.157,6, ncp = 0, lower.tail = FALSE, log.p = FALSE)
2 p.value
3 1.379546e-14
```

A Tabela 7.11 mostra as estatísticas do teste da razão de verossimilhança marginal global para testar presença de DIF nos dados para o modelo com dados politômicos:

Tabela 7.11: *Estatística Qui-quadrado (χ^2): teste do efeito de grupos entre o escore dos periódicos.*

Estimativas	χ^2	Grau de liberdade (g.l.)	valor- p
Valores coeficientes	77.176	6	1.379546e-14

Fonte: Elaborado pelos autores (2017).

Conforme mostra a Tabela 7.11, o teste é altamente significativo, indicando que há diferença entre os grupos, mesmo após o uso da padronização dos indicadores em cada área (Estatística, Matemática Aplicada, Matemática e Outra Área).

A origem e a unidade da escala são definidas de modo que a média e o desvio padrão do grupo de referência (baseline) sejam 0 e 1, respectivamente. Um gráfico das distribuições latentes estimadas é dado abaixo.

Como uma análise final sobre a existência do DIF nos dados trabalhados apresenta-se um gráfico comparativo das distribuições dos escores de proficiência englobando todas as áreas e curvas separadas para cada uma das áreas. Para mais detalhes consulte o **Apêndice I**.

Na presença de DIF as distribuições latentes não representam a mesma variável latente e não é possível uma comparação significativa das duas distribuições de acordo com a Figura 7.11 as distribuições dos escores nos diferentes grupos apresentam diferenças apreciáveis.

Representação gráfica das distribuições latentes estimadas

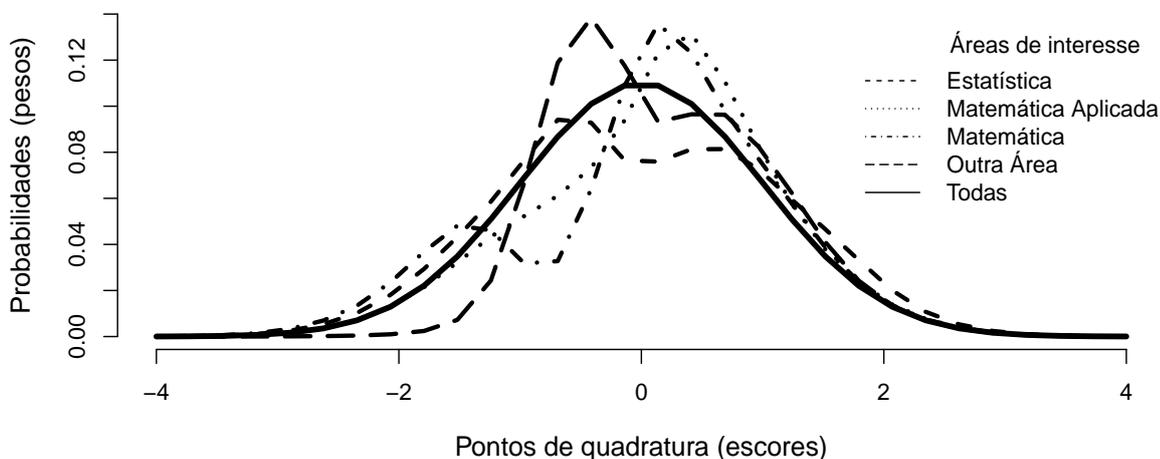


Figura 7.11: *Representação gráfica das duas distribuições latentes estimadas, isto é, com grupos e sem grupos. Elaborado pelos autores (2017).*

7.8 Elaboração da escala de medida

Contrariamente aos processos realizados anteriormente, nas análises que seguem utilizou-se um subconjunto do banco de dados da CAPES do ano-base de 2012 considerando-se apenas os periódicos que continham informação em pelo menos um dos itens. Dessa forma, a base de dados é formada por $N=882$ periódicos nessa condição. Dentre os 882 referidos periódicos em que há alguma informação faltante, tratou-se o item como “não apresentado”. As informações sobre os itens foram previamente padronizadas dentro de cada um dos grupos (Estatística, Matemática, Matemática Aplicada e Outras área), na tentativa de remover “efeitos de área de concentração”.

Os valores dos parâmetros dos itens e os escores de proficiência dos periódicos foram estimados em uma mesma métrica, sendo representados na escala $(0;1)$, ou seja, média 0 e desvio padrão 1,0. Considerando-se o Modelo de Samejima (1969), $D = 1$ e os valores estimados dos parâmetros do modelo, calculou-se a distribuição acumulada relativa a cada categoria dos itens (vide expressão (3.10)).

Na Tabela 7.12 apresenta-se os resultados da estimação dos parâmetros de cada item.

Tabela 7.12: Estimativas dos parâmetros dos itens AIS, FI e HL da base de dados sem imputação, CAPES/Qualis ano-base 2012, $N = 882$ periódicos. EP (erro padrão da estimativa).

Parâmetros	Item		
	AIS	FI	HL
a_i (EP)	4.71 (0.29)	1.59 (0.11)	0.67 (0.09)
$b_{i,1}$ (EP)	-1.09 (0.05)	-1.40 (0.12)	-2.73 (0.39)
$b_{i,2}$ (EP)	-0.54 (0.04)	-0.63 (0.08)	-1.26 (0.22)
$b_{i,3}$ (EP)	-0.11 (0.03)	0.02 (0.07)	-0.14 (0.15)
$(b_{i,4})$ (EP)	0.30 (0.03)	0.74 (0.09)	1.08 (0.20)

Fonte: Elaborado pelos autores (2017).

Os itens com os maiores valores estimados dos parâmetros $b_{i,k}$ são relacionados ao FI e ao HL, o que indica que os itens sobre o Fator de Impacto e Meia Vida dos periódicos são os mais “difíceis” de serem observados nas categorias mais altas e que, portanto, só serão atingidas por periódicos de maior sucesso ou relevância. Já o AIS é o item que apresenta os maiores valores estimados dos parâmetros $b_{i,k}$ nas categorias mais baixas. Tal resultado pode indicar uma característica importante do AIS na discriminação de periódicos numa região da escala latente em que nem o FI e nem o HL são capazes de discriminar.

De acordo com a Figura 7.12, a CCI associada ao AIS indica que os periódicos com nível de traço latente menores do que -1 possuem uma maior probabilidade de pertencerem à primeira categoria ordenada. Já os periódicos com habilidade entre -1 e -0.5 aproximadamente, possuem maior probabilidade de estarem situados na categoria 2. Periódicos com habilidade entre -0.5 e 0.0, possuem maior probabilidade de estarem situados na categoria 3. Periódicos com habilidade entre 0.0 e 0.5, possuem maior probabilidade de estarem situados na categoria 4 e, finalmente, periódicos com habilidade superior a 0.5 apresentam maior probabilidade de estarem situados na categoria 5. De acordo com o lado direito da Figura 1, a Curva de Informação do Item (CII) (vide Seção 3.5) atinge o seu valor máximo para os traços latentes no intervalo de -1 a 0.5, aproximadamente.

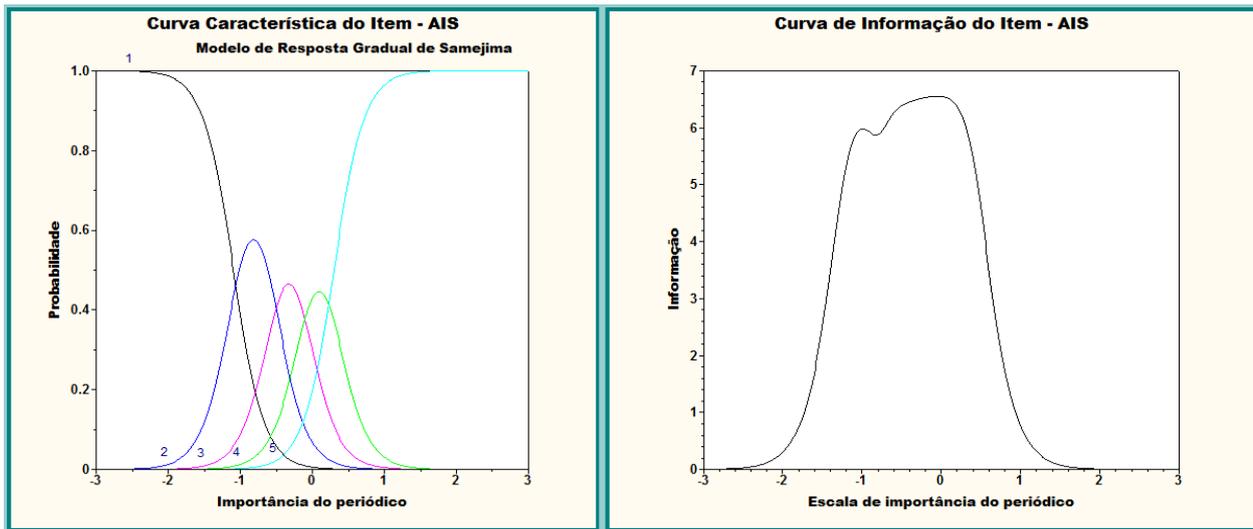


Figura 7.12: Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

Com base na Figura 7.13, a CCI associada ao FI indica que os periódicos com nível de traço latente menores do que -1 possuem uma maior probabilidade de pertencerem à primeira categoria ordenada. Já os periódicos com habilidade entre -1 e -0.5 aproximadamente, possuem maior probabilidade de estarem situados na categoria 2. Periódicos com habilidade entre -0.5 e 0.0, possuem maior probabilidade de estarem situados na categoria 3. Periódicos com habilidade entre 0.0 e 0.5, possuem maior probabilidade de estarem situados na categoria 4 e, finalmente, periódicos com habilidade superior a 0.5 apresentam maior probabilidade de estarem situados na categoria 5. Na Figura 2 a curva de informação item FI cai drasticamente para os escores menores do que -1 ou maiores do que 1.

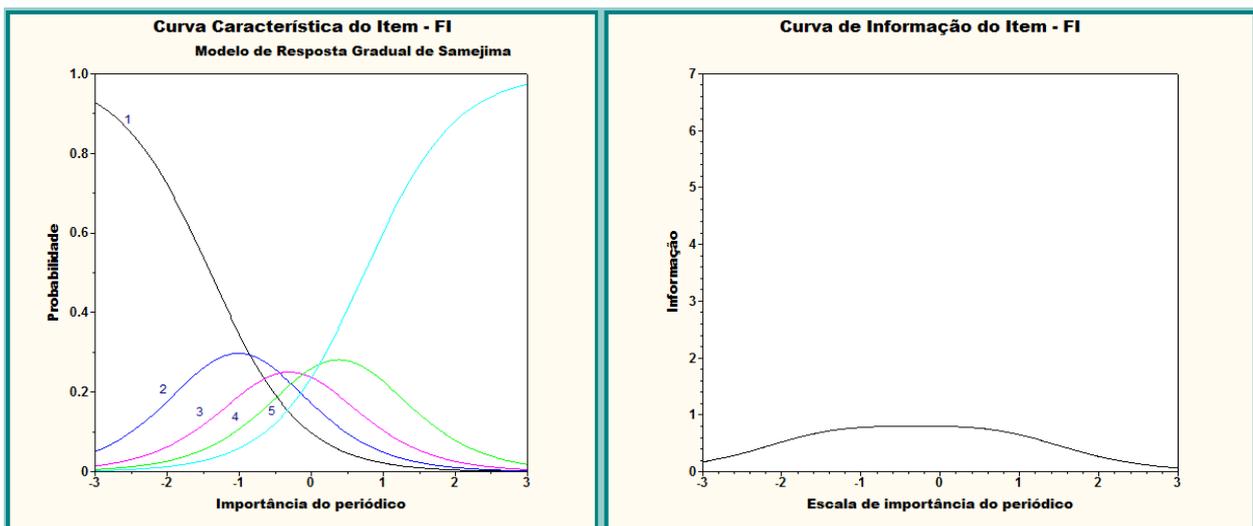


Figura 7.13: Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

A Figura 7.14 descreve a CCI associada ao item HL, indicando que os periódicos com traço latente até -1.0 têm maior probabilidade de pertencerem à categoria 1. Já os periódicos com traço latente maiores ou iguais a -1.0 têm mais chance de pertencerem à categoria 5.

Na Figura 3, a curva de informação fornecida pelo item HL é platicúrtica.

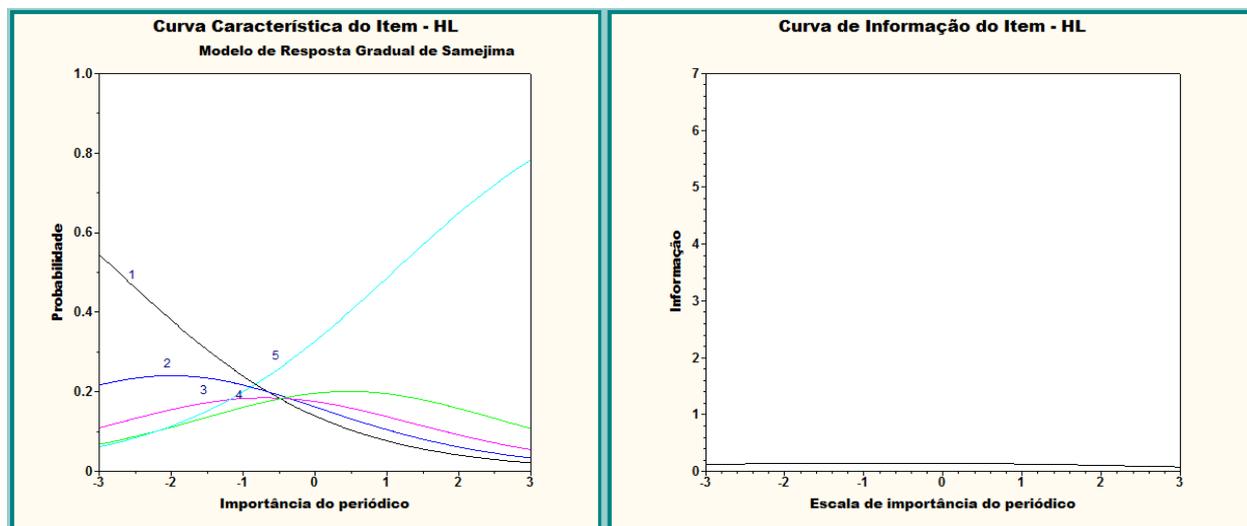


Figura 7.14: Curva Característica do Item (CCI) e a Função de Informação do Item (FII) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

A soma das funções de informação de todos os itens resulta na Função de Informação do Teste - FIT, que pode ser visualizada na Figura 7.15.

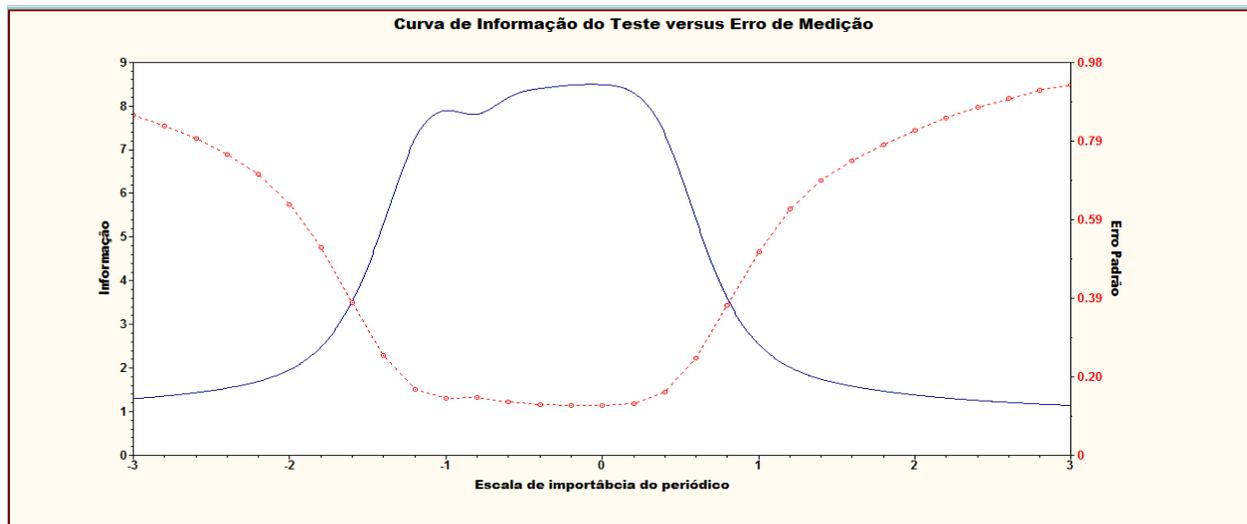


Figura 7.15: Curva de Informação do Teste (CIT) e o erro padrão de medição entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

Conforme pode ser visualizado na Figura 7.15, a função de informação do teste ou informação total do conjunto de itens possui maior quantidade de informação no intervalo de -1.0 a 0.5 da escala. A linha pontilhada mostra o erro padrão das estimativas dos escores, configurando uma relação oposta à quantidade de informação. O erro padrão mais baixo situa-se também neste intervalo. Assim, para quantidades maiores de informações, menor o erro padrão.

De uma forma geral, observa-se através da Figura 7.15 que boa parte da cauda da distribuição (valores do traço latentes menores do que -1.5 ou maiores do que 1.5) encerra grande incerteza nos valores estimados. Tal resultado evidencia a falta de itens que sejam

informativos para os periódicos cujo traço latente está situado nas referidas caudas. Claramente, o fato de, até o momento estarmos trabalhando com apenas 3 itens, não possibilita uma cobertura adequada da distribuição associada ao traço latente. No entanto, um estudo acerca do impacto da incorporação de novos indicadores a um subconjunto de periódicos será abordado no Capítulo 8 desta dissertação.

7.8.1 Interpretação da escala em ordem decrescente

Os valores estimados dos parâmetros do Modelo de Samejima e as probabilidades acumuladas em cada categoria também são apresentados na Tabela 7.13. A referida tabela mostra, também, a identificação dos níveis âncora ou quase-âncora (vide Seção 5.2), que estão destacados em vermelho. Algumas pequenas adaptações às condições ou requisitos necessários para a descrição dos níveis âncora foram necessárias. Dessa forma, o requisito 1 (níveis âncora apresentam probabilidade em determinado nível âncora da escala superior a 0.65), o requisito 2 (probabilidade menor que 0.5 no nível de habilidade anterior) e o requisito 3 (diferença maior que 0.3 entre estes níveis) foram ligeiramente alterados de modo a propiciar melhor interpretação da escala de habilidades.

De acordo com a Tabela 7.13 observa-se quanto ao item, categorias e níveis âncora que tanto para o AIS quanto para o FI, classifica-se como nível âncora na categoria 2 o escore -1.0; como nível âncora na categoria 3 o escore -0.3; como nível âncora na categoria 4 o escore 0.3, e como nível âncora na categoria 5 o escore 1.1. As mesmas especificações de níveis âncora seguem para o HL, exceto para a categoria 5, em que o escore do nível âncora associado é bem mais alto, atingindo o valor de 3.0.

As interpretações da escala de valores latentes são dadas a seguir

1. Escores iguais ou superiores a 3.0 indicam periódicos na categoria mais elevada de HL.
2. Escores iguais ou superiores a 1.1 indicam periódicos na categoria mais elevada de AIS e FI.
3. Escores iguais ou superiores a 0.3 indicam periódicos na categoria 4 ou superior de AIS, FI e HL.
4. Escores iguais ou superiores a -0.3 indicam periódicos na categoria 3 ou superior de AIS, FI e HL.
5. Escores iguais ou superiores a -1.0 indicam periódicos na categoria 2 ou superior de AIS, FI e HL.

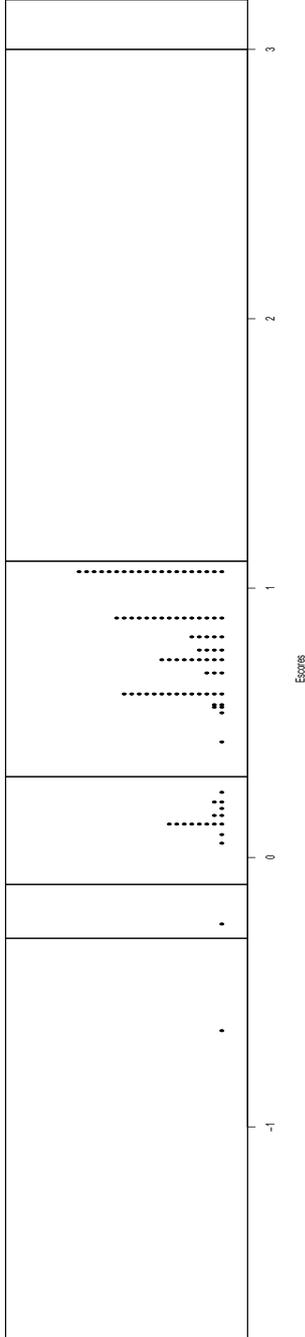
É interessante apresentar o posicionamento dos $N = 882$ periódicos utilizados nesta análise, na escala das habilidades, indicando o estrato atribuído pela CAPES-2012 a cada um destes. As Figuras 7.16 e 7.17 apresentam uma série de 7 gráficos, um para cada estrato, A1 a B5, descrevendo tais posicionamentos. Observa-se que há grande intersecção na faixa de escores dos periódicos classificados como A1, A2 e B1 pela CAPES. Tal resultado revela que ainda há um longo caminho a ser trilhado até que se possa, de fato, classificar os periódicos de maneira satisfatória.

Tabela 7.13: Escala (0,1) dos escores e definição dos níveis âncora e quase-âncora associados às categorias dos itens em estudo.

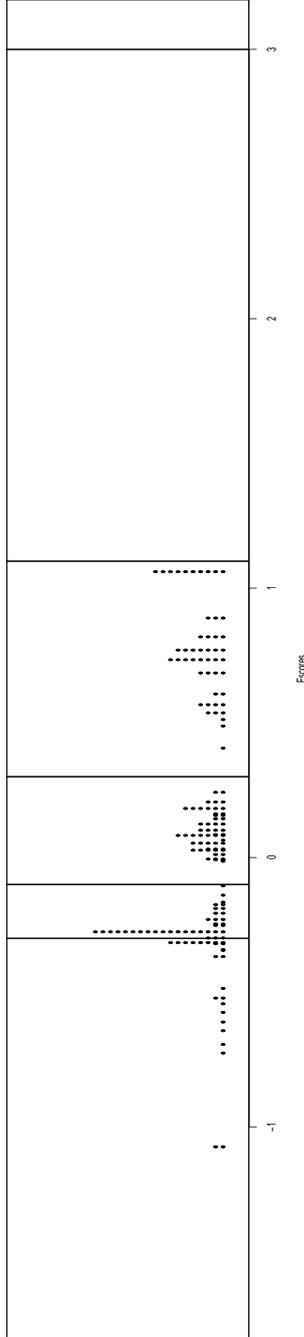
Item	a	Categoria	b	-3,7	-1	-0,3	0,3	1,1	3
AIS	4,70803	2 ou +	-1,09344	0,000000	0,608240	0,976696	0,998587	0,999967	1
		3 ou +	-0,53558	0,000000	0,100969	0,751965	0,980809	0,999548	1
		4 ou +	-0,10784	0,000000	0,014769	0,288087	0,872152	0,996620	1
		5 ou +	0,29969	0,000000	0,002196	0,056075	0,500365	0,977421	0,999997
FI	1,58921	2 ou +	-1,39947	0,025184	0,653586	0,851611	0,937075	0,981516	0,999081
		3 ou +	-0,62734	0,007517	0,356122	0,627197	0,813625	0,939636	0,996873
		4 ou +	0,01560	0,002719	0,166034	0,377175	0,611108	0,848558	0,991361
		5 ou +	0,74247	0,000858	0,059014	0,160205	0,331108	0,638345	0,973082
HL	0,66695	2 ou +	-2,73161	0,343921	0,760404	0,835038	0,883081	0,927942	0,978599
		3 ou +	-1,26169	0,164347	0,543523	0,655070	0,739153	0,828514	0,944921
		4 ou +	-0,14076	0,085190	0,360529	0,473474	0,572967	0,695831	0,890390
		5 ou +	1,08111	0,039591	0,199728	0,284727	0,372627	0,503150	0,782416

Fonte: Elaborado pelos autores (2017).

Escores dos periódicos classificados como A1-CAPES-2012 e níveis âncora associados aos itens: AIS e FI (-0,1,-0,3,0,3,1,1), e HI (-0,1,-0,3,0,3,3)



Escores dos periódicos classificados como A2-CAPES-2012 e níveis âncora associados aos itens: AIS e FI (-0,1,-0,3,0,3,1,1), e HI (-0,1,-0,3,0,3,3)



Escores dos periódicos classificados como B1-CAPES-2012 e níveis âncora associados aos itens: AIS e FI (-0,1,-0,3,0,3,1,1), e HI (-0,1,-0,3,0,3,3)



Figura 7.16: Posicionamento dos periódicos A1, A2 e B1 divulgados pela CAPES-2012 na escala de habilidades construída com base nos 882 periódicos. As linhas verticais referem-se às posições dos níveis âncora. Elaborado pelos autores (2017).

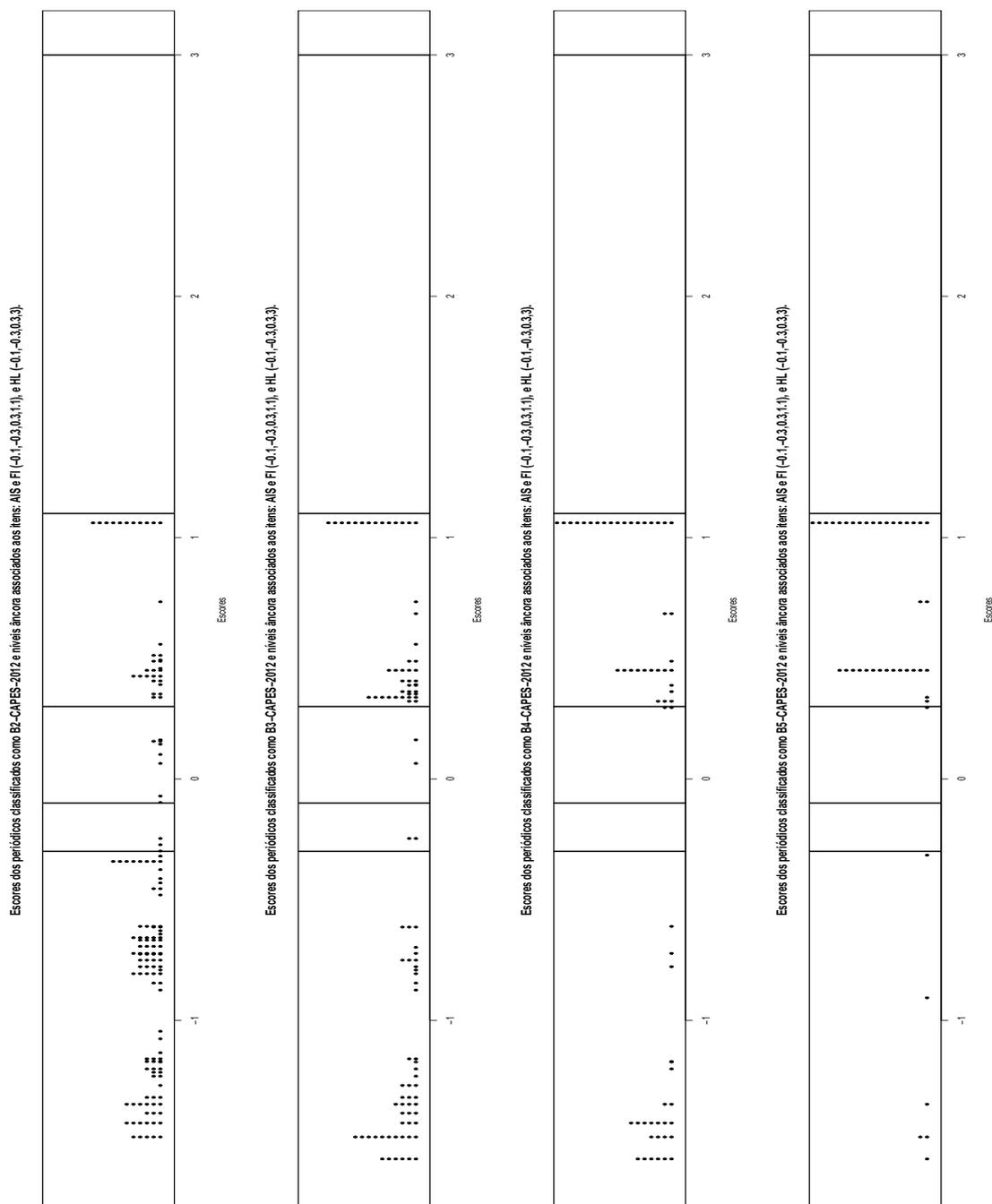


Figura 7.17: Posicionamento dos periódicos B2, B3, B4 e B5 divulgados pela CAPES-2012 na escala de habilidades construída com base nos 882 periódicos. As linhas verticais referem-se às posições dos níveis âncora. Elaborado pelos autores (2017).

Capítulo 8

Análises via TRI para uma base de dados com mais indicadores

O Scimago (<http://www.scimagojr.com/>) e a Web Of Science (<http://wokinfo.com/>), entre outros, são sites que disponibilizam indicadores, utilizando diferentes métricas, para a classificação de periódicos nas mais diversas áreas. Tais métricas são importantes na avaliação da qualidade de publicações científicas.

Como as referidas duas bases de dados disponibilizam distintos indicadores da qualidade dos periódicos, onde cada um destes indicadores exploram algum aspecto distinto utilizado para a ordenação dos jornais, objetiva-se propor um novo indicador da qualidade de periódicos, baseado em uma análise conjunta dos referidos indicadores, via Teoria da Resposta ao Item (TRI), nas diferentes bases. Para tanto, necessita-se criar uma base de dados comum que congregue todos os indicadores disponíveis. Desta forma, faz-se necessário criar mecanismos de pareamento dos registros das informações de um mesmo periódico nas diferentes bases. No entanto, tal tarefa é bastante desafiadora devido a inúmeras inconsistências das bases e a limitações dos softwares disponíveis. De forma que o uso do *software* R na elaboração dos códigos computacionais necessários revelou-se uma ferramenta extremamente eficaz para esse tipo de problema.

A partir das informações disponíveis nos sites citados, técnicas de pareamentos determinísticos foram aplicadas no intuito de compor uma base consistente, utilizando-se como chave de pareamento entre as bases, o código identificador de periódicos, ISSN (pareamento exato).

Neste capítulo a base de dados considerada é resultado de um processo de pareamento ou de cruzamento de informações envolvendo várias bases de dados. Vale ressaltar que uma vez que os diversos tipos de indicadores bibliométricos não estão descritos em uma base de dados comum, mas sim em 3 bancos de dados distintos, a saber: Qualis (<http://qualis.capes.gov.br/webqualis>), Web Of Science (<http://wokinfo.com/>) e Scimago (<http://www.scimagojr.com/>), faz-se necessário construir um banco de dados que contemple, para um mesmo periódico, a diversidade de indicadores bibliométricos existentes.

Considerando os anos de 2010 ou 2012 a base de final, isto é, a base produto dos pareamentos, é composta por 28 variáveis, no entanto, após a realização de algumas análises descritivas considerou-se apenas os indicadores (itens) listados abaixo.

No **Apêndice E** apresenta-se descritivamente as bases de dados supra citadas. Estas análises descritivas foram realizadas considerando apenas algumas das variáveis que formam o banco de dados. Por fim, neste capítulo, considerou-se apenas as variáveis que serão trabalhadas na TRI.

No que concerne as variáveis que serão trabalhadas na TRI, apresenta-se, a saber:

- FISC: fator de impacto sem auto citações (Web Of Science);
- FIJ: fator de impacto (Web Of Science);
- FI5A: fator de impacto 5 anos (Web Of Science);
- MPFI: média percentual do fator de impacto (Web Of Science);
- ES: eigenfactor score (Web Of Science);
- AIS: Article Influence Score (Web Of Science);
- CHL: cited half life (Web Of Science);
- II: immediacy index (Web Of Science);
- REF/DOC: número médio de referências por documentos em um dado ano (Scimago);
- CITES/DOC: média de citações por documento no período de 2 anos (Scimago);
- IH: índice H (Scimago);
- FIS: fator de impacto (Scimago).

Neste capítulo, primeiramente, será apresentada a análise descritiva dos dados (antes do processo de imputação). Em seguida, são descritos os resultados da análise de TRI. Ressalta-se que a ordenação e a classificação dos periódicos, não fazem parte dos objetivos deste capítulo. Objetiva-se, tão somente, avaliar o impacto da inclusão de novos indicadores na curva de informação do teste.

8.1 Base de dados resultante dos pareamentos

Algumas informações genéricas sobre os bancos de dados são apresentadas na Tabela 8.1.

Tabela 8.1: *Informações sobre os periódicos das bases de dados: Classificação de Periódicos no Qualis/CAPES, Web Of Science e Scimago Journal & Country Rank, ano base 2012.*

Base de dados	Banco de dados	Total de observações
1	Qualis	1189
2	Scimago	1032
3	Web Of Science	504

Fonte: Elaborado pelos autores (2017).

Observando a Tabela 8.1, verifica-se que a base de dados do Qualis/CAPES, em comparação as demais bases, é a base que apresenta mais informações e, quando comparada com a base de dados da Web Of Science o número de registro é o dobro.

Nos pareamentos que foram realizados entre as três bases (Qualis *vs.* Scimago *vs.* Web Of Science), levou-se em conta o ISSN como chave comum entre as bases de dados.

Verificou-se que apenas 369 registros são comuns às três bases de dados. Deste montante, apenas 342 periódicos têm informação sobre área ou grupo comuns nos anos de 2010 e 2012. Além disso, constatou-se que 51 registros não apresentam informações sobre a classificação (estrato ao qual o periódico está classificado) advinda da CAPES. Resultando assim, uma base final composta de 291 registros.

Portanto, para as análises que seguem considerou-se a base formada por 291 periódicos. Ressalta-se que, deste montante, 280 registros possuem informações para todos os indicadores, restando apenas 11 ($291 - 280 = 11$) registros que apresentam pelo menos uma informação faltante em seus indicadores.

Apresenta-se a seguir algumas análises descritivas para a bases de dados formada pelos 291 registros, sem considerar o processo de imputação.

8.1.1 Análise descritiva dos indicadores

Nas figuras que se seguem, tem-se a representação gráfica dos indicadores bibliométricos da base de dados formada pelos 291 registros antes do processo de imputação.

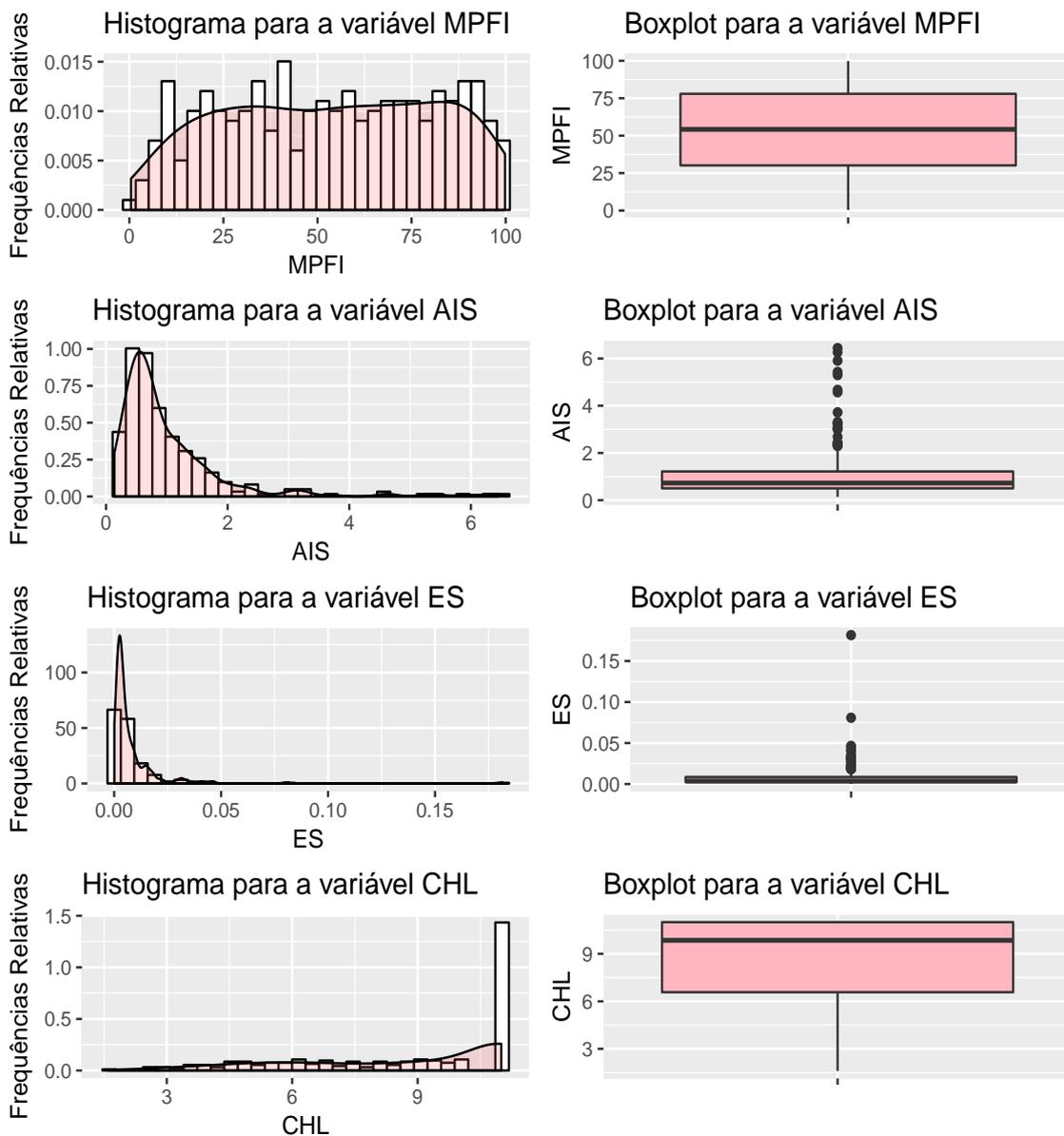


Figura 8.1: Representações gráficas para os indicadores bibliométricos: *MPFI*, *AIS*, *ES* e *CHL* antes do processo de imputação, $N = 291$ periódicos.

Analisando os histogramas verifica-se que os indicadores Fator AIS e ES apresentam comportamento assimétrico à direita, os boxplots confirmam tais evidências. Já o Fator de Meia Vida (*CHL*) descreve um comportamento assimétrico à esquerda e uma gande

concentração de informações abaixo da mediana como observa-se no boxplot referente a este indicador. Visualizando o histograma do indicador MPFI verifica-se um comportamento uniforme nos dados.

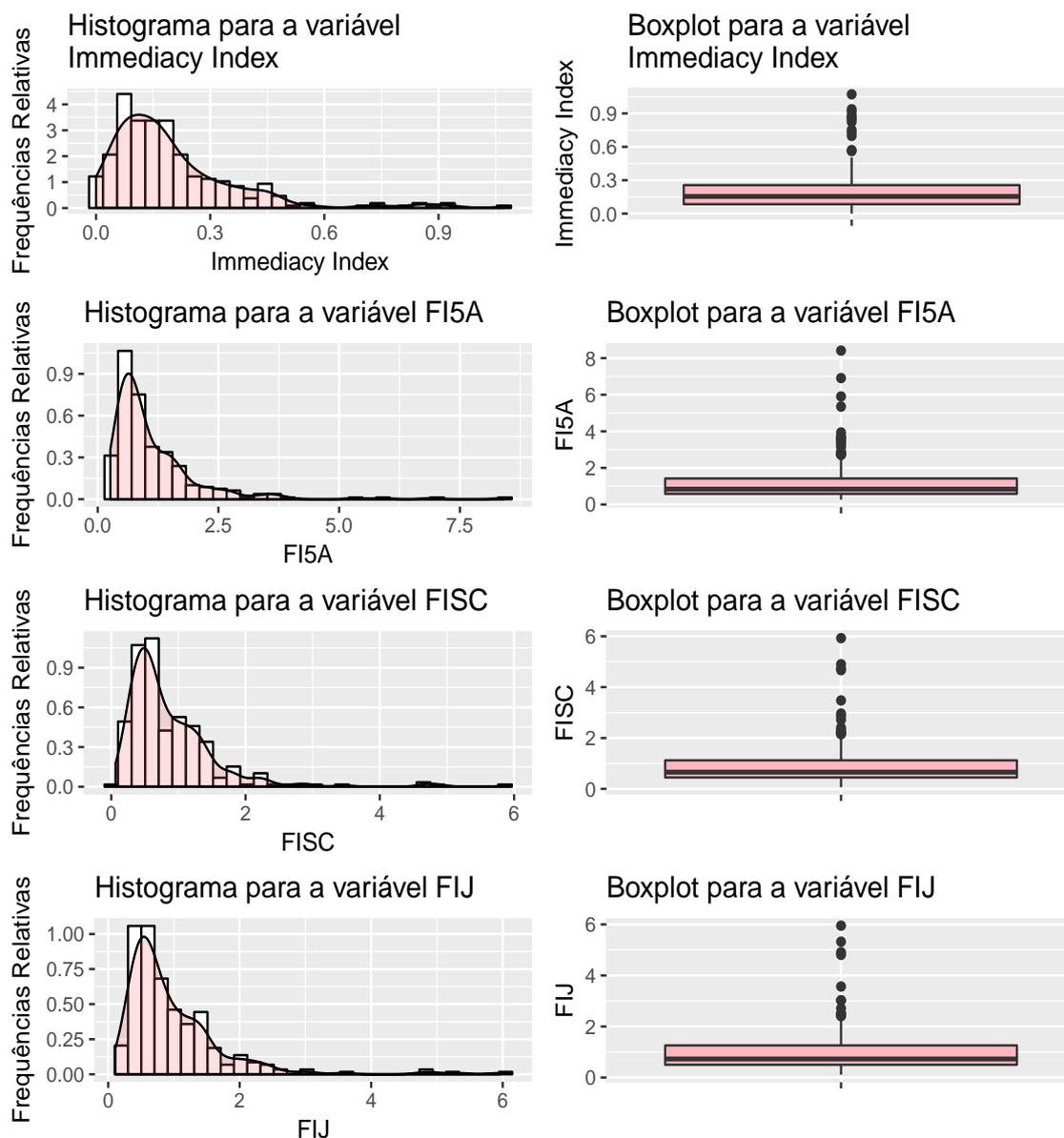


Figura 8.2: Representações gráficas para os indicadores bibliométricos: II, FI5A, FISC e FIJ antes do processo de imputação, $N = 291$ periódicos.

Visualizando a Figura 8.2, observa-se-se que os histogramas dos indicadores II, FI5A, FISC e FIJ apresentam comportamento assimétrico à direita. Os boxplots confirmam tais evidências.

Objetivando verificar se existe associação entre estes indicadores, na sequência apresenta-se algumas análises de Scatter Plot.

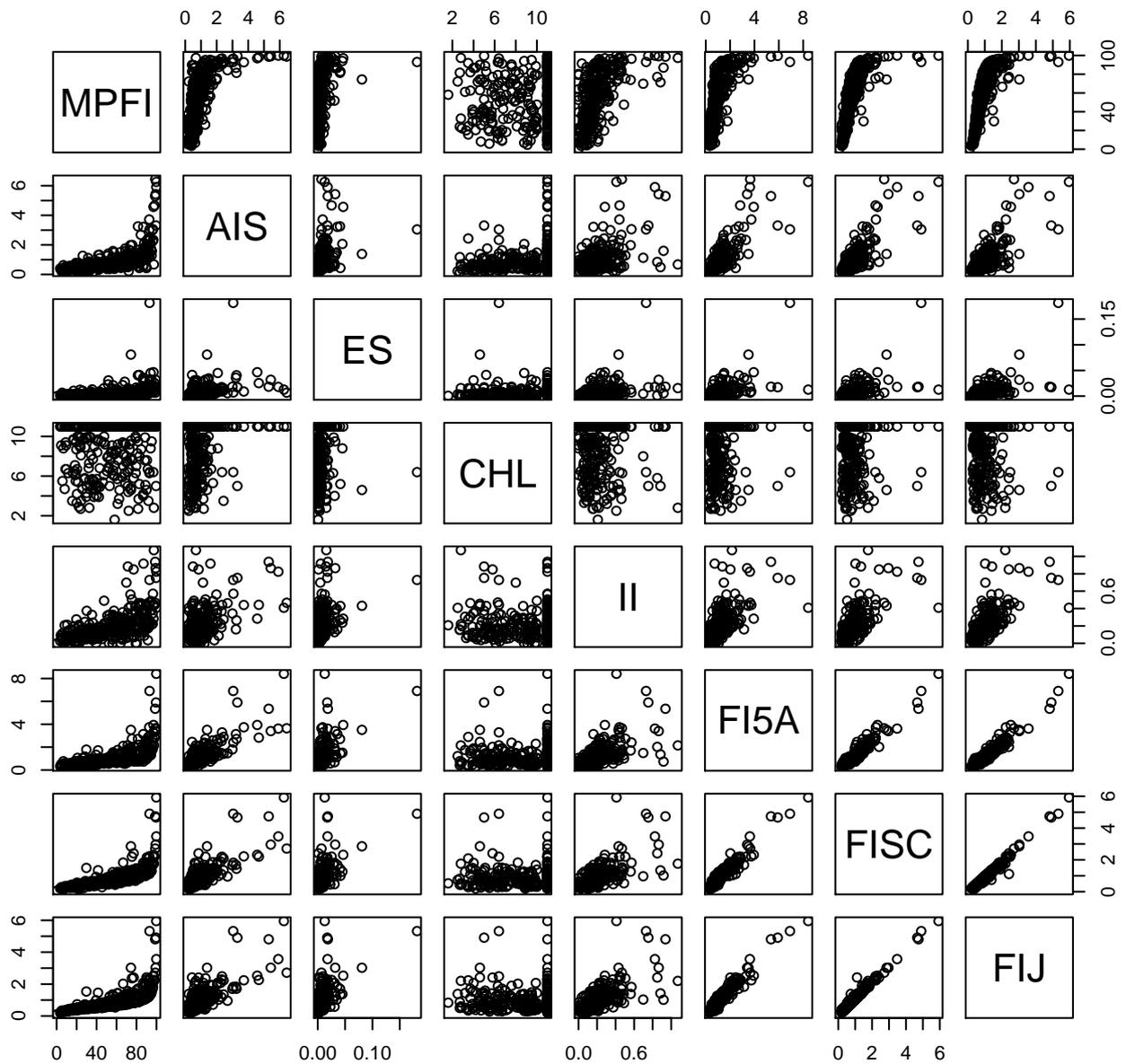


Figura 8.3: Scatter Plot para os indicadores bibliométricos observados: MPFI, AIS, ES, CHL, II, FI5A, FISC e FIJ antes o processo de imputação, dados na escala original, $N = 291$ periódicos.

Na Figura 8.3, referente aos scatter plots para os indicadores bibliométricos MPFI, AIS, ES, CHL, II, FI5A, FISC e FIJ (*antes* do processo de imputação). Observa-se que há correlação entre alguns indicadores, em especial entre àqueles indicadores relativos aos fatores de impacto: FISC e FIJ.

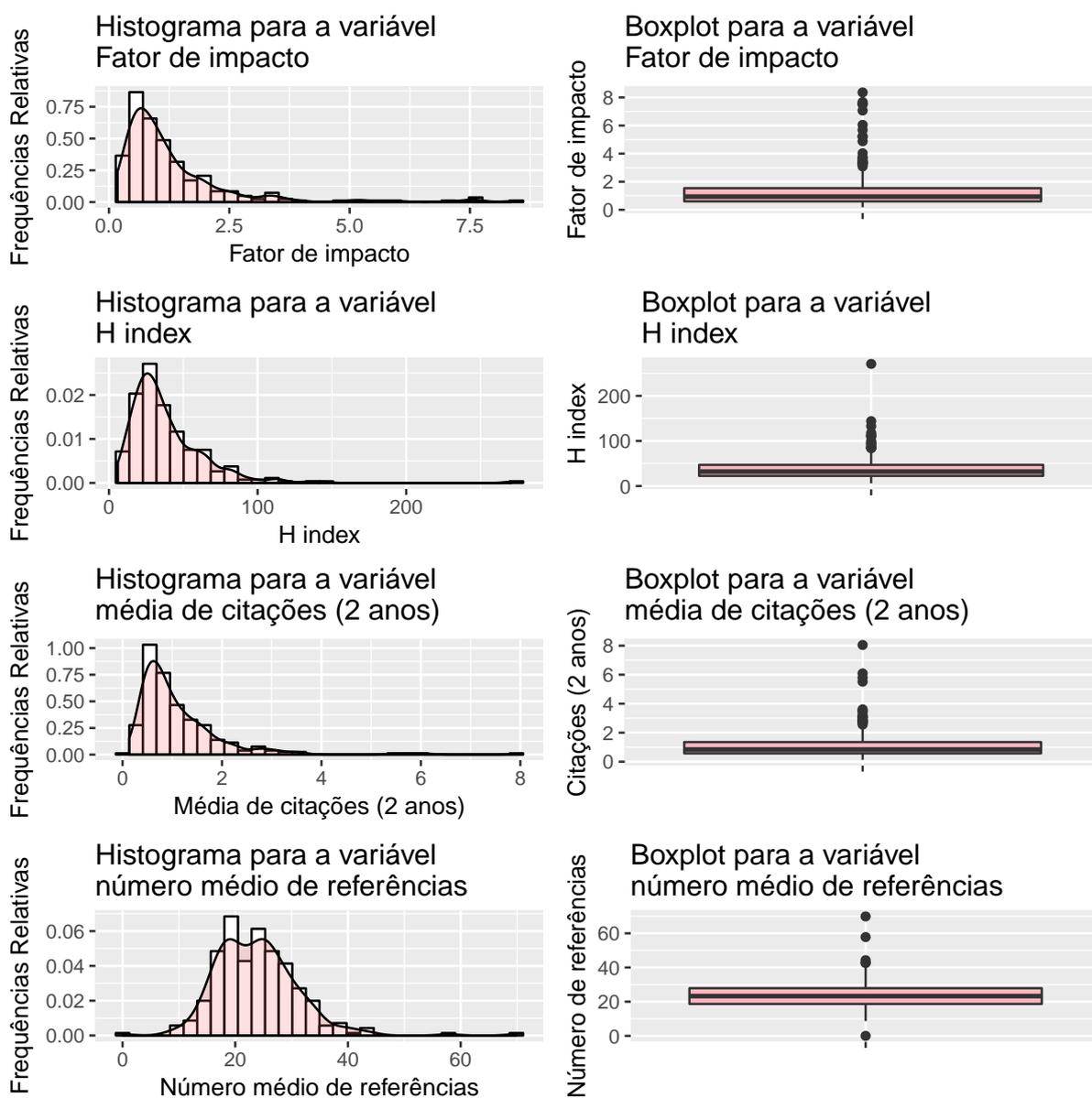


Figura 8.4: Representações gráficas para os indicadores bibliométricos: FIS, IH, REF/DOC e CITES/DOC antes do processo de imputação, $N = 291$ periódicos.

Visualizando a Figura 8.4, observa-se-se que os histogramas dos indicadores FIS, IH, REF/DOC e CITES/DOC apresentam, aparentemente, comportamento assimétrico à direita, os boxplots confirmam tais evidências. Já o histograma do indicador fator de impacto (FIS) evidencia comportamento simétrico em torno do número médio de citações recebidas pelo jornal quanto a importância e o prestígio dos jornais dos quais as citações são originárias.

Na tentativa de verificar se existe associação entre estes indicadores, na sequência apresenta-se algumas análises de Scatter Plot.

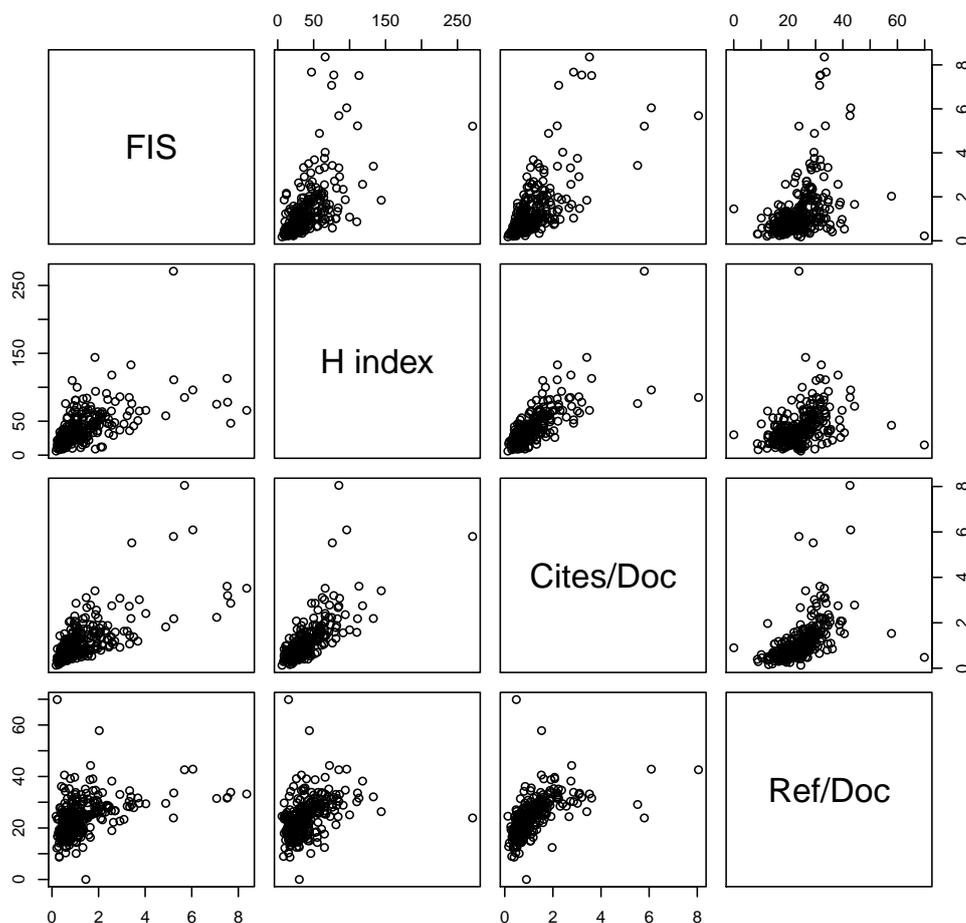


Figura 8.5: *Scatter Plot para os indicadores bibliométricos observados: FIS, IH, REF/DOC e CITES/DOC antes o processo de imputação, dados na escala original, $N = 291$ periódicos.*

Conforme pode ser visualizado na Figura 8.5, identifica-se que os indicadores FIS, IH, REF/DOC e CITES/DOC são fortemente linearmente relacionados.

Baseando-se nas análises discutidas, enfatizando-se, principalmente, o fato dos indicadores serem bastante correlacionados, tem-se que proceder da seguinte maneira: para cada uma das áreas (separadamente) tomar-se os escores padronizados para cada um dos indicadores. Desta feita, os escores separados por área passam a ser comparáveis, isto é, agora estão descritos na mesma escala, ou escala universal.

8.2 Descrição da base $N = 291$

Em relação às áreas de interesse desta pesquisa, ou seja, as áreas nos quais os periódicos estão classificados, a Tabela 8.2 apresenta a distribuição de frequência dos periódicos antes do processo de imputação, formando assim a base de dados composta por $N = 291$ periódicos.

Tabela 8.2: Representação tabular dos periódicos científicos por área de interesse, base de dados antes do processo de imputação, $N = 291$ periódicos.

Área de interesse	Frequências
Matemática	156 (54%)
Matemática Aplicada	76 (26%)
Estatística	48 (16%)
Outra área	11 (4%)

Fonte: Elaborado pelos autores (2017).

Baseado na Tabela 8.2, verifica-se que o número de periódicos classificados em Matemática representa à uma parte considerável dos dados, contendo 54% do total dos registros, ao passo que Outra área é a que comporta o menor percentual dos registros, representando pouco mais de 4% do total.

Na Tabela 8.3 apresenta-se algumas estatísticas descritivas para a base de dados, $N = 291$ periódicos, após o processo de imputação via média.

Tabela 8.3: Análise descritiva dos indicadores bibliométricos após o processo de imputação via média, $N = 291$ periódicos.

Indicador	Core	Medidas resumo						
		Min.	1st. Qu.	Mediana	Média	3rd. Qu.	Max.	NA's
MPFI	Est.	0.427	22.220	37.180	46.040	70.510	99.540	-
	Mat.	3.547	31.330	51.690	52.670	78.970	99.830	-
	Mat. Apl.	6.795	47.440	66.000	61.460	81.280	99.800	-
	Out. Área	5.877	25.270	39.480	39.710	41.880	93.030	-
AIS	Est.	0.292	0.574	0.782	1.297	5.300	5.3000	1
	Mat.	0.1770	0.4440	0.6805	0.9798	1.1840	6.4430	-
	Mat. Apl.	0.1370	0.5052	0.7420	0.9022	1.1040	6.2650	-
	Out. Área	0.141	0.269	0.608	0.778	0.749	3.047	-
ES	Est.	0.000090	0.002095	0.003610	0.007644	0.009395	0.046550	-
	Mat.	0.000390	0.002335	0.003475	0.006654	0.007645	0.044980	-
	Mat. Apl.	0.000690	0.002720	0.004600	0.007469	0.008835	0.041320	-
	Out. Área	0.000610	0.001335	0.001810	0.025330	0.003045	0.181600	-
CHL	Est.	3.300	7.650	9.700	8.847	11.000	11.000	1
	Mat.	2.700	7.250	11.000	9.188	11.000	11.000	1
	Mat. Apl.	1.600	6.300	8.650	8.166	11.000	11.000	-
	Out. Área	3.900	4.600	5.500	6.118	6.550	11.000	-
II	Est.	0.0000	0.0755	0.1490	0.1999	0.2802	0.9380	-
	Mat.	0.0000	0.0710	0.1305	0.1792	0.2152	1.0720	2
	Mat. Apl.	0.0000	0.1398	0.2085	0.2453	0.3060	0.8840	-
	Out. Área	0.0000	0.1005	0.1310	0.2029	0.2020	0.7290	-
FI5A	Est.	0.4260	0.6410	0.9220	1.4140	1.4650	5.9070	1
	Mat.	0.2620	0.4802	0.6755	0.8825	0.9735	3.6500	-
	Mat. Apl.	0.4170	0.7928	1.2040	1.3920	1.6690	8.4140	-
	Out. Área	0.3820	0.5145	0.7480	1.6890	1.7520	6.9110	-

Análise descritiva dos indicadores bibliométricos antes da imputação (dados originais), N = 291 periódicos. Continuação da Tabela 8.3.

Indicador	Core	Medidas resumo						
		Min.	1st. Qu.	Mediana	Média	3rd. Qu.	Max.	NA's
FISC	Est.	0.0650	0.4782	0.6510	1.0100	1.2510	4.7410	-
	Mat.	0.1930	0.3850	0.5570	0.7226	0.8878	3.4780	-
	Mat. Apl.	0.2830	0.6533	1.0340	1.0830	1.3330	5.9290	-
	Out. Área	0.240	0.465	0.706	1.298	1.420	4.903	-
FIJ	Est.	0.1090	0.5305	0.7230	1.0850	1.2960	4.9100	-
	Mat.	0.2060	0.4550	0.6010	0.7956	0.9612	3.5670	-
	Mat. Apl.	0.3140	0.7255	1.1380	1.2140	1.4570	5.9520	-
	Out. Área	0.2810	0.4955	0.7540	1.4080	1.4980	5.3230	-
REF/DOC	Est.	14.61	22.12	26.10	26.14	29.69	44.28	-
	Mat.	0.00	17.29	19.95	21.32	25.08	57.83	-
	Mat. Apl.	12.41	22.82	25.85	26.50	30.27	42.61	-
	Out. Área	14.87	19.00	23.90	29.64	37.66	69.88	-
CITES/DOC	Est.	0.1300	0.6475	0.8100	1.2720	1.5250	6.0900	-
	Mat.	0.1600	0.5000	0.7000	0.8582	0.9850	3.5200	-
	Mat. Apl.	0.3300	0.9475	1.3150	1.4680	1.6930	8.0500	-
	Out. Área	0.4500	0.7000	1.0300	1.6480	1.8000	5.8000	-
IH	Est.	6.00	25.00	34.50	41.71	53.50	133.00	-
	Mat.	8.00	21.00	27.00	32.04	40.00	110.00	-
	Mat. Apl.	14.00	31.75	43.00	48.36	64.25	118.00	-
	Out. Área	15.00	22.00	24.00	56.91	32.50	271.00	-
FIS	Est.	0.1780	0.5585	0.9800	1.5300	1.9520	7.5160	-
	Mat.	0.2080	0.5878	0.9045	1.3290	1.5300	8.3600	-
	Mat. Apl.	0.2320	0.6942	1.0110	1.1720	1.3620	5.6900	-
	Out. Área	0.2220	0.4265	0.5350	1.0870	0.9140	5.2140	-

Fonte: Elaborado pelos autores (2017).

Visualizando a Tabela 8.3, verifica-se que após o processo de imputação, os indicadores AIS, CHL, II e FI5A ainda apresentam NAs (missings) em suas informações.

A título de informação, enfatiza-se que a base imputada apresenta alguns missings porque o processo de imputação tinha como referência a combinação das categorias *versus* áreas e por este motivo, haviam células deste cruzamento que não possuíam informações de maneira a permitir imputação para todos os elementos.

Analogamente ao apresentado na Seção 6.4 desta pesquisa, no processo de discretização dos dados para base formada por 291 registros considerou-se 8 categorias (escala discretizada considerando 8 categorias, cada uma representando 12.5% da distribuição). Feito isso, uma categoria adicional foi criada para representar dados faltantes. Dessa forma, criou-se 9 categorias para representar os dados discretizados, sendo a categoria 1 a mais baixa das categorias ordenadas. Os dados foram previamente padronizados dentro de cada área ou grupo.

8.3 Estimação dos parâmetros TRI

Nesta seção apresenta-se os resultados preliminares da estimação dos parâmetros da TRI politômica aplicada ao conjunto de 10 indicadores bibliométricos, FI5A, MPFI, ES, AIS, CHL, II, REF/DOC, CITES/DOC, IH e FIS, dos 291 periódicos após as etapas de pré-processamento (vide Capítulo 6).

A base formada por $N = 291$ periódico contém 14 variáveis, no entanto, apenas 12 destas são úteis para este trabalho. Vale ressaltar que dessas 12 variáveis tem-se que uma variável representa o ISSN, uma outra variável representa a área ou grupo, e duas outras variáveis, a saber: FISC (Impactor Factor without Journal) e FIJ (Journal Impact Factor) por apresentarem alta correlação e baixo poder de discriminação foram desconsideradas nas análises que seguem. Por fim, as 10 variáveis, restantes, caracterizam os indicadores bibliométricos que compõem a base considerada.

Inicialmente ajustou-se um modelo politômico de Resposta Gradual de Samejima (1969) considerando-se as 9 categorias (vide Seção 6.4). Após uma sucessão extensa de análises, obteve-se que o modelo mais parsimonioso e que proporciona um melhor ajuste (melhor descrição dos dados), incorpora um número bem menor de categorias para cada um dos itens (FI5A, MPFI, ES, AIS, CHL, II, REF/DOC, CITES/DOC, IH e FIS).

O modelo inicial era composto de um parâmetro de discriminação para cada um dos itens (a_i , $i = 1, 2, \dots, 10$) e 8 parâmetros de dificuldade para cada item, ($b_{i,j}$ $j = 1, \dots, 8$). Utilizou-se o *software* Multilog para as análises.

Com junções de categorias adjacentes para os itens, necessárias à obtenção de um ajuste adequado, o modelo final é composto de três categorias ordenadas para o item CHL, de modo que há dois parâmetros de dificuldade para este item ($b_{8,1}, b_{8,2}$), cinco categorias ordenadas para o item FI5A, implicando em quatro parâmetros de dificuldade para o item FI5A ($b_{7,1}, b_{7,2}, b_{7,3}, b_{7,4}$), e 8 categorias ordenadas para os demais itens, implicando em oito parâmetros de dificuldade.

Na Tabela 8.4 estão descritas as novas categorias estabelecidas para cada um dos itens ao combinarmos categorias adjacentes, considerando-se as categorias preliminares c_1 a c_9 .

Tabela 8.4: *Categorias estabelecidas para cada um dos itens ao combinarmos categorias adjacentes considerando-se as categorias preliminares C_1 a C_9 .*

Item	Novas categorias								
	Cat. 1	Cat. 2	Cat. 3	Cat. 4	Cat. 5	Cat. 6	Cat. 7	Cat. 8	Cat. 9
MPFI	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8 e C_9	-
CITES/DOC	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8 e C_9	-
FIS	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8 e C_9	-
AIS	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8 e C_9	-
IH	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8 e C_9	-
ES	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8 e C_9	-
II	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8 e C_9	-
REF/DOC	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8 e C_9	-
FI5A	C_1 e C_2	C_3 e C_4	C_5 e C_6	C_7 e C_8	C_9	-	-	-	-
CHL	C_1, C_2 e C_3	C_4, C_5 e C_6	C_7, C_8 e C_9	-	-	-	-	-	-

Fonte: Elaborado pelos autores (2017).

Na Tabela 8.5 apresenta-se os resultados da estimação dos parâmetros de cada item

Tabela 8.5: *Estimativas dos parâmetros dos itens FI5A, MPFI, ES, AIS, CHL, II, REF/DOC, CITES/DOC, IH e FIS da base de dados imputada, produto dos pareamentos ano-base 2012, $N = 291$ periódicos. EP (erro padrão da estimativa).*

Item	Parâmetro de discriminação a_i (EP)	Parâmetros de Dificuldade						
		$b_{i,1}$ (EP)	$b_{i,2}$ (EP)	$b_{i,3}$ (EP)	$b_{i,4}$ (EP)	$b_{i,5}$ (EP)	$b_{i,6}$ (EP)	$b_{i,7}$ (EP)
MPFI	4.62 (0.46)	-1.27 (0.09)	-0.78 (0.07)	-0.41 (0.07)	-0.07 (0.06)	0.25 (0.06)	0.65 (0.08)	1.20 (0.09)
FIS	3.94 (0.37)	-1.28 (0.10)	-0.83 (0.07)	-0.47 (0.07)	-0.11 (0.07)	0.26 (0.07)	0.67 (0.09)	1.22 (0.10)
CITES/DOC	3.90 (0.38)	-1.35 (0.10)	-0.78 (0.07)	-0.42 (0.07)	-0.08 (0.07)	0.30 (0.09)	0.70 (0.08)	1.22 (0.10)
AIS	2.92 (0.25)	-1.35 (0.11)	-0.88 (0.09)	-0.48 (0.08)	-0.13 (0.08)	0.23 (0.09)	0.68 (0.12)	1.29 (0.13)
IH	2.16 (0.22)	-1.46 (0.15)	-0.96 (0.12)	-0.53 (0.10)	-0.11 (0.10)	0.29 (0.12)	0.80 (0.14)	1.47 (0.17)
ES	2.06 (0.23)	-1.51 (0.17)	-0.94 (0.12)	-0.51 (0.11)	-0.07 (0.11)	0.36 (0.12)	0.84 (0.13)	1.48 (0.16)
II	1.64 (0.20)	-1.69 (0.21)	-1.05 (0.15)	-0.56 (0.13)	-0.10 (0.12)	0.36 (0.12)	0.90 (0.15)	1.60 (0.21)
REF/DOC	1.61 (0.21)	-1.71 (0.21)	-1.09 (0.14)	-0.60 (0.14)	-0.11 (0.13)	0.37 (0.14)	0.93 (0.18)	1.65 (0.22)
FI5A	0.57 (0.15)	-0.20 (0.35)	1.65 (0.44)	1.80 (0.53)	6.56 (4.84)	-	-	-
CHL	0.28 (0.16)	-1.95 (1.27)	5.93 (3.38)	-	-	-	-	-

Fonte: Elaborado pelos autores (2017).

Considerando a Tabela 8.5, verifica-se que os itens com maior poder de discriminar os periódicos são o MPFI, FIS e CITES/DOC. Esses itens, como pode ser visualizado na

Figura 7.2, são os que possuem as Curvas Características do Item (CCI) mais acentuadas (Vide Seção 3.2). Dessa forma, estes itens trazem informações importantes no cômputo do *valor* de um periódico. Os itens AIS, IH, ES, II, e REF/DOC também proporcionaram contribuição importante na descrição do valor do periódico. Já os itens FI5A e CHL são os que apresentam os menores valores do parâmetro de discriminação.

Como mencionado anteriormente, as Curvas Características do Item (CCI) (Vide Seção 3.2) associadas ao ajuste em discussão estão apresentadas nas figuras que seguem.

De acordo com a Figura 8.6, a CCI associada ao MPFI indica que os periódicos com nível de traço latente até -1 possuem uma maior probabilidade de pertencer às categorias 1, 2, 3 e 4 ordenadas. Já os periódicos com habilidade entre -1 e 0.5 aproximadamente, possuem maior probabilidade de estarem situados nas categorias 5 e 6. Jornais com desempenho entre 0.5 e 1.5, aproximadamente, apresentam maior probabilidade de estar situada na categoria 7. Já os jornais com desempenho superior a 1.5, apresentam maior probabilidade de estarem situados na categoria 8.

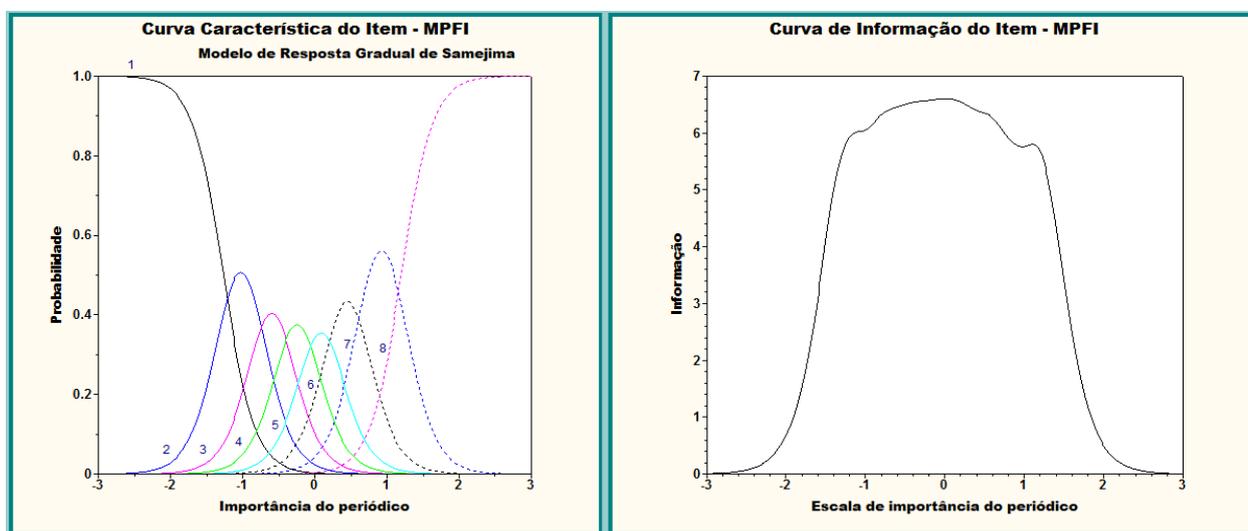


Figura 8.6: *Curva Característica do Item (CCI) e a Função de Informação do Item (MPFI) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).*

Com base na Figura 8.6 a Curva de Informação do Item (CII) (vide Seção 3.5) o seu valor máximo atinge para os traços latentes no intervalo de -1 a 0.2, aproximadamente.

De acordo com a Figura 8.7, a CCI associada ao item FI indica que os periódicos com nível de traço latente até -1.5 têm maior probabilidade de pertencerem à primeira categoria ordenada. Já os periódicos com traço latente entre -1.5 e 0, têm mais chance de pertencerem às categorias 1, 2, 3 e 4 ordenadas. Já os periódicos com habilidade entre -1 e 0.5 aproximadamente, possuem maior probabilidade de estarem situados nas categorias 5 e 6. Jornais com desempenho entre 0.5 e 1.5, aproximadamente, apresentam maior probabilidade de estar situada na categoria 7. Já os jornais com desempenho superior a 1.5, apresentam maior probabilidade de estarem situados na categoria 8. Na Figura 8.7 a curva à direita indica que a informação fornecida pelo item FI cai drasticamente para os escores menores do que -1 ou maiores do que 1.

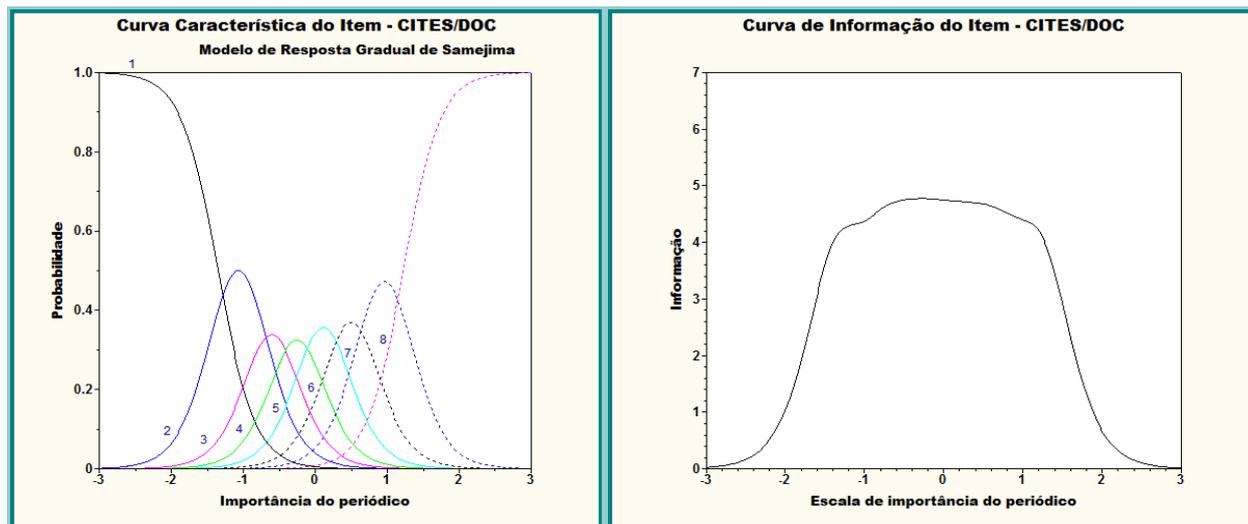


Figura 8.7: *Curva Característica do Item (CCI) e a Função de Informação do Item (CITES/DOC) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).*

Conforme pode ser visualizado na Figura 8.8 associada ao item FIS, verifica-se que a maior probabilidade encontra-se nas categorias 1 e 8. A categoria 3 não possui maior probabilidade para desempenhos entre -1.5 a 0, que é a escala apresentada no gráfico. Dessa forma, periódicos com desempenho de até -0.5, aproximadamente, apresentam maior probabilidade de situar-se na categoria 1, 2, 3, e 4 e periódicos com desempenho superior a 0.5, apresentam maior probabilidade de situar-se nas categoria 7 e 8. Na Figura 8.8, a curva à direita indica que a informação fornecida pela item FIS apresenta a maior quantidade de informação, atingindo valor superior a 4.5 na escala de quantidade de informação.

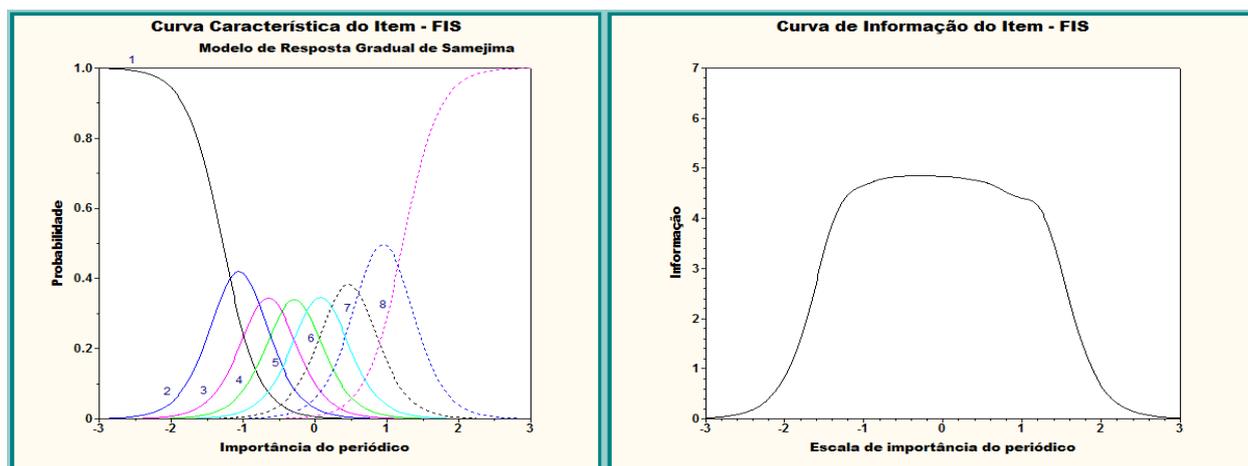


Figura 8.8: *Curva Característica do Item (CCI) e a Função de Informação do Item (FIS) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).*

De acordo com a Figura 8.9, a CCI associada ao item AIS indica que os periódicos com nível de traço latente até -1.5 têm maior probabilidade de pertencerem à primeira categoria ordenada. Já os periódicos com traço latente entre -1.5 e 0, têm mais chance de pertencerem à categoria 2. Finalmente, os periódicos com traço latente acima de zero têm mais chance de estarem associados à categoria 3. Na Figura 8.9 a curva à direita indica que a informação fornecida pelo item FI cai drasticamente para os escores menores do que 1.5.

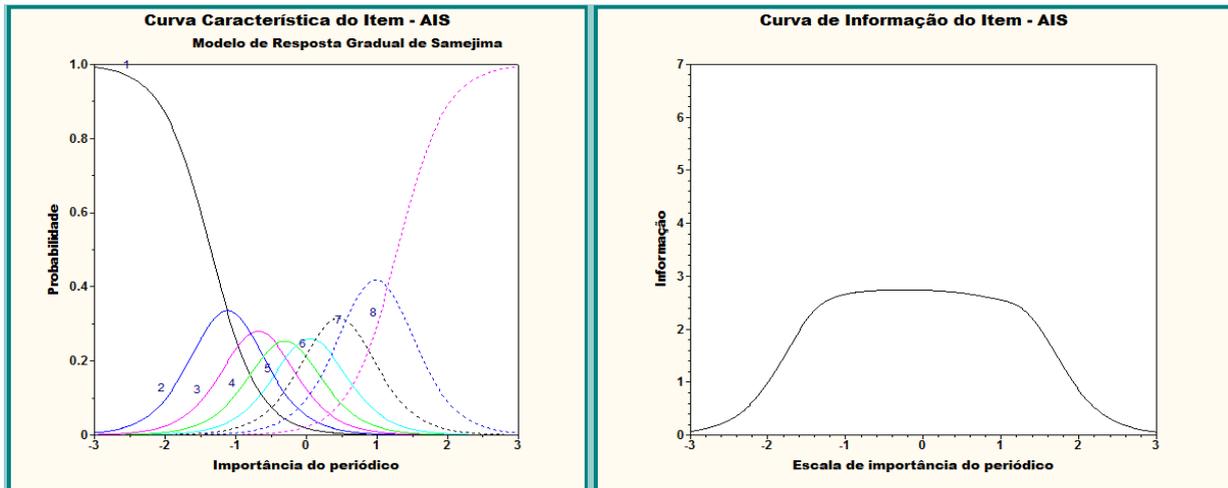


Figura 8.9: Curva Característica do Item (CCI) e a Função de Informação do Item (AIS) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

Na Figura 8.10 descreve a CCI associada ao item IH, indicando que os periódicos com traço latente até -0.5 têm maior probabilidade de pertencerem às categoria ordenadas de 1 a 5. Já os periódicos com traço latente entre -0.5 e 0.5 , têm mais chance de pertencerem às categorias 6 e 7. Finalmente, os periódicos com traço latente acima de 0.5 têm mais chance de estarem associados à categoria 8. Na Figura 8.10, a curva à direita indica que a informação fornecida pelo item HL é bastante modesta, descrevendo uma curva platicúrtica.

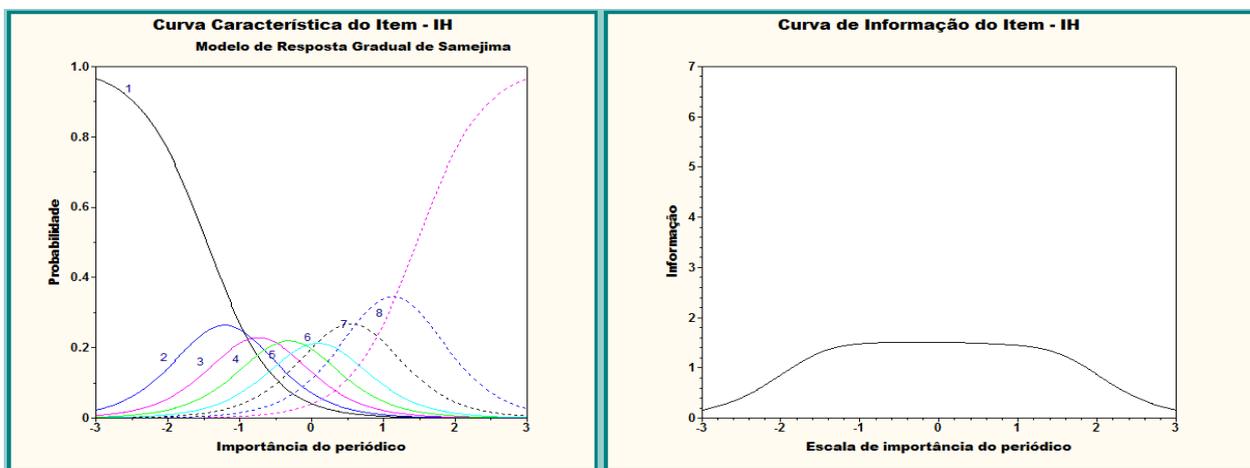


Figura 8.10: Curva Característica do Item (CCI) e a Função de Informação do Item (IH) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

Conforme pode ser visualizado na Figura 8.11 descreve a CCI associada ao item ES, indicando que os periódicos com traço latente até -0.5 têm maior probabilidade de pertencerem às categoria ordenadas de 1 a 5. Já os periódicos com traço latente entre -0.5 e 0.5 , têm mais chance de pertencerem às categorias 6 e 7. Finalmente, os periódicos com traço latente acima de 0.5 têm mais chance de estarem associados à categoria 8. Na Figura 8.11, a curva à direita indica que a informação fornecida pelo item HL é bastante modesta, descrevendo uma curva platicúrtica.

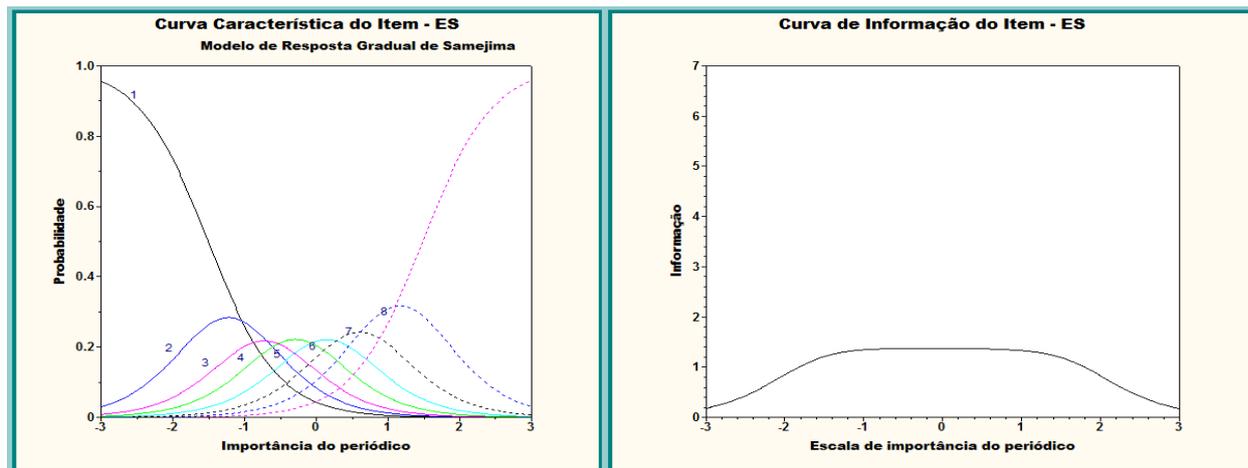


Figura 8.11: *Curva Característica do Item (CCI) e a Função de Informação do Item (ES) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).*

Na Figura 8.12 descreve a CCI associada ao item II, indicando que os periódicos com traço latente até -0.5 têm maior probabilidade de pertencerem às categoria ordenadas de 1 a 5. Já os periódicos com traço latente entre -0.5 e 0.5 , têm mais chance de pertencerem às categorias 6 e 7. Finalmente, os periódicos com traço latente acima de 0.5 têm mais chance de estarem associados à categoria 8. Na Figura 8.12, a curva à direita indica que a informação fornecida pelo item HL é bastante modesta, descrevendo uma curva platicúrtica.

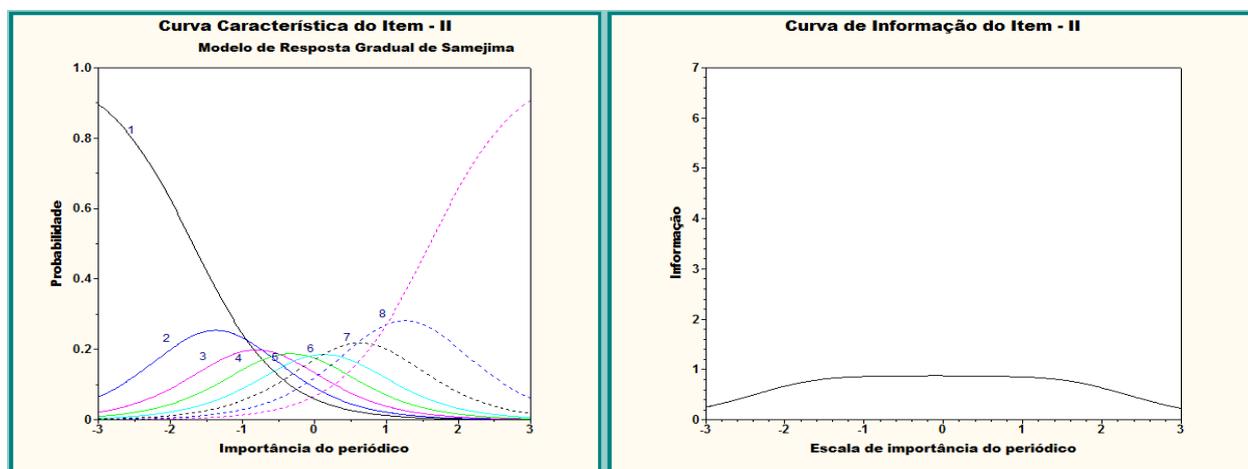


Figura 8.12: *Curva Característica do Item (CCI) e a Função de Informação do Item (II) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).*

Conforme pode ser visualizado na Figura 8.13 descreve a CCI associada ao item ES, indicando que os periódicos com traço latente até -0.5 têm maior probabilidade de pertencerem às categoria ordenadas de 1 a 5. Já os periódicos com traço latente entre -0.5 e 0.5 , têm mais chance de pertencerem às categorias 6 e 7. Finalmente, os periódicos com traço latente acima de 0.5 têm mais chance de estarem associados à categoria 8. Na Figura 8.13, a curva à direita indica que a informação fornecida pelo item HL é bastante modesta, descrevendo uma curva platicúrtica.

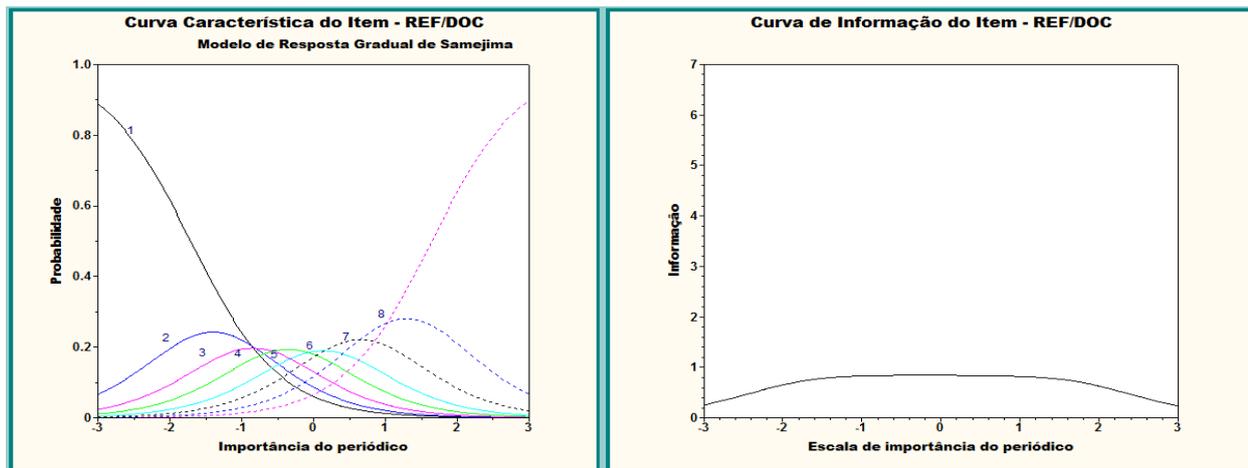


Figura 8.13: *Curva Característica do Item (CCI) e a Função de Informação do Item (REF/DOC) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).*

No que concerne à interpretação da CCI nas Figuras 8.14 e 8.15 associadas aos itens FI5A e CHL, respectivamente, como se pode perceber nos gráficos, os referidos itens são os que possuem as curvas características dos itens menos acentuadas. Ressalta-se que junções de categorias adjacentes foram realizadas para os itens.

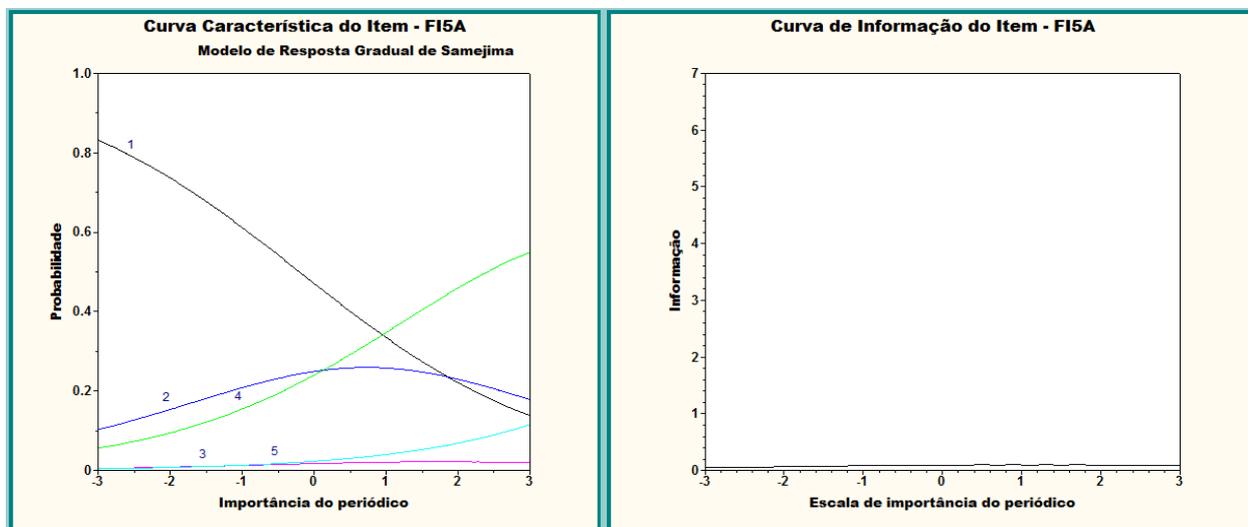


Figura 8.14: *Curva Característica do Item (CCI) e a Função de Informação do Item (FI5A) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).*

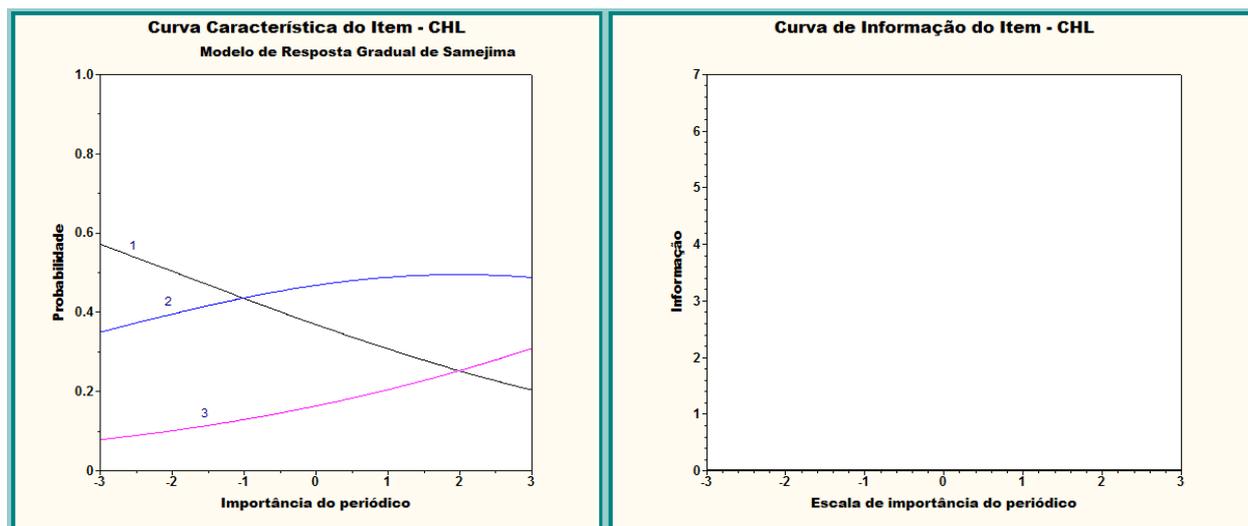


Figura 8.15: Curva Característica do Item (CCI) e a Função de Informação do Item (CHL) com maior poder de discriminação entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

Considerando o exposto, destaca-se que a soma das funções de informações de cada item resulta na informação do conjunto de itens, conhecida como função ou curva de informação do teste - FIT, como mostra a Figura 8.16.

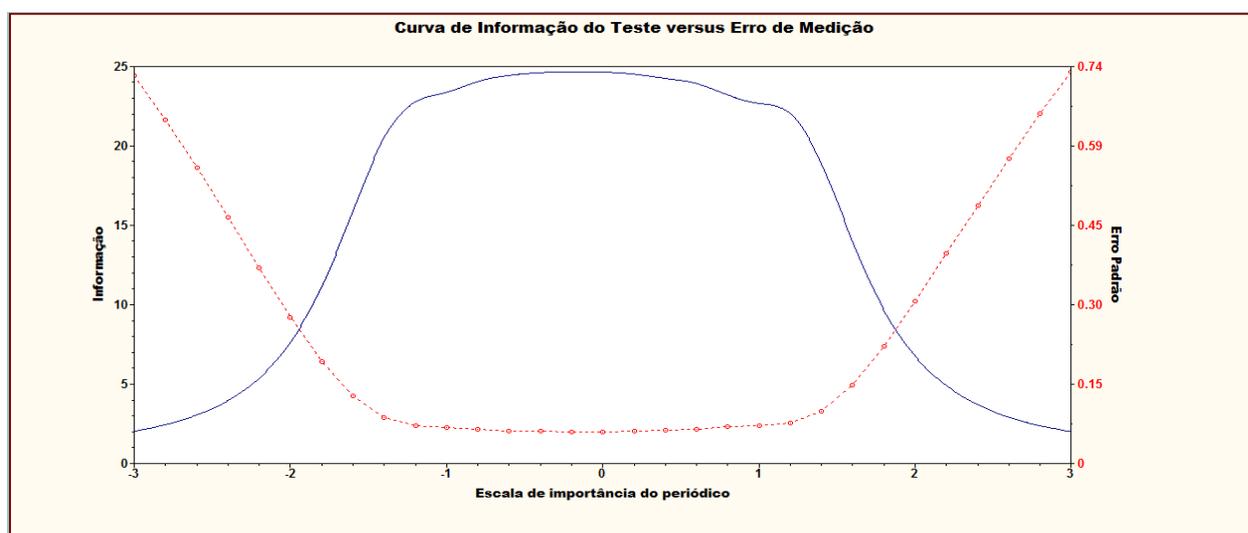


Figura 8.16: Curva de Informação do Teste (CIT) e o erro padrão de medição entre os periódicos e os itens selecionados. Elaborado pelos autores (2017).

Conforme pode ser visualizado na Figura 8.16, a função de informação do teste ou informação total do conjunto de itens possui maior quantidade de informação no intervalo de -1.0 a 1.0 da escala. A linha pontilhada mostra o erro padrão, configurando uma relação oposta à quantidade de informação. O erro padrão mais baixo situa-se também neste intervalo. Assim, para quantidades maiores de informações, menor o erro padrão.

Comparativamente à Figura 7.6, a inclusão de novos itens proporciona substancial ganho de informação. Observa-se na Figura 8.16 que as caudas da distribuição dos escores estão melhor descritas do que as que se observa na Figura 7.6. Além disso, nota-se que a curva descrita pela Figura 8.16 segue uma forma menos anômala que a observada na Figura 7.6.

Trabalhos futuros discutiremos a ordenação dos periódicos com base nos escores de *valor do periódico* estimados com base no modelo apresentado neste trabalho.

Capítulo 9

Considerações finais

O presente trabalho teve como objetivo geral criar um “índice de proficiência” ou valor dos periódicos na área de Matemática, Estatística e Matemática Aplicada baseado em Modelos Politômicos da Teoria da Resposta ao Item. Na busca desse objetivo, estabeleceram-se os seguintes objetivos específicos: identificar os componentes que influenciam a ordenação dos periódicos científicos; investigar o impacto dos indicadores descritos nas bases de dados desta pesquisa no que tange à classificação dos periódicos; desenvolver uma escala de medida padronizada para a classificação dos periódicos científicos nas áreas de Estatística, Matemática e Matemática Aplicada; avaliar a consistência das classificações apresentadas pelo Qualis CAPES.

Neste contexto, a sistemática proposta, que se refere ao objetivo geral, foi apresentada e detalhada no Capítulo 6 e sua aplicação foi desenvolvida nos capítulos 7 e 8. Ressalta-se que o desenvolvimento da escala de medida padronizada para a classificação dos periódicos científicos nas áreas de Estatística, Matemática e Matemática Aplicada, resultado da sistemática proposta, consiste no principal produto da pesquisa.

Destaca-se que a Teoria da Resposta ao Item possibilitou a criação de uma escala de medida padronizada e interpretável para a ordenação e classificação dos periódicos científicos, permitindo a realização de comparações entre os periódicos e a identificação e compreensão dos níveis de desempenho nos quais os periódicos estão posicionados.

Considerando o exposto, a pesquisa caracterizou-se como original em função de dois aspectos primordiais: o desenvolvimento de um modelo conceitual que considere a aplicação da TRI para a ordenação e classificação dos periódicos científicos e a criação de uma escala de medida padronizada para a classificação dos periódicos científicos nas áreas de Estatística, Matemática e Matemática Aplicada.

Ressalta-se, também, que se tem como propósito, à construção de uma base de dados final consistente, composta por diversos indicadores bibliométricos, que atendem aos objetivos desta pesquisa e que, também, contribuirá para outros estudos futuros.

Apresentam-se, na sequência, as conclusões concernentes aos objetivos da pesquisa e as sugestões para futuras pesquisas.

9.1 Conclusões

A sistemática proposta para a ordenação e classificação dos periódicos científicos, que consistiu no objetivo geral do presente trabalho, abrangeu algumas etapas: pré processamento dos dados; escolha da estratégia de imputação dos dados; classificação dos periódicos via CART; proposta para a classificação dos periódicos via TRI; comparação entre as classificações CAPES e via TRI; análise da persistência de efeito de grupo; e o monitoramento do

modelo, com a possibilidade de inserção de novas variáveis.

Na identificação dos componentes que influenciam a ordenação dos periódicos científicos e investigação do impacto dos indicadores descritos nas bases de dados desta pesquisa no que tange à classificação dos periódicos, realizou-se um levantamento bibliográfico. Com a realização deste estudo, foi possível identificar que inexistiu uma padronização na medição da classificação e rankings dos periódicos científicos. Há uma lacuna existente no que concerne a uma metodologia única de ordenação para a mensuração e classificação de periódicos. É importante observar também, o potencial existente para o desenvolvimento de modelos que possibilitam a ordenação de revistas científicas que considerem todos os indicadores bibliométricos para tal avaliação.

Para atender o objetivo específico de desenvolver uma escala de medida padronizada para a ordenação e classificação dos periódicos científicos, após a seleção dos itens, foi aplicada à TRI que possibilitou a estimação dos parâmetros de discriminação e de dificuldade dos itens e o posicionamento dos periódicos na escala quanto aos estratos de classificação.

Referente à etapa - escolha da estratégia de imputação dos dados. Para decidir acerca da melhor estatística de sumarização, se média ou mediana, a ser utilizada no processo de imputação dos dados, foram empreendidas algumas análises preliminares de modelos politômicos de TRI. Tais modelos foram ajustados tanto considerando imputações feitas utilizando a média quanto a mediana. Para tanto, considerou-se os dados de cada indicador bibliométrico e o cruzamento das informações referentes ao grupo ou área de dado periódico e o estrato ao qual o periódico foi classificado pela CAPES em 2012 (vide Tabela 6.3). Após a imputação os dados foram padronizados dentro de cada grupo (vide Seção 6.3) e discretizados para uso nos modelos politômicos.

O desenvolvimento da proposta para a classificação dos periódicos via TRI está descrito detalhadamente na seção 7.4.

Na discretização dos dados (vide Seção 6.4) utilizou-se 8 categorias, cada uma representando 12.5% dos dados. Adicionalmente, criou-se uma categoria artificial para acomodar 15 observações para as quais não foi possível efetuar a imputação. A esta categoria adicional atribuiu-se o valor 1. Às 8 categorias resultantes da discretização atribuiu-se valores de 2 a 9.

Na tentativa de avaliar a qualidade/deficiências dos dois métodos de imputação dos dados (vide Seção 6.2), modelos de TRI politômicos foram ajustados aos dados discretizados e previamente imputados (pela média/mediana). Os dados foram, então padronizados (vide Seção 6.3).

No desenvolvimento da presente pesquisa, advêm questões que merecem um maior destaque e investigações complementares, que possibilitam a elaboração de novas pesquisas, conforme destacados no Capítulo 6.6 e na Seção 7.3.1.

Uma comparação da aplicação das duas metodologias dos dois modelos (classificações CAPES e via TRI) permite salientar que as análises realizadas no Capítulo 7 não são definitivas, uma vez que os dados são esparsos e sofreram processos de imputação, com consequências ainda não plenamente ou satisfatoriamente estudadas. Além disso, estamos cientes da fragilidade de um método TRI que incorpora apenas 3 itens. Deve-se considerar, no entanto, que os resultados aqui descritos são fruto de um procedimento piloto, ou seja, fruto de um protótipo, a ser aperfeiçoado em análises futuras, e de posse de bases de dados mais informativas. Em resumo, os resultados aqui descritos precisam ser analisados com cautela.

Com a finalização da análise do modelo politômico, partiu-se para estudos cujo objetivo foi analisar a persistência de efeito de grupo. Na Seção 7.5 mencionou-se a possibilidade de que a padronização dos dados dentro dos grupos não tenha sido suficiente para eliminar possíveis efeitos de grupo nos indicadores bibliométricos. Neste trabalho os grupos são descritos

pelas áreas: Matemática pura, Matemática Aplicada, Estatística e Outras áreas. Adotou-se a área de Estatística como o *grupo de referência* e as demais áreas como *grupos focais*. As análises de DIF foram realizadas com o auxílio do *software* Parscale. Os programas (syntaxes) utilizados estão documentados no **Apêndice I**. Considerando o exposto, há evidências de que os efeitos diferenciais ainda estão presentes a despeito do processo de padronização.

Os métodos e modelos propostos nesta dissertação permitem o desenvolvimento e aprimoramento dos mesmos. Faz-se necessário, no entanto, a construção de uma base de dados abrangentes que incorpore indicadores bibliométricos importantes, como alguns já evidenciados no Capítulo 8 e o aprimoramento dos métodos de imputação de dados.

Referências Bibliográficas

- Adami(2016)** Anna Adami. Infoescola codecademy. <http://www.infoescola.com/informatica/codecademy/>, 2016. Último acesso em 10/2/2016. Citado na pág. 158
- Andersen(1980)** E. B. Andersen. *Discrete statistical models with social science applications*. North-Holland Publishing Company Amsterdam. Citado na pág. 54
- Andrade et al.(2000)** D. F. de Andrade, H. R. Tavares e da C. R. Valle. Teoria da resposta ao item: conceitos e aplicações. *ABE, São Paulo*. Citado na pág. xxv, 4, 23, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 36, 38, 39, 40, 42, 43, 45, 51, 52, 54, 55, 59, 60, 64, 66, 67
- Andrich(1978)** D. Andrich. A rating formulation for ordered response categories. *Psychometrika*, 43(4):561–573. Citado na pág. 24, 37
- Araujo et al.(2009)** E. A. C. de Araujo, D. F. de Andrade e S. L. V. Bortolotti. Item response theory. *Revista da Escola de Enfermagem da USP*, 43(SPE):1000–1008. Citado na pág. 23, 25, 27, 35
- Azevedo(2003)** Caio Lucidius Naberezny Azevedo. Métodos de estimação na teoria de resposta ao item. Dissertação de Mestrado, Instituto de Matemática e Estatística da Universidade de São Paulo, 27/02/2003. Citado na pág. 72, 73, 76, 77
- Baeza-Yates e Ribeiro-Neto(2013)** R. Baeza-Yates e B. Ribeiro-Neto. *Recuperação de Informação: Conceitos e Tecnologia das Máquinas de Busca*. Bookman Editora. Citado na pág. 204, 205
- Baker(1992)** F. B. Baker. *Item response theory: Parameter estimation techniques*. New York: Marcel Dekker, Inc. Citado na pág. 71
- Baker e Kim(2004)** F. B. Baker e Seock-Ho Kim. *Item response theory: Parameter estimation techniques*. CRC Press. Citado na pág. xxv, 41, 42, 52, 54, 60, 68
- Barra(2013)** Fernanda Barra. A importância de saber programar. <http://blogdopetcivil.com/2013/03/06/a-importancia-de-saber-programar/>, 2013. Último acesso em 10/2/2016. Citado na pág. 158
- Bartolucci et al.(2015)** F. Bartolucci, V. Dardanoni e F. Peracchi. Ranking scientific journals via latent class models for polytomous item response data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(4):1025–1049. Citado na pág. 3, 89
- Bechger et al.(2003)** T. M. Bechger, G. Maris, H. H. F. M. Verstralen e A. A. Béguin. Using classical test theory in combination with item response theory. *Applied Psychological Measurement*, 27(5):319–334. Citado na pág. 25

- Birnbaum(1968)** A. Birnbaum. Some latent trait models and their use in inferring an examinee's ability. *Statistical theories of mental test scores*, páginas 395–479. Citado na pág. 24, 29
- Bock(1972)** R. D. Bock. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37(1):29–51. Citado na pág. 24, 35, 71, 72
- Bock e Aitkin(1981)** R. D. Bock e M. Aitkin. Marginal maximum likelihood estimation of item parameters: Application of an em algorithm. *Psychometrika*, 46(4):443–459. Citado na pág. 25, 54
- Bock e Lieberman(1970)** R. D. Bock e M. Lieberman. Fitting a response model for dichotomously scored items. *Psychometrika*, 35(2):179–197. Citado na pág. 25, 55, 60, 63, 74, 77
- Bock e Zimowski(1997)** R. Darrell Bock e Michele F. Zimowski. Multiple group IRT. Em *Handbook of modern item response theory*, páginas 433–448. Springer. Citado na pág. 76, 77
- Bornia et al.(2009)** A. C. Bornia, D. F. Andrade, O. Possamai, P. M. R. Mafra e V. L. Almeida. Satisfação do congressista em relação ao congresso brasileiro de custos por meio da teoria da resposta ao item. Em *Anais do XVI Congresso Brasileiro de Custos-ABC*. Citado na pág. 65, 66
- Bortolotti et al.(2012)** S. L. V. Bortolotti, F. J. Moreira Junior, A. C. Bornia, A. F. Sousa Júnior e D. F. Andrade. Avaliação do nível de satisfação de alunos de uma instituição de ensino superior: uma aplicação da teoria da resposta ao item. *Gestão & Produção, São Carlos*, 19(2):287–302. Citado na pág. 26
- Breiman et al.(1984)** L. Breiman, J. Friedman, C. J. Stone e R. A. Olshen. *Classification and regression trees*. CRC press. Citado na pág. 95
- Brookes(1970)** B. C. Brookes. The growth, utility, and obsolescence of scientific periodical literature. *Journal of documentation*, 26(4):283–294. Citado na pág. 10
- Burton e Kebler(1960)** R. E. Burton e R. W. Kebler. The “half-life” of some scientific and technical literatures. *American documentation*, 11(1):18–22. Citado na pág. 10
- CAPES(2014)** CAPES. Coordenação de aperfeiçoamento de pessoal de nível superior. <http://www.capes.gov.br/>, Maio 2014. Último acesso em 1/5/2016. Citado na pág. xxix, 1, 2, 3
- Ceriani e Verme(2012)** Lidia Ceriani e Paolo Verme. The origins of the gini index: extracts from *variabilità e mutabilità* (1912) by Corrado Gini. *The Journal of Economic Inequality*, 10(3):421–443. Citado na pág. 95
- Cohen et al.(2003)** W. W. Cohen, P. D. Ravikumar, S. E. Fienberg et al. A comparison of string distance metrics for name-matching tasks. Em *IJWeb*, volume 2003, páginas 73–78. Citado na pág. 208, 209
- Contributors(2014)** Wikiversidade Contributors. Introdução às linguagens de programação em python. <http://pt.slideshare.net/brunomaspoli/trabalho-linguagem-python>, 2014. Último acesso em 10/2/2016. Citado na pág. 158

- Da-Silva et al.(2010)** A. D. Da-Silva, A. de M. Frota, F. P. da Silva, O. S. M. Romeo e T. S. Soares. Inovações no sistema de pareamento de domicílios e pessoas para a pesquisa de avaliação da cobertura da coleta do censo 2010. *Trabalho apresentado no XVII Encontro Nacional de Estudos Populacionais, Minas Gerais*. Citado na pág. 209, 215
- Dados(2016)** Ciências de Dados. Data science for professionals. <http://www.cienciaedados.com/r-ou-python-para-analise-de-dados/>, 2016. Último acesso em 11/2/2016. Citado na pág. 157
- De Ayala(2009)** R. J. De Ayala. The theory and practice of item response theory. new york, 2009. Citado na pág. 66
- Dempster et al.(1977)** A. P Dempster, N. M. Laird e D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, páginas 1–38. Citado na pág. 25
- Dias e Albieri(1992)** A. Dias e S Albieri. Uso de imputação em pesquisas domiciliares. *Encontro Nacional de Estudos Populacionais, VIII*, páginas 1–26. Citado na pág. 221, 222
- Embretson e Reise(2000)** S. E. Embretson e S. P. Reise. Item response theory for psychologists. Citado na pág. xxv, 4, 26, 27, 28, 29, 34, 35, 36, 38, 40, 42, 43, 52, 54, 60, 114, 115
- Fellegi e Sunter(1969)** I. P. Fellegi e A. B. Sunter. A theory for record linkage. *Journal of the American Statistical Association*, 64(328):1183–1210. Citado na pág. 208, 209
- Gill(2001)** L. Gill. *Methods for automatic record matching and linkage and their use in national statistics*. Number 25. Office for National Statistics. Citado na pág. 203, 204
- Gini e Lexis(1912)** C Gini e W Lexis. Variabilità e mutabilità. contributo allo studio delle distribuzioni e delle relazioni statistiche. fascicolo i, 1912. Citado na pág. 95
- Guerrero-Botea e Moya-Anegónb(2012)** V. P. Guerrero-Botea e F. Moya-Anegónb. A further step forward in measuring journals' scientific prestige: The sjr2 indicator. *Journal of Informetrics*, 6:674–688. Citado na pág. 15
- Gustafsson(1980)** J. E. Gustafsson. A solution of the conditional estimation problem for long tests in the rasch model for dichotomous items. *Educational and Psychological Measurement*, 40(2):377–385. Citado na pág. 62
- Guttman et al.(1950)** L. Guttman, S. A. Stouffer, E. A. Suchman, P. F. Lazarsfeld, S. A. Star e J. A. Clausen. The basis for scalogram analysis. 4:60–90. Citado na pág. 23
- Hair et al.(2005)** J. F. Hair, R. E. Anderson, R. L. Tatham e W. C. Black. *Análise multivariada de dados. 5ª Edição*. Porto Alegre: Bookman. Citado na pág.
- Hambleton e Jones(1993)** R. K. Hambleton e R. W. Jones. Theory and their applications to test development. *Educational Measurement: Issues and Practice*. Citado na pág. 25
- Hambleton e Swaminathan(1985)** R. K. Hambleton e H. Swaminathan. *Item Response Theory: Principles and Applications*. Springer Netherlands, 1 edição. Citado na pág. 43, 45, 52, 54, 60, 61, 62, 68

- Hambleton et al.(1991)** R. K. Hambleton, H. Swaminathan, H. J. Rogers e R. M. Jaeger. Fundamentals of item response theory. Sage publications. *Inc, Newbury Park, CA*. Citado na pág. 25, 26
- Ilyankou(2014)** I. Ilyankou. Comparison of jaro-winkler and ratcliff/obershelp algorithms in spell check. *IB Extended Essay Computer Science*. Citado na pág. 205, 207
- Issac e Keller(1966)** E. Issac e H. B. Keller. Analysis of numerical methods, 1966. Citado na pág. 42
- Jaccard(1901)** P. Jaccard. Etude comparative de la distribution florale dans une portion des alpes et du jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37:547–579. Citado na pág. 209
- Jaro(1989)** M. A. Jaro. Advances in record-linkage methodology as applied to matching the 1985 census of tampa, florida. *Journal of the American Statistical Association*, 84 (406):414–420. Citado na pág. 203, 206, 208
- Jaro(1995)** M. A. Jaro. Probabilistic linkage of large public health data files. *Statistics in medicine*, 14(5-7):491–498. Citado na pág. 206, 208, 209
- Johnson e Wichern(2002)** R. A. Johnson e D. W. Wichern. *Applied multivariate statistical analysis*, volume 5. Prentice hall Upper Saddle River, NJ. Citado na pág.
- Karishma e Singh(2012)** Shukla Karishma e Mahesh Singh. Personalized web crawler for invisible web. *International Journal of Mathematical Archive (IJMA) ISSN 2229-5046*, 3 (5). Citado na pág. 218
- Lattin et al.(2011)** J. Lattin, J. D. Carroll e P. E. Green. *Análise de dados multivariados*. Citado na pág.
- Lawley(1943)** D. N. Lawley. Xxiii.ãon problems connected with item selection and test construction. *Proceedings of the Royal Society of Edinburgh. Section A. Mathematical and Physical Sciences*, 61(03):273–287. Citado na pág. 23
- Lazarsfeld et al.(1950)** P. F. Lazarsfeld, S. A. Stouffer, L. Guttman, E. A. Suchman, S. A. Star e J. A. Clausen. The logical and mathematical foundation of latent structure analysis. 4:362–412. Citado na pág. 23
- Levenshtein(1966)** V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. Em *Soviet physics doklady*, volume 10, páginas 707–710. Citado na pág. 210
- Likert(1932)** R. A. Likert. Technique for the measurement of attitudes. *Archives of Psychology*, (140):44–53. Citado na pág. 35
- Lord(1952)** F. M. Lord. A theory of test score (psychometric monograph no. 7). *Iowa City, IA: Psychometric Society*, página 35. Citado na pág. 23, 24, 29, 34
- Lord(1980)** F. M. Lord. *Applications of item response theory to practical testing problems*. Routledge. Citado na pág. 29, 30
- Lord et al.(1968)** F. M. Lord, M. R. Novick e A. Birnbaum. Statistical theories of mental test scores. Citado na pág. 23, 24, 25, 27, 43

- Mafra(2010)** P. M. R. Mafra. *Proposta de uma sistemática para a modelagem de risco de crédito sob a perspectiva da teoria da criação do conhecimento: uma abordagem da teoria da resposta ao item. 2010*. Tese de Doutorado, Tese (Doutorado em Engenharia e Gestão do Conhecimento)-Universidade Federal de Santa Catarina, Florianópolis. Citado na pág. 60
- Masters(1982)** G. N. Masters. A rasch model for partial credit scoring. *Psychometrika*, 47(2):149–174. Citado na pág. 39
- Mingoti(2005)** S. A. Mingoti. *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Editora UFMG. Citado na pág.
- Mislevy(1986)** R. J. Mislevy. Bayes modal estimation in item response models. *Psychometrika*, 51(2):177–195. Citado na pág. 42
- Mislevy e Bock(1982)** RJ Mislevy e RD Bock. Bilog, maximum likelihood item analysis and test scoring: Logistic model [computer software]. mooreville, in: Scientific software, 1982. Citado na pág. 62
- Moed(2010)** H. F. Moed. Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3):265–277. Citado na pág. 17
- Muraki(1992)** E. Muraki. A generalized partial credit model: Application of an em algorithm. *ETS Research Report Series*, 16:159–176. Citado na pág. 24, 39, 116
- Nascimento(2011)** C. Nascimento. Um arcabouço para busca e recomendação de artigos científicos. Dissertação de Mestrado, (Mestrado em Ciência da Computação - Instituto de Ciências Exatas - Universidade Federal de Minas Gerais, Belo Horizonte. Citado na pág. 218
- Navarro(2001)** G. Navarro. A guided tour to approximate string matching. *ACM computing surveys (CSUR)*, 33(1):31–88. Citado na pág. 210
- Nunes et al.(2009)** L.N. Nunes, M. M. Kluck e J. M. G. Fachel. Uso da imputação múltipla de dados faltantes: uma simulação utilizando dados epidemiológicos. *Cadernos de saúde pública. Rio de Janeiro. Vol. 25, no. 2 (fev. 2009), p. 268-278*. Citado na pág. 221, 222
- Nunes et al.(2008)** S. da-S. Nunes, C. Henrique, C. Primi, M. F. Oliveira Nunes, M. Muniz, T. da-F. Cunha e G. Couto. Teoria de resposta ao item para otimização de escalas tipo likert-um exemplo de aplicação. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 1(25):51–79. Citado na pág. 27
- Onoda e Ebecken(2001)** Mauricio Onoda e Nelson FF Ebecken. Implementação em java de um algoritmo de árvore de decisão acoplado a um sgbd relacional. Em *SBBD*, páginas 55–64. Citado na pág. 95
- Ostini e Nering(2006)** R. Ostini e M. L. Nering. Polytomous item response theory models. series: Quantitative applications in the social sciences, 2006. Citado na pág. 63
- Pet-Tele(2011)** Grupo Pet-Tele. Tutorial de introdução ao python. <http://pt.wikipedia.org/w/index.php?title=Python&oldid=40719582>, 2011. Último acesso em 10/2/2016. Citado na pág. 158
- Rao(1973)** C. R. Rao. 1973. linear statistical inference and its applications, 1973. Citado na pág. 42

- Rasch(1960)** G. Rasch. Probabilistic models for some intelligence and achievement tests. *Copenhagen: Danish Institute for Educational Research*. Citado na pág. 24, 29, 34
- Ratcliff e Metzener(1988)** J. W. Ratcliff e D. E. Metzener. Pattern matching: the gestalt approach. *Dr Dobbs Journal*, 13(7):46. Citado na pág. 203, 204, 205
- Ratcliffe(1983)** J. W. Ratcliffe. Notions of validity in qualitative research methodology. *Science Communication*, 5(2):147–167. Citado na pág. 205
- Regazzi(2000)** A. J. Regazzi. *Análise multivariada*, volume 2. Notas de aula INF 766, Departamento de Informática da Universidade Federal de Viçosa. Citado na pág.
- Rodichevski(2013)** A. Rodichevski. Approximate string-matching algorithms, part 2. <http://www.morfoedro.it/doc.php?n=223&lang=en>, 2013. Último acesso em 11/4/2016. Citado na pág. 207
- Rubin(1987)** D. B. Rubin. Multiple imputation for nonresponse in surveys, 1987. Citado na pág. 222
- Samejima(1969)** F. A. Samejima. Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, 17. Citado na pág. 23, 24, 36, 37, 77
- Samejima(1997)** F. A. Samejima. Graded response model. *Handbook of Modern Item Response Theory*. Citado na pág. 75, 76, 77
- Scheidt et al.(2008)** R. F. Scheidt, R. Fernandes e A. Maria. Desenvolvimento de um sistema de busca de conteúdo científico baseado em bibliometria e mineração de textos. *Anais SULCOMP*, 4. Citado na pág. 218
- Sörensen(1948)** T. Sörensen. {A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons}. *Biol. Skr.*, 5:1–34. Citado na pág. 210, 211
- Sutton(2005)** C. D. Sutton. Classification and regression trees, bagging, and boosting. *Handbook of statistics*, 24:303–329. Citado na pág. 202
- Teicher e Chow(1978)** H. Teicher e Y. S. Chow. *Probability theory: independence, interchangeability, martingales*. Springer. Citado na pág. 56
- Thissen et al.(1993)** D. Thissen, L. Steinberg e H. Wainer. Detection of differential item functioning using the parameters of item response models. Citado na pág. 115
- Turik(2010)** C. Turik. Análise de atitudes de alunos universitários em relação à estatística por meio da teoria da resposta ao item (tri). Citado na pág. xxv, 27
- Urry(1976)** V. W. Urry. Ancillary estimators for the item parameters of mental tests. *Washington, DC: Personnel Research and Development Center, US Civil Service Commission*. Citado na pág. 62
- van Belle e Arnold(2000)** G. van Belle e A. Arnold. Reliability of cognitive tests used in alzheimer’s disease. *Statistics in medicine*, 19(11-12):1411–1420. Citado na pág. 25
- Van Der Linden e Hambleton(1997)** W. J. Van Der Linden e R. K. Hambleton. Item response theory: Brief history, common models, and extensions. Em *Handbook of modern item response theory*, páginas 1–28. Springer. Citado na pág. 63, 68

- Van Rossum e De Boer(1991)** G. Van Rossum e J. De Boer. Linking a stub generator (ail) to a prototyping language (python). Em *Proceedings of the Spring 1991 EurOpen Conference, Troms, Norway*, páginas 229–247. Citado na pág. 155
- Vargas et al.(2007)** V. C. C. Vargas et al. *Medida padronizada para avaliação de intangíveis organizacionais por meio da teoria da resposta ao item*. Tese de Doutorado, Tese (Doutorado em Engenharia de Produção) - Programa de Pós Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis. Citado na pág. 66, 67
- Varin et al.(2016)** C. Varin, M. Cattelan e D. Firth. Statistical modelling of citation exchange between statistics journals. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(1):1–63. Citado na pág. 3
- Vieira et al.(2008)** K. Vieira, L. Barbosa, J. Freire e A. Silva. Siphon++: a hidden-webcrawler for keyword-based interfaces. Em *Proceedings of the 17th ACM conference on Information and knowledge management*, páginas 1361–1362. ACM. Citado na pág. 219
- Waltman et al.(2013)** L. Waltman, N. J. van Eck, T. N. van Leeuwen e M. S. Visser. Some modifications to the snip journal impact indicator. *Journal of Informetrics*, 7(2):272–285. Citado na pág. 17
- Wingersky(1983)** M. S. Wingersky. Logist: A program for computing maximum likelihood procedures for logistic test models. *Applications of item response theory*, páginas 45–56. Citado na pág. 61
- Wingersky et al.(1982)** M. S. Wingersky, M. A. Barton e F. M. Lord. *LOGIST user's guide: LOGIST 5, version 1.0*. Educational Testing Service. Citado na pág. 62
- Winkler(1990)** W. E. Winkler. String comparator metrics and enhanced decision rules in the fellegi-sunter model of record linkage. Citado na pág. 206, 208
- Winkler(1999)** W. E. Winkler. The state of record linkage and current research problems. Em *Statistical Research Division, US Census Bureau*. Citeseer. Citado na pág. 203, 204, 209
- Wright(1968)** B. D. Wright. *Sample-free test calibration and person measurement*. publisher not identified. Citado na pág. 24
- Wright(1977)** B. D. Wright. Misunderstanding the rasch model. *Journal of Educational Measurement*, páginas 219–225. Citado na pág. 29, 34
- Wright e Mead(1976)** B. D. Wright e R. J. Mead. Bical: Calibrating rating scales with the rasch model. *Research Memorandum*, 23. Citado na pág. 61, 62
- Zealand(2013)** S. N. Zealand. Data integration manual: Tuoto, t. relais: Donot get lost in a record linkage project. <http://www.fcsm.gov/07papers/Tuoto.VI-C.pdf>, 2013. Último acesso em 23/2/2016. Citado na pág. 158, 204

Apêndice A - Apresentação da linguagem *Python*

A.1 Apresentação da linguagem *Python*

O *Python* é uma linguagem de programação orientada a objetos de alto nível, e com grau de abstração relativamente elevado, longe do código de máquina e mais próximo à linguagem humana. A linguagem foi criada por [Van Rossum e De Boer \(1991\)](#). O *Python* é uma linguagem de programação com uma sintaxe muito simples e intuitiva, mas ao mesmo tempo muito poderosa, que pode ser usada por todo tipo de usuários, por ser mais próxima da linguagem humana, tornando-se uma ferramenta ideal para iniciantes em programação. Adicionalmente, a linguagem prioriza a forma prática de leitura a velocidade.

Atualmente, o desenvolvimento da linguagem é comunitário, participativo, aberto, coordenado por uma agência sem fins lucrativos a, *Python Software Foundation* como gestor do projeto e apresenta diferentes tipos de dialetos.

Nesta conjectura, tem-se que dialeto corresponde a forma de como uma língua é realizada numa região específica. No caso dos dialetos *Python*, tem-se uma camada de escrita *Python* que executa sobre a camada de recursos de um ambiente. Como exemplos de dialetos *Python*, pode-se elencar: *i)* IronPython: é a implementação da linguagem para a plataforma .NET; *ii)* PyPy: corresponde a implementação de *Python* escrita em *Python*; *iii)* Jython: implementação de *Python* para máquinas virtuais Java.

A.2 Características básicas da linguagem

Python é uma linguagem de programação interpretada, de código-fonte aberto e disponível para vários sistemas operacionais e não precisa ser compilada (traduzida para uma linguagem da máquina), mas sim “lida” por um outro programa, chamado de interpretador, que traduz para a máquina o que seu programa quer dizer.

O interpretador para *Python* é interativo, ou seja, é possível executá-lo sem fornecer um script (programa) para ele. Ao invés disso, o interpretador disponibiliza uma interface interativa onde é possível inserir os comandos desejados, um por um e observar o efeito de cada um deles, ou seja, a cada comando inserido no interpretador interativo, pode-se observar o resultado gerado.

O *Python* é ideal para scripting e para o desenvolvimento rápido de aplicações em diversas áreas e na maioria das plataformas, pois a linguagem foi projetada com a filosofia de enfatizar a importância do esforço do programador sobre o esforço computacional. Um das características do *Python* é priorizar a legibilidade do código.

A.3 Plataformas disponíveis

A linguagem e seu interpretador estão disponíveis para as mais diversas plataformas, desde Unix (Linux, FreeBSD, Solaris, MacOS X, etc.), Windows, .NET, versões antigas de MacOS até consoles de jogos eletrônicos ou mesmo alguns celulares, como a série 60, N8xx(PyMaemo) da Nokia e palmtops.

Para algum sistema operacional não incluso na relação acima, basta que exista um compilador C disponível e a geração do *Python* a partir da fonte. O código fonte é traduzido pelo interpretador para o formato bytecode, que é multiplataforma e pode ser executado e distribuído sem fonte original.

A.4 Características da linguagem *Python*

- É uma linguagem de programação em que não é necessário compilar seu código fonte para que a máquina o entenda.
- Roda em ambientes Linux, Windows, MacOS, smartphones, celulares, e outra infinidade de sistemas.
- É uma linguagem totalmente orientada a objetos e permite que o programador desenvolva os códigos de forma procedural ou funcional.
- É um Software Livre
- Possui código aberto.
- É uma linguagem multiuso, pois permite criar desde aplicativos para desktop a web sites.

A.5 Algumas das vantagens de programar utilizando o *Python*

- (i) O *Python* não possui complicações, tem uma sintaxe simples, transparente e intuitiva. Seus blocos de código são definidos através de indentação *Python*. Também possui um código limpo o que torna a leitura mais fácil;
- (ii) Foi feito para ser usado na internet;
- (iii) Por não precisar declarar variáveis, economiza tempo: o interpretador faz essa tarefa;
- (iv) Uma facilidade muito importante é a possibilidade de somar uma string com um inteiro, gerando assim um erro, que é possível também ser tratado. Isso porque a linguagem não mistura os vários tipos de variáveis;
- (v) Comunicação Padronizada: XML;
- (vi) O *Python*, por ser uma linguagem de alto nível, facilita a sua utilização.

A.6 Instalação do *Python*

O *Python* já vem instalado em programas como **OpenOffice.org** e o executável **python.exe** é encontrado no mesmo diretório que os outros aplicativos. A versão mais atual, 3.4.2, está disponível em www.python.org/downloads/ para download gratuito para usuários Windows. A instalação é muito simples, basta seguir as instruções que aparecem na tela e clicar em avançar sempre. Para o usuário Linux, o *Python* já vem instalado na maioria das versões. Caso o usuário não disponha do *Python*, é possível digitar `sudo apt-get install` para a instalação.

A.7 A linguagem *Python* na criação de bases de dados

Devido a sua versabilidade, tem sido frequente a opção pela linguagem *Python* não só no desenvolvimento Web, mas em muitos outros tipos de aplicação. *Python* tem sido utilizado nos servidores de mega-sites como o Google e YouTube, nos clusters de computação gráfica em laboratórios da NASA, da indústria e em games como o Civilization IV e EVE-Online.

Além de sua sintaxe simples e clara, o *Python* oferece os seguintes recursos, disponíveis, também, em linguagens mais complexas como Java e C++:

- programação orientada a objetos (incluindo herança múltipla, um conceito apenas parcialmente presente em Java);
- por “exceções” (exceptions); um moderno mecanismo para o tratamento de erros denominado;
- programação modular, uma forma inteligente de acessar e organizar código a ser reutilizado;
- coleta de lixo automática, sistema que elimina os erros causados pelo acúmulo de dados inúteis na memória do computador (característica presente também em Java, mas não em C++);
- recursos avançados de manipulação de textos, listas e outras estruturas de dados;
- universalidade: é possível executar o mesmo programa sem modificações em várias plataformas de hardware e sistemas operacionais (difícil de se conseguir em C++).

A.8 R ou *Python* para análise de dados?

Quando se trata de Ciência de Dados, uma pergunta sempre vem à tona: R ou *Python* para análise de dados? Apesar de existirem muitas outras possibilidades, estas duas linguagens tem polarizado as discussões sobre qual ferramenta utilizar para análise. As duas linguagens são simples (e gratuitas) para instalar e relativamente fáceis de começar a usar. Um iniciante em Ciência de Dados em geral não tem experiência com programação, optando por aprender R ou *Python* primeiro (Dados, 2016).

Embora o *Python* não disponha de um conjunto de pacotes e bibliotecas tão abrangente como os disponíveis para a linguagem R, a combinação de *Python* com ferramentas como Pandas, Numpy, Scipy, Scikit-learn e Seaborn, torna a linguagem uma das principais escolhas entre os Cientistas de Dados. A linguagem também está lentamente tornando-se útil

para tarefas em aprendizado de máquinas e base para o trabalho estatístico intermediário (anteriormente apenas sob o domínio de R).

O *Python* é um linguagem de programação mais intuitiva que o R, visto que seu objetivo é aproximar a linguagem humana. Desta forma, apresenta-se como uma maneira mais simples e flexível de programar.

No que concerne a linguagem de programação *Python*, para um estudo mais detalhado, veja [Adami \(2016\)](#), [Barra \(2013\)](#), [Pet-Tele \(2011\)](#), [Zealand \(2013\)](#) e [Contributors \(2014\)](#).

Apêndice B - Comandos utilizando o *Python*

B.1 Programação

Neste apêndice, encontra-se todos os códigos fonte utilizados na geração dos pareamentos realizados na linguagem *Python*, assim como o código fonte do *Crawler* e do estudo do ponto ótimo, disponibilizados no endereço eletrônico (<http://bit.ly/1XZk5VI>).

Para utilizar os códigos, basta baixar os arquivos mantendo a estrutura das pastas, observando que na pasta “Scripts in Python” existem mais duas pastas, respectivamente “Source” e “Setup”.

A pasta “Source” contém todos os arquivos fonte utilizados:

- **spider_CWTS.py**: código utilizado para realizar a busca na web dos códigos de ISSN da amostra de periódicos retirados do site da CWTS. Para executar este código execute o comando:
 - **python spider_CWTS.py runspider -a ../../files/input/banco_dados_CWTS_area_matematica_2014_utf8.csv -o ISSN_CWTS.csv --nolog**
 - **OBSERVAÇÃO**: repare que o comando acima informa o arquivo de *entrada* após o argumento “-a”, o arquivo de *saída* após o argumento “-o”, sendo o argumento “--nolog” utilizado para desligar a visualização de acesso e retorno das chamadas ao protocolo HTTP (*Hypertext Transfer Protocol*), em português *Protocolo de Transferência de Hipertexto*, isto é, um protocolo de comunicação utilizado para sistemas de informação de hipermídia, distribuídos e colaborativos, usado como base para a comunicação de dados da World Wide Web.
- **Pareador_v_1.4.py**: código utilizado para realizar os pareamentos, com base nos arquivos de configuração disponibilizados. Para executar este código execute o comando:
 - **python Pareador_v_1.4.py [arquivo de configuração]**
 - **OBSERVAÇÃO**: para que o código execute sem problemas, é necessário que as bibliotecas *difflib*, *Levenshtein* e *distance*, estejam instaladas no ambiente em que for executar. Portanto, é importante checar sua instalação antes de iniciar os pareamentos.
- **Parser_Pareamento.py**: biblioteca criada para tratar o parseamento (em ciência da computação e linguística, análise sintática, também conhecida pelo termo em inglês *parsing*). É o processo de analisar uma sequência de entrada, lida de um arquivo de computador ou do teclado, por exemplo, para determinar sua estrutura gramatical

segundo uma determinada gramática formal) do arquivo de configuração a ser utilizado, mantidos na pasta “*Setup*”, também disponibilizada neste trabalho.

A pasta “*Setup*” contém todos os arquivos de configuração utilizados. Cada arquivo tem minimamente 4 sessões, a saber:

- **File_IN:** sessão para registrar os arquivos de entrada a serem pareados. É possível parear quantos arquivos quiser, numerando cada argumento (mnemônico formado pelo prefixo “*fd_*” concatenado ao número indicando o arquivo de entrada).
- **File_OUT:** sessão para registrar o arquivo a ser gerado pelo pareamento. Utiliza-se de alguns *argumentos* (não importando a ordem em que são declarados) que funcionam como parâmetros para definir ações a serem consideradas pelo arquivo *Pareador_v_1.4.py* ao realizar os pareamentos. Os seguintes argumentos podem ser utilizados (todos devem ser escritos **estritamente** com letras minúsculas):
 - **fd_out:** utilizado para informar o nome do arquivo a ser gerado.
 - **path:** utilizado para informar o nome da pasta onde o arquivo informado na sessão, deverá ser gravado (é importante colocar o “/” no final do nome).
 - **sep:** caractere a ser utilizado como separador de campos no arquivo da sessão (é necessário declarar o caractere entre aspas).
 - **unique:** utilizado para informar o tipo de pareamento realizado. Assume valor **1** se o pareamento é realizado por chave direta, ou **0** se o pareamento é realizado via medidas de similaridade (este trabalho considerou a variável *título* para avaliar a similaridade entre duas bases de dados).
 - **prob_min:** utilizado para definir um *ponto de corte* que representa o valor mínimo para o qual considera-se os registros pareados, segundo a medida de similaridade calculada entre as variáveis *título* das bases de dados a serem pareadas.
 - **method:** utilizado para definir a medida de similaridade considerada (as opções são: *jaccard*, *jaro*, *jaro_winckler*, *levenshtein*, *ratcliff*, *sorensen* e *best* - este último é utilizado para gerar um arquivo para se estudar o ponto de corte ótimo).
 - **prob_max:** utilizado unicamente quando o argumento “*method=best*”, indica o ponto de corte *máximo* a ser considerado na busca pelo ótimo.
 - **prob_step:** utilizado unicamente quando o argumento “*method=best*”, indica o valor dos saltos entre os pontos de corte considerados (neste trabalho utilizou-se 0.1).
- **File [#arquivo de entrada]:** sessão que descreve detalhes de cada arquivo de entrada (a expressão “*#arquivo de entrada*” refere-se ao número que identifica o arquivo na sessão *File_IN*. Os seguintes argumentos são especificados:
 - **path:** utilizado para informar o nome da pasta onde o arquivo informado na sessão, deverá ser gravado (é importante colocar o “/” no final do nome).
 - **sep:** caractere a ser utilizado como separador de campos no arquivo da sessão (é necessário declarar o caractere entre aspas).
 - **cab:** utilizado para descrever o nome dos campos do arquivo em questão (se estiver em branco considera os nomes como estão no cabeçalho do arquivo de entrada).

- **keys:** utilizado para informar a posição do campo que será considerado como chave de pareamento (no caso de pareamento por similaridade, neste trabalho, remete-se à posição dos campos onde estão os títulos dos periódicos) no arquivo de entrada.
- **vars:** utilizado para informar a posição dos campos do arquivo de entrada, que deverão estar presentes no arquivo de saída.
- **prefix_cab:** utilizado para renomear campos (no geral o código gera automaticamente no arquivo de saída um prefixo associado ao arquivo de entrada, para que se possa identificar de onde a informação no arquivo de saída é proveniente. Este argumento é opcional para que se possa escolher este prefixo.).

Apêndice C - Comandos utilizando o *Software R*

C.1 Programação

Neste apêndice, encontram-se os scripts na linguagem R utilizados para a geração dos pareamentos entre as bases de dados desta pesquisa, disponibilizados no endereço eletrônico (<http://bit.ly/1VE0dac>).

A pasta contém os seguintes scripts:

- **Pareamentos_COM_ISSN.R:** script utilizado para realizar os pareamentos com as bases de dados do trabalho, considerando que *todas*, possuem o código do ISSN (este script foi testado após a obtenção dos códigos de ISSN da base do CWTS pelo *crawler*).
- **OBSERVAÇÃO:** não foram tratados no R pareamentos por similaridade, sendo estes tratados pelos códigos no **Apêndice B**.
- **Analise_Otimo.R:** script utilizado para gerar boxplots para avaliação do ponto de corte ótimo, resultante da saída do código *Gera_Dados_Grafico.py*, descrito no **Apêndice B**.

Apêndice D - Bases de dados desta pesquisa

D.1 Bases de dados

Neste apêndice, encontra-se as bases de dados utilizadas nesta pesquisa, disponibilizados no endereço eletrônico (<http://bit.ly/1Xsh9SY>). Estas bases de dados estão apresentadas no escopo deste trabalho de dissertação.

Apêndice E - Algumas análises descritivas das bases de dados desta pesquisa

E.1 Introdução

A análise descritiva é uma das áreas da Estatística que tem por objetivo descrever e sintetizar um conjunto de dados, correspondendo a uma fase muito importante de uma pesquisa.

Nos estudos que envolvem análise descritiva, é sempre interessante apresentar-se os resultados da pesquisa em tabelas e gráficos, pois a compreensão se torna mais fácil para o público interessado. É nesta fase que conseguiu-se organizar e resumir as informações de interesse a uma única variável, referente a um conjunto de dados.

Na sequência, apresenta-se uma das bases de dados que fazem parte dos estudos da dissertação em desenvolvimento. No que concerne à referida apresentação, tem-se: a descrição das suas variáveis, principais características e procedência dos dados utilizados.

Vale ressaltar que os referidos bancos de dados foram verificados até mesmo manualmente, na investigação de problemas de inconsistência de natureza diversa.

Tal procedimento foi motivado devido às diferentes formatações dos dados e as necessidades específicas deste trabalho.

E.2 Periódicos no Qualis/CAPES

Completando o exposto, na Tabela E.1, tem-se algumas informações sobre a base de dados Qualis. Tal base de dados, foi retirada do site da CAPES (<http://qualis.capes.gov.br/webqualis>).

Tabela E.1: *Resultados apurados na base do PDF (dados originais CAPES).*

Informações	Totais
Registros no arquivo	1.197
Registros sem ISSN	5
Registros considerados <i>outliers</i>	3
Registros sem informações nos indicadores	307
Registros para análise	885
Indicadores TODOS preenchidos	318
Registros do indicador (AIS) preenchido	679
Registros do indicador (FI) preenchido	782
Registros do indicador (HL) preenchido	726
Registros do indicador (SciNet) preenchido	460

Fonte: Elaborado pelos autores (2017).

Na sequência, discorre-se à análise descritiva sobre às variáveis, supra citadas, que compõem a base de dados em estudos, enfatizando-se às variáveis: Core, Avaliação Qualis 2012, AIS, FI e Half-Life (HL).

E.3 Distribuição conjunta de frequências

É muito frequente, nas Análises Descritivas, o interesse em verificar se duas variáveis qualitativas se apresentam associadas num certo conjunto de elementos. Uma distribuição conjunta de frequências das duas variáveis pode satisfazer a este objetivo. As distribuições conjuntas de frequências, geralmente, são apresentadas nas chamadas *tabelas cruzada* ou *tabelas de dupla entrada*.

Neste contexto, tem-se que para a construção da distribuição conjunta de frequências numa tabela cruzada, cada elemento deve pertencer a uma e apenas uma casela (cruzamento de uma linha com uma coluna). Nesta conjectura, na Tabela E.2, verifica-se que os totais das colunas formam a distribuição de frequência da variável *avaliação dos periódicos, quanto aos tipos de extratos, realizada pelo Qualis em 2012*, quando observada isoladamente, enquanto os totais das linhas constituem a distribuição da variável *Core*.

Tabela E.2: *Distribuição conjunta de frequências da avaliação dos periódicos, quanto aos tipos de extratos, realizada pelo Qualis em 2012 e o core (natureza das áreas dos periódicos).*

Core	Avaliação dos Periódicos								Total
	A1	A2	B1	B2	B3	B4	B5	C	
Estatística	16	27	26	17	5	6	5	0	102
Matemática	42	52	63	61	38	28	30	0	314
Matemática Aplicada	17	38	41	16	15	5	4	0	136
Outra área	19	49	111	54	41	30	25	1	330
Total	94	166	241	148	99	69	64	1	882

Fonte: Elaborado pelos autores (2017).

No que concerne as tabelas cruzada, tem-se que para facilitar a análise, pode-se incluir frequências relativas, que podem ser calculadas em relação aos totais das linhas ou colunas, dependendo do objetivo.

Neste contexto, na sequência, tem-se que a Tabela E.3 mostra a Tabela E.2, acrescida de porcentagens em relação aos totais das colunas.

Tabela E.3: Distribuição conjunta de frequências da avaliação dos periódicos, quanto aos tipos de extratos, realizada pelo Qualis em 2012 e o core (natureza das áreas dos periódicos) (%).

Core	Avaliação dos Periódicos							
	A1	A2	B1	B2	B3	B4	B5	C
Estatística	0.17	0.16	0.11	0.11	0.05	0.09	0.08	0.00
Matemática	0.45	0.31	0.26	0.41	0.38	0.41	0.47	0.00
Matemática Aplicada	0.18	0.23	0.17	0.11	0.15	0.07	0.06	0.00
Outra área	0.20	0.30	0.46	0.36	0.41	0.43	0.39	1.00
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Fonte: Elaborado pelos autores (2017).

Na Tabela E.3, acima, verifica-se que os dados da amostra parecem sugerir uma associação entre a avaliação dos periódicos e o core Matemática, pois, em todas as categorias da classificação por extrato, este core apresenta um nível de avaliação mais alto do que os demais. Neste contexto, enfatiza-se que, o core Estatística apresenta o segundo maior nível de associação no que diz respeito ao tipo de avaliação sugerida pela Qualis/2010.

Completando às informações expostas pela Tabela E.2, tem-se que uma análise estatística mais elaborada, como o Teste de Associação Qui-Quadrado, poderá detectar se esta associação é realmente válida para toda a base de dados em estudo.

Complementarmente às análises descritivas expostas, na sequência, discorre-se representações gráficas para às variáveis, já citadas neste relatório, que compõem a base de dados em estudos, enfatizando-se às variáveis: Core, Avaliação Qualis 2012, AIS, FI e Half-Life (HL).

E.4 Representações gráficas

As representações gráficas fornecem, em geral, uma visualização mais sugestiva do que as tabelas. Elas constituem-se numa forma alternativa de representação de distribuições de frequências.

Nesta seção, apresentaremos o gráfico de histograma, o gráfico de boxplot e o gráfico de setores, nesta ordem, que são particularmente importantes na representação de distribuições de frequências de dados categorizados.

E.4.1 Gráfico Histograma

Histograma é uma representação gráfica (um gráfico de barras verticais ou barras horizontais) da distribuição de frequências por meio de retângulos justapostos de um conjunto de dados quantitativos contínuos. O histograma pode ser um gráfico por valores absolutos ou frequência relativa ou densidade.

E.4.2 Gráfico Boxplot

Este tipo de gráfico estatístico representa os dados através de um retângulo construído com os quartis. Neste contexto, o objetivo deste gráfico é fornecer informações sobre a variabilidade dos dados e valores atípicos, “outliers”, que podem influenciar o cálculo de medidas como a média aritmética, por exemplo.

Corroborando com o exposto, tem-se que o boxplot utiliza cinco medidas estatísticas: mínimo, máximo, mediana, primeiro quartil, terceiro quartil. Em que, o conjunto destas medidas fornece evidência acerca da posição, dispersão, assimetria e valores extremos (atípicos).

E.4.3 Gráfico em Setores

É a representação gráfica de uma série estatística, em um círculo, por meio de setores. É utilizado principalmente quando se pretende comparar cada valor da série com o total. Para construí-lo, divide-se o círculo em setores, cujas áreas serão proporcionadas aos valores da série.

Complementarmente às definições gráficas expostas, segue-se, alguns gráficos referentes às variáveis: Core, Avaliação Qualis 2012, AIS, FI e HL. Vale ressaltar que, todos os gráficos, a seguir, foram gerados utilizando o pacote *ggplot2* (<http://ggplot2.org>) do *software* R.

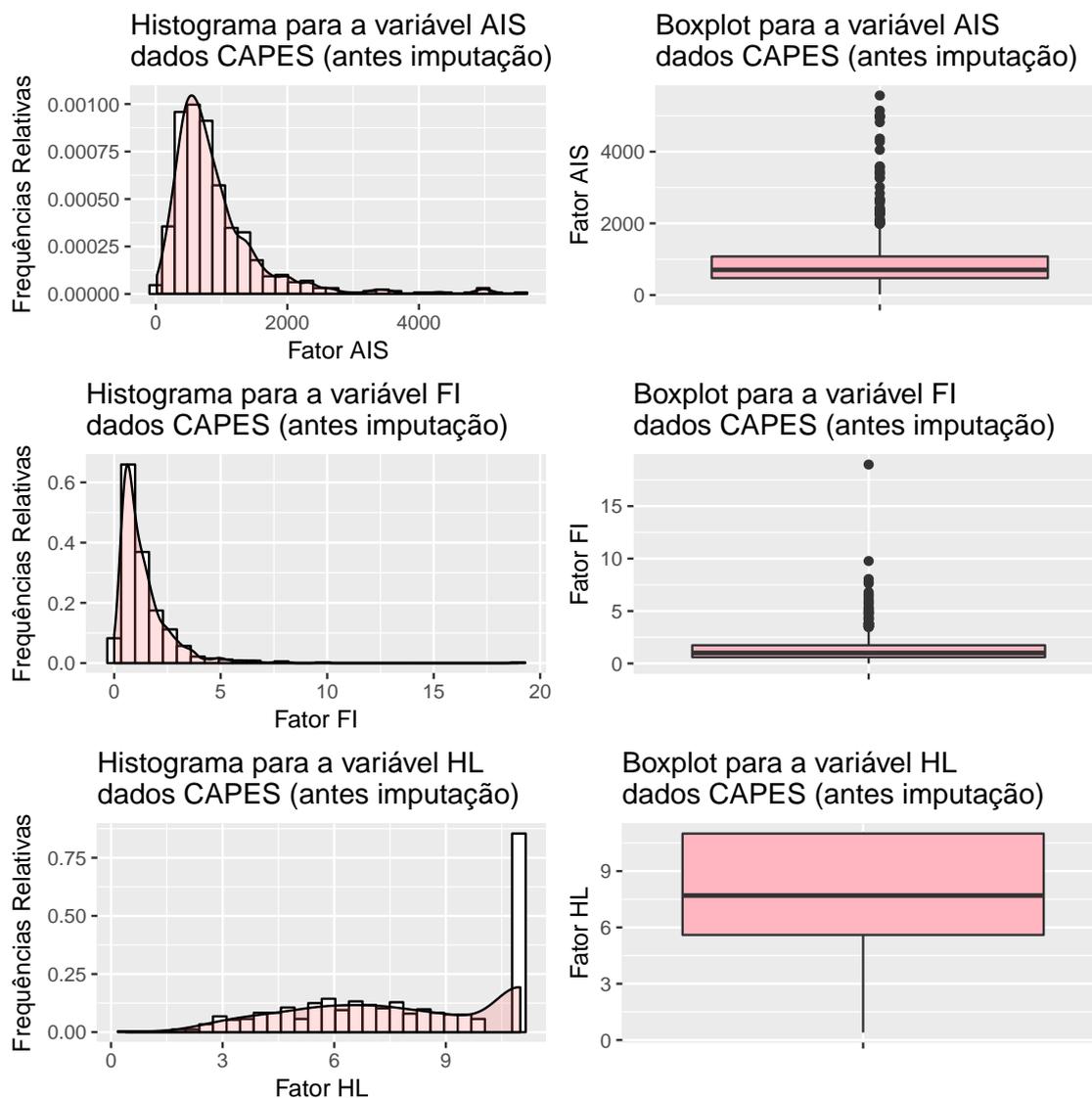


Figura E.1: Representações gráficas para os indicadores bibliométricos: Fator AIS, Fator FI e o Fator de Meia Vida.

Classificação dos periódicos por área, base de dados Qualis

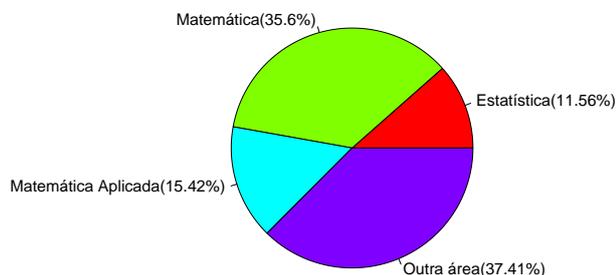


Figura E.2: Distribuição de frequências dos periódicos segundo a natureza das suas áreas, Qualis/CAPES.

Avaliação dos periódicos, quanto aos tipos de extratos, realizada pelo Qualis em 2012

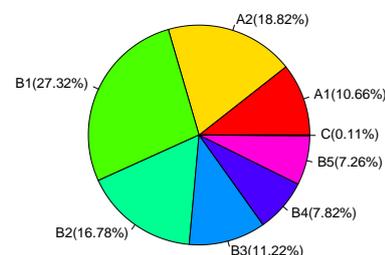


Figura E.3: Distribuição de frequências da avaliação dos periódicos divulgada pela Qualis/CAPES em 2012.

E.5 Web Of Science

Complementarmente às definições gráficas supra citadas, segue-se, alguns gráficos, Figuras E.4, E.5 e E.6, referentes às variáveis contínuas que compem a base de dados da Web Of Science: Fator de impacto de 5 anos, Fator de impacto, Fator de impacto sem as auto citações, Immediacy Index, Eigenfactor Score, Article Influence Score e Percentual de artigos em itens citáveis. Vale ressaltar que, todos os gráficos, a seguir, foram gerados utilizando o pacote *ggplot2* (<http://ggplot2.org>) do software R:

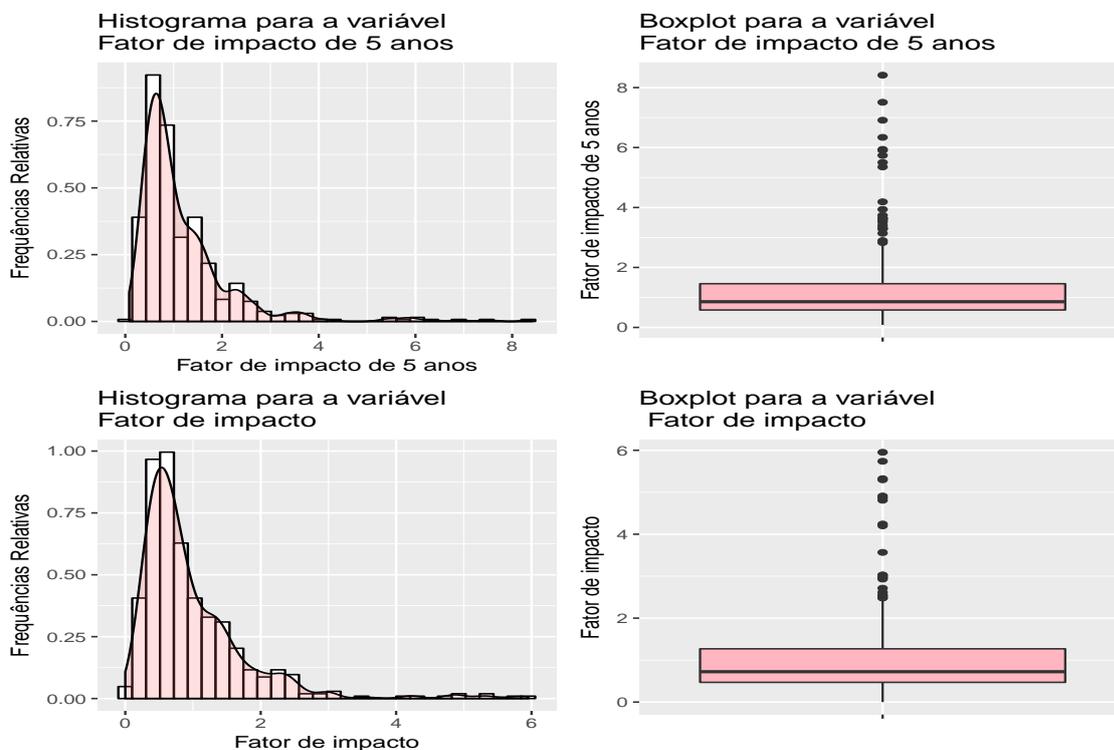


Figura E.4: Representações gráficas para os indicadores bibliométricos: Fator de impacto de 5 anos e Fator de impacto.

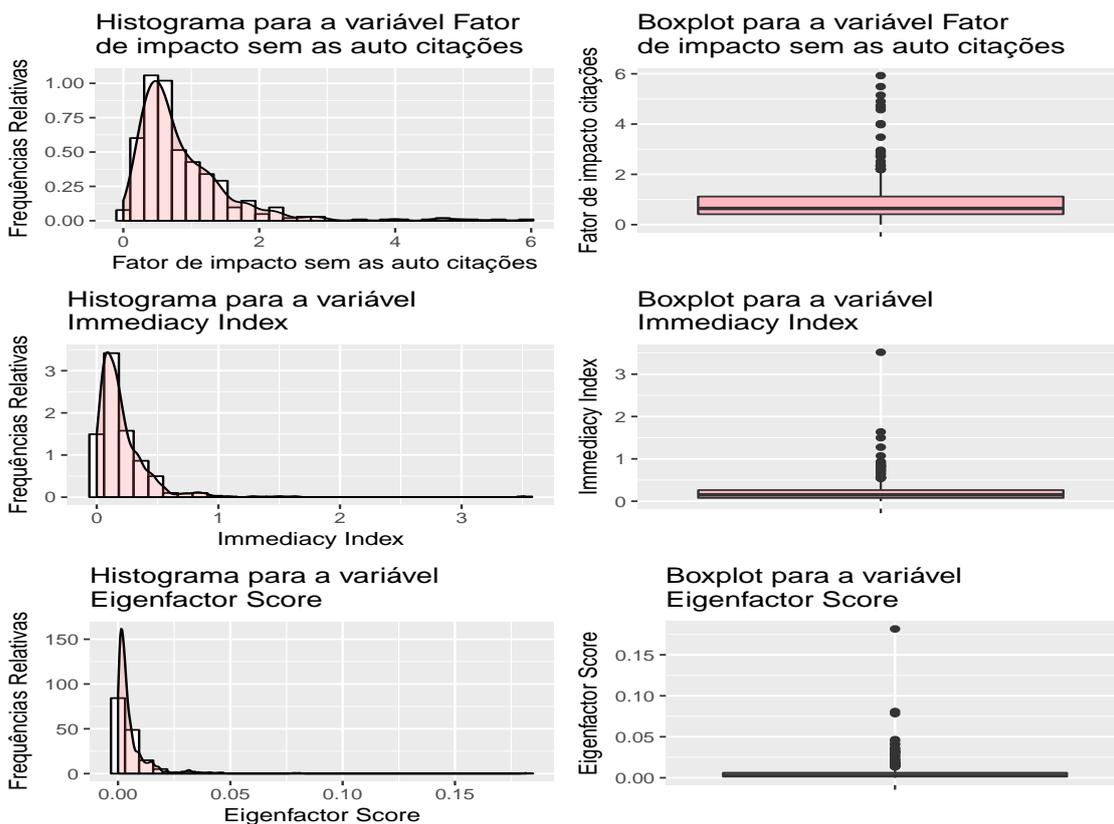


Figura E.5: Representações gráficas para os indicadores bibliométricos: Fator de impacto sem as auto citações, Immediacy Index e Eigenfactor Score.

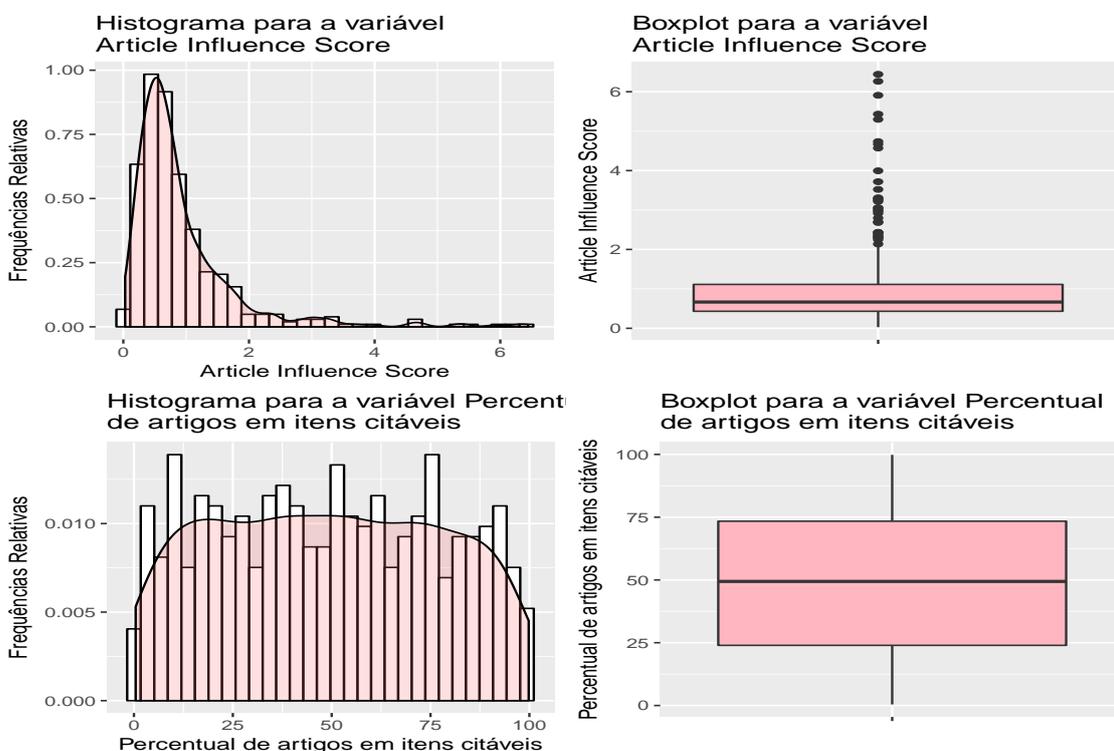


Figura E.6: Representações gráficas para os indicadores bibliométricos: Article Influence Score e Percentual de artigos em itens citáveis.

No que tange as áreas de interesse desta pesquisa, isto é, área no qual os periódicos

estão classificados, na Tabela E.4 de frequência, a seguir, tem-se algumas informações sobre os registros, por área de interesse, que formam a base de dados da Web Of Science, ano considerado 2012, composta por 504 periódicos:

Tabela E.4: *Representação tabular dos periódicos científicos por área para a base de dados da Web Of Science, ano considerado 2012.*

Área de interesse	Frequências
Matemática	190
Matemática Aplicada	136
Estatística	78
Outra área	100

Fonte: Elaborado pelos autores (2017).

Baseado na Tabela E.4, verifica-se que o número de periódicos classificados em Matemática representa a grande maioria com 190 dos jornais, a Matemática Aplicada 136, a Estatística com 78 e Outra área com 100 periódicos. Vale salientar que o core Outra área para a base de dados da Web Of Science 2012 é formada por Matemática Interdisciplinar, Biologia e Computação.

Analogamente, apresenta-se na sequência, algumas análises descritivas para as bases de dados do Scimago Journal & Country Rank, ano considerado 2012.

E.6 Scimago Journal & Country Rank

Segue-se, alguns gráficos, Figuras E.7, referentes a algumas das variáveis contínuas que compem a base de dados Scimago Journal & Country Rank: SJR e média de citações por documento no período de dois anos:

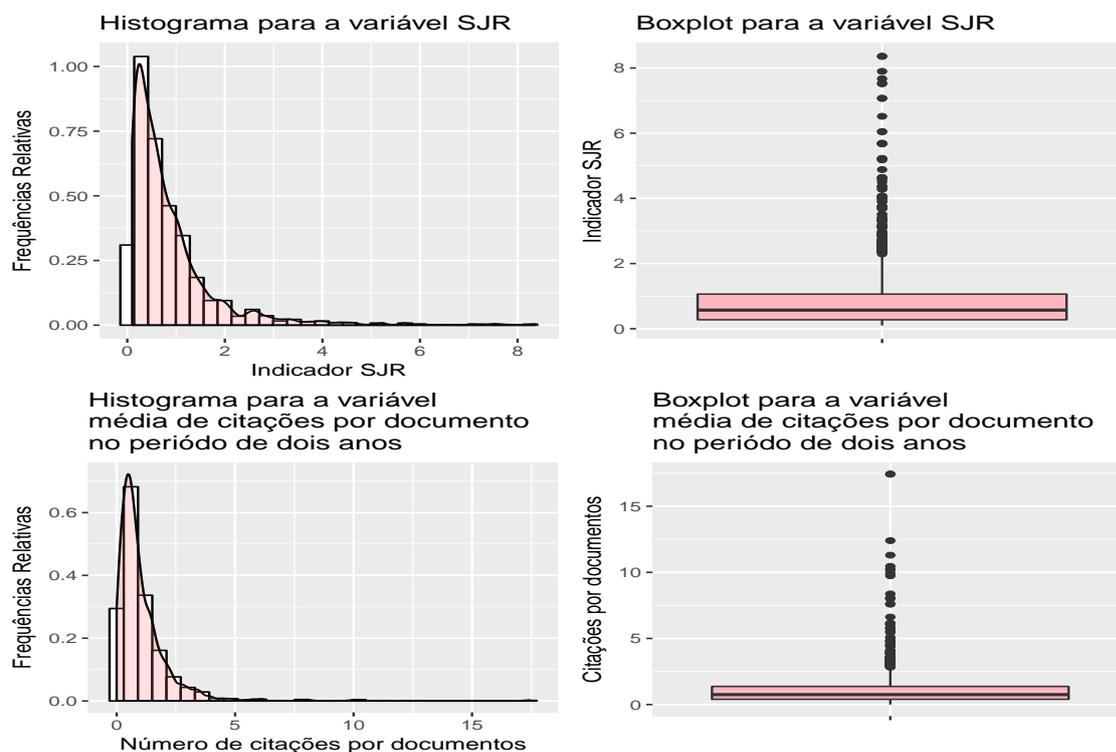


Figura E.7: *Representações gráficas para os indicadores bibliométricos: SJR e média de citações por documento no período de dois anos.*

Na Tabela E.5 de frequências, a seguir, tem-se algumas informações sobre os registros, por área de interesse, que formam a base de dados do Scimago Journal & Country Rank, considerando 2012 ano base:

Tabela E.5: *Representação tabular dos periódicos científicos por área para a base de dados do Scimago Journal & Country Rank, considerando 2012 ano base.*

Área de interesse	Frequências
Matemática	1032
Matemática Aplicada	516
Estatística	186
Outra área	-

Fonte: Elaborado pelos autores (2017).

Visualizando a Tabela E.5, verifica-se a existência de uma distância distoante entre o número de periódicos da Matemática e da Estatística, visto que a primeira área é composta por 1032 registros e a segunda por apenas 186 periódicos. Dentro do universo formador para a Matemática, tem-se um total de 59 registros duplicados. Para está base de dados não considerou-se “Outra área” ao baixar os registros do site (<http://www.scimagojr.com/>).

Apêndice F - Resultado das simulações para a calibração dos parâmetros dos itens

F.1 Resultados - Etapa I

Neste apêndice, apresenta-se os resultados das simulações para a calibração dos parâmetros dos itens via modelo politômico de Resposta Gradual de Samejima. No que tange a Etapa I, esta corresponde a fase da pesquisa em que a base de dados considerada é a fornecida pela CAPES constituída pelos seguintes indicadores bibliométricos: AIS, FI, HL e SciNet.

Vale ressaltar que o indicador SciNet, por falta de informações na literatura, este item não foi considerado nas análises, isto é, como um consoante na ordenação e classificação dos periódicos.

F.1.1 Calibração dos parâmetros dos itens para os indicadores AIS, FI, HL e SciNet - base sem imputação

Tabela F.1: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 1189$ periódicos considerando 8 categorias.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	8.66547	-5.47396	-0.88072	-0.45393	-0.10367	0.16223	0.48768	0.85764	1.34265
FI	2.60267	-4.98539	-1.25168	-0.69027	-0.27763	0.10015	0.47782	0.90385	1.49015
HL	0.52116	-10.52151	-3.66374	-2.03914	-0.87418	0.13799	1.11707	3.03433	3.9392
SciNet	3.17079	-4.69553	-1.19814	-0.73674	-0.41085	-0.0628	0.27788	0.69424	1.27503

Fonte: Elaborado pelos autores (2017).

Tabela F.2: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 882$ periódicos considerando 8 categorias.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	8.66547	-5.47396	-0.88072	-0.45393	-0.10367	0.16223	0.48768	0.85764	1.34265
FI	2.60267	-4.98539	-1.25168	-0.69027	-0.27763	0.10015	0.47782	0.90385	1.49015
HL	0.52116	-10.52151	-3.66374	-2.03914	-0.87418	0.13799	1.11707	3.03433	3.9392
SciNet	3.17079	-4.69553	-1.19814	-0.73674	-0.41085	-0.0628	0.27788	0.69424	1.27503

Fonte: Elaborado pelos autores (2017).

Tabela F.3: Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 1189$ periódicos considerando 8 categorias e agrupando as categorias 4 e 5 do item HL.

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	7.46584	-5.19632	-0.88989	-0.45791	-0.12051	0.15928	0.47994	0.84907	1.34114
FI	2.67754	-5.13256	-1.24582	-0.6897	-0.28032	0.09423	0.46887	0.89131	1.47321
HL	0.26854	-4.93715	-1.41622	1.54204	18.60126	-16.88731	-2.54303	6.28842	-
SciNet	3.24078	-7.38426	-1.19593	-0.73887	-0.41557	-0.07031	0.2673	0.68055	1.25804

Fonte: Elaborado pelos autores (2017).

Tabela F.4: Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 882$ periódicos considerando 8 categorias e agrupando as categorias 4 e 5 do item HL.

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	7.46584	-5.19632	-0.88989	-0.45791	-0.12051	0.15928	0.47994	0.84907	1.34114
FI	2.67754	-5.13256	-1.24582	-0.6897	-0.28032	0.09423	0.46887	0.89131	1.47321
HL	0.26854	-4.93715	-1.41622	1.54204	18.60126	-16.88731	2.54303	6.28842	-
SciNet	3.24078	-7.38426	-1.19593	-0.73887	-0.41557	-0.07031	.2673	0.68055	1.25804

Fonte: Elaborado pelos autores (2017).

Tabela F.5: Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 1189$ periódicos considerando a discretização pelos quintis.

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$
AIS	8.00411	-4.77787	-0.60207	-0.04995	0.43512	1.0228
FI	2.49672	-4.84943	-0.90571	-0.20051	0.40669	1.12772
HL	0.50495	-10.61424	-2.76525	-0.73281	0.93159	3.09959
SciNet	3.32856	-4.54825	-0.8433	-0.29976	0.19962	0.89707

Fonte: Elaborado pelos autores (2017).

Tabela F.6: Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 882$ periódicos considerando a discretização pelos quintis.

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$
AIS	8.00411	-4.77787	-0.60207	-0.04995	0.43512	1.0228
FI	2.49672	-4.84943	-0.90571	-0.20051	0.40669	1.12772
HL	0.50495	-10.61424	-2.76525	-0.73281	0.93159	3.09959
SciNet	3.32856	-4.54825	-0.8433	0.29976	0.19962	0.89707

Fonte: Elaborado pelos autores (2017).

Tabela F.7: Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 1189$ periódicos considerando a discretização pelos quintis e agrupando as categorias 5 e 6 do item HL.

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$
AIS	7.78885	-5.33447	-0.59912	-0.04512	0.44156	1.03074
FI	2.49535	-4.87342	-0.9005	-0.19348	0.41401	1.13554
HL	0.50458	-10.69142	-2.74419	-0.71774	0.94386	-
SciNet	3.35056	-4.52735	-0.83527	-0.29198	0.20746	0.905

Fonte: Elaborado pelos autores (2017).

Tabela F.8: Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 882$ periódicos considerando a discretização pelos quintis e agrupando as categorias 5 e 6 do item HL.

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$
AIS	7.78885	-5.33447	-0.59912	-0.04512	0.44156	1.03074
FI	2.49535	-4.87342	-0.9005	-0.19348	0.41401	1.13554
HL	0.50458	-10.69142	-2.74419	-0.71774	0.94386	-
SciNet	3.35056	-4.52735	-0.83527	-0.29198	0.20746	0.905

Fonte: Elaborado pelos autores (2017).

Na sequência, apresenta-se as simulações realizadas na base de dados sem o processo de imputação e, desta vez, considerando apenas três indicadores: AIS, FI e HL.

F.1.2 Calibração dos parâmetros dos itens para os indicadores AIS, FI e HL - base sem imputação

Tabela F.9: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, N = 1189 periódicos considerando 8 categorias sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	5.97095	-4.57182	-1.01805	-0.55278	-0.21749	0.07574	0.39386	0.76604	1.26149
FI	2.74516	-5.06772	-1.34293	-0.78392	-0.37393	-0.00063	0.37343	0.79689	1.37843
HL	0.51101	-10.72077	-3.82091	-2.17013	-0.98599	0.04321	1.03847	2.98865	3.91004

Fonte: Elaborado pelos autores (2017).

Tabela F.10: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, N = 882 periódicos considerando 8 categorias sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	5.97095	-4.57182	-1.01805	-0.55278	-0.21749	0.07574	0.39386	0.76604	1.26149
FI	2.74516	-5.06772	-1.34293	-0.78392	-0.37393	-0.00063	0.37343	0.79689	1.37843
HL	0.51101	-10.72077	-3.82091	-2.17013	-0.98599	0.04321	1.03847	2.98865	3.91004

Fonte: Elaborado pelos autores (2017).

Tabela F.11: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, N = 1189 periódicos considerando 8 categorias e agrupando as categorias 4 e 5 do item HL sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	4.0033	-4.97907	-1.07474	-0.58722	-0.23882	0.07316	0.39968	0.79105	1.31335
FI	3.38287	-5.02163	-1.28109	-0.75225	-0.36224	-0.00696	0.34854	0.74959	1.2998
HL	0.25719	-5.23469	-1.54126	1.54919	19.7916	-18.17776	-2.79984	6.43707	-

Fonte: Elaborado pelos autores (2017).

Tabela F.12: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, N = 882 periódicos considerando 8 categorias e agrupando as categorias 4 e 5 do item HL sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	4.0033	-4.97907	-1.07474	-0.58722	-0.23882	0.07316	0.39968	0.79105	1.31335
FI	3.38287	-5.02163	-1.28109	-0.75225	-0.36224	-0.00696	0.34854	0.74959	1.2998
HL	0.25719	-5.23469	-1.54126	1.54919	19.7916	-18.17776	-2.79984	6.43707	-

Fonte: Elaborado pelos autores (2017).

Tabela F.13: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, N = 1189 periódicos considerando a discretização pelos quintis sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$
AIS	5.69174	-5.19719	-0.7456	-0.16028	0.33539	0.95036
FI	2.6085	-4.93613	-0.9953	-0.29731	0.30373	1.02044
HL	0.51077	-10.59028	-2.81404	-0.80829	0.83739	2.98332

Fonte: Elaborado pelos autores (2017).

Tabela F.14: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 882$ periódicos considerando a discretização pelos quintis sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$
AIS	5.69174	-5.19719	-0.7456	-0.16028	0.33539	0.95036
FI	2.6085	-4.93613	-0.9953	-0.29731	0.30373	1.02044
HL	0.51077	-10.59028	-2.81404	-0.80829	0.83739	2.98332

Fonte: Elaborado pelos autores (2017).

Tabela F.15: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 1189$ periódicos considerando a discretização pelos quintis e agrupando as categorias 5 e 6 do item HL sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$
AIS	5.62909	-5.16322	-0.74583	-0.1595	0.33712	0.95318
FI	2.6148	-5.02442	-0.99348	-0.29526	0.30583	1.02207
HL	0.50472	-10.71956	-2.84967	-0.81931	0.84304	-

Fonte: Elaborado pelos autores (2017).

Tabela F.16: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados sem o processo de imputada, $N = 882$ periódicos considerando a discretização pelos quintis e agrupando as categorias 5 e 6 do item HL sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$
AIS	5.62909	-5.16322	-0.74583	-0.1595	0.33712	0.95318
FI	2.6148	-5.02442	-0.99348	-0.29526	0.30583	1.02207
HL	0.50472	-10.71956	-2.84967	-0.81931	0.84304	-

Fonte: Elaborado pelos autores (2017).

Na sequência, apresenta-se as simulações realizadas na base de dados após o processo de imputação via média para os três indicadores: AIS, FI e HL.

F.1.3 Calibração dos parâmetros dos itens para os indicadores AIS, FI e HL - base após imputação via média

Tabela F.17: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados após o processo de imputada via média, $N = 1189$ periódicos considerando 8 categorias sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	6.50087	-5.14233	-1.00243	-0.6127	-0.17295	0.13145	0.45701	0.79396	1.29463
FI	3.03443	-4.73521	-1.34826	-0.83591	-0.41338	-0.14391	0.37166	0.8053	1.36334
HL	0.85969	-8.96377	-2.41157	-1.35427	-0.59367	0.15366	0.74575	1.59454	2.59033

Fonte: Elaborado pelos autores (2017).

Tabela F.18: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados após o processo de imputada via média, $N = 1189$ periódicos considerando 8 categorias sem o indicador SciNet e agrupando as categorias 3 e 4, 6 e 7 do HL.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade							
		$(b_{i,1})$	$(b_{i,2})$	$(b_{i,3})$	$(b_{i,4})$	$(b_{i,5})$	$(b_{i,6})$	$(b_{i,7})$	$(b_{i,8})$
AIS	4.3756	-4.61694	-0.99085	-0.59095	-0.17344	0.13059	0.44981	0.79901	1.32107
FI	3.84697	-4.73489	-1.24803	-0.77015	-0.37868	-0.1274	0.34928	0.75213	1.27401
HL	0.24944	-5.44341	1.86878	20.89764	-19.25586	-2.15472	19.34489	-	-

Fonte: Elaborado pelos autores (2017).

Tabela F.19: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados após o processo de imputada via média, $N = 1189$ periódicos considerando a discretização pelos quintis sem o indicador SciNet.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		($b_{i,1}$)	($b_{i,2}$)	($b_{i,3}$)	($b_{i,4}$)	($b_{i,5}$)
AIS	8.24489	-5.40475	-0.49788	0.00239	0.42613	0.94789
FI	3.5572	-4.80639	-0.84572	-0.26072	0.33438	0.94732
HL	1.0901	-7.82669	-1.15788	-0.28212	0.5051	2.13445

Fonte: Elaborado pelos autores (2017).

Tabela F.20: *Estimativa dos parâmetros dos itens para algumas das variáveis da base de dados após o processo de imputada via média, $N = 1189$ periódicos considerando a discretização pelos quintis sem o indicador SciNet e agrupando as categorias 4 e 5 do HL.*

Item	Parâmetro de discriminação (a_i)	Parâmetros de Dificuldade				
		($b_{i,1}$)	($b_{i,2}$)	($b_{i,3}$)	($b_{i,4}$)	($b_{i,5}$)
AIS	4.69209	-5.68342	-0.5469	-0.03027	0.39863	0.96424
FI	4.59617	-4.96765	-0.82567	-0.27214	0.29956	0.89115
HL	0.38399	-4.84495	-1.42481	0.45827	13.30387	-

Fonte: Elaborado pelos autores (2017).

As dezenove tabelas apresentadas neste apêndice, correspondem aos resultados de simulações para a obtenção das estimativas dos parâmetros dos itens consideradas nesta etapa da pesquisa, bem como já apresentada no Capítulo 7 deste estudo de dissertação. Todas as simulações foram realizadas até se obter a Tabel 7.2, Capítulo 7.

Apêndice G - Dados históricos das distribuições dos periódicos por categoria de classificação da CAPES

G.1 Resultados e discussão da proposta - Etapa I

As análises que seguem neste apêndice, são análogas as apresentadas no Capítulo 7, Seção 7.4.

Assim, para ter maior comparabilidade com as classificações do Qualis/CAPES, uma possibilidade é utilizar os valores percentuais de 2012 dados pela Tabela H.1:

Tabela G.1: Representação tabular da classificação dos periódicos científicos por ano em termos percentuais (%) segundo o método apresentado pelo Qualis/CAPES.

Estrato de classificação	Classificação por ano em termos percentuais (%) Qualis/CAPES				Média
	2012	2013	2014	2015	
A1	7.205982	11.90053	11.36767	9.413854	9.972011
A2	13.9361	12.25577	12.07815	11.54529	12.453828
B1	20.05439	24.51155	23.80107	19.98224	22.087309
B2	13.32427	19.36057	23.44583	16.0746	18.051316
B3	10.94494	9.413854	8.525755	11.98934	10.195721
B4	12.44052	6.216696	6.039076	9.058615	8.438726
B5	19.57852	15.63055	13.85435	8.525755	14.397294
C	2.515296	0.71048	0.8881	13.4103	4.381044

Fonte: Elaborado pelos autores (2017).

G.1.1 Comparação do desempenho estimado no modelo da TRI versus estabelecido pela CAPES

Considerando as proporções de 2012, baseado na Tabela H.1, e seja $N = 1189$ número de registro da base de dados ordenada pela TRI, os próximos passos estabelecem critérios para propor uma nova classificação de periódicos científicos, alcançando assim, um dos objetivos desta pesquisa:

A1 TRI

$(1189 * 7.205982)/100 = 85.53501 \Rightarrow$ tomar os 85 periódicos de posto mais alto;

A2 TRI

$((7.205982 + 13.9361)/100) * 1189 = 251.3794 \Rightarrow$ considerar até o periódico de posto 251;
Os periódicos A2 são os que ocupam os postos 85 a 251. Logo, classificados como A2 tem-se 166 artigos.

B1 TRI

$((7.205982 + 13.9361 + 20.05439)/100) * 1189 = 489.8261 \Rightarrow$ considerar até o periódico de posto 490;
Ressalta-se que os periódicos B1 são os que ocupam os postos 252 a 490. Portanto, classificados como B1 tem-se 238 artigos.

B2 TRI

$((7.205982 + 13.9361 + 20.05439 + 13.32427)/100) * 1189 = 648.2516 \Rightarrow$ considerar até o periódico de posto 648;
Os periódicos B2 são os que ocupam os postos 491 a 648. Desta feita, classificados como B2 tem-se 157 artigos.

B3 TRI

$((7.205982 + 13.9361 + 20.05439 + 13.32427 + 10.94494)/100) * 1189 = 778.387 \Rightarrow$ considerar até o periódico de posto 778;
Os periódicos B3 são os que ocupam os postos 649 a 778. Portanto, classificados como B3 tem-se 129 artigos.

B4 TRI

$((7.205982 + 13.9361 + 20.05439 + 13.32427 + 10.94494 + 12.44052)/100) * 1189 = 926.3047 \Rightarrow$ considerar até o periódico de posto 925.
Os periódicos B4 são os que ocupam os postos 779 a 925. Assim, classificados como B4 tem-se 146 artigos.

B5 TRI

$((7.205982 + 13.9361 + 20.05439 + 13.32427 + 10.94494 + 12.44052 + 19.57852)/100) * 1189 = 1159.093 \Rightarrow$ considerar até o periódico de posto 1158;
Os periódicos B5 são os que ocupam os postos 926 a 1158. Portanto, classificados como B5 tem-se 232 artigos.

C TRI

todos os demais periódicos.

Tabela G.2: *Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área por área - método de classificação via TRI, considerando as proporções do ano de 2012, N = 1189 periódicos.*

Categoria	Número de periódicos por área de classificação				Total
	Estatística	Matemática	Mat. Apl.	Outra área	
A1 TRI	9	28	15	33	85
A2 TRI	17	60	12	77	166
B1 TRI	23	75	39	101	238
B2 TRI	16	86	26	31	159
B3 TRI	8	43	26	53	130
B4 TRI	12	27	8	101	148
B5 TRI	46	73	27	87	233
C TRI	3	17	10	0	30
				Total	1189

Fonte: Elaborado pelos autores (2017).

Tabela G.3: *Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo os critérios considerados pelo Qualis/-CAPES, ano base 2012.*

Categoria	Número de periódicos por área de classificação				Total
	Estatística	Matemática	Mat. Apl.	Outra área	
A1	16	44	18	19	97
A2	28	56	41	50	175
B1	26	66	45	111	248
B2	20	64	18	61	163
B3	9	40	18	60	127
B4	14	53	13	61	141
B5	20	82	10	104	216
C	1	4	0	17	22
				Total	1189

Fonte: Elaborado pelos autores (2017).

Tabela G.4: *Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área em valores percentuais por área - método de classificação via TRI, considerando as proporções do ano de 2012, N = 1189 periódicos.*

Categoria	Número de periódicos por área de classificação			
	Estatística	Matemática	Mat. Apl.	Outra área
A1 TRI	0.16	0.45	0.19	0.20
A2 TRI	0.16	0.32	0.23	0.29
B1 TRI	0.10	0.27	0.18	0.45
B2 TRI	0.12	0.39	0.11	0.37
B3 TRI	0.07	0.31	0.14	0.47
B4 TRI	0.10	0.38	0.09	0.43
B5 TRI	0.09	0.38	0.05	0.48
C TRI	0.05	0.18	0.00	0.77

Fonte: Elaborado pelos autores (2017).

Tabela G.5: Representação tabular da classificação Qualis/CAPES versus classificação proposta pela TRI, considerando as proporções do ano de 2012, $N = 1189$ periódicos.

Classificação CAPES	Proposta de classificação via TRI							
	A1 TRI	A2 TRI	B1 TRI	B2 TRI	B3 TRI	B4 TRI	B5 TRI	C TRI
A1	60	31	6	0	0	0	0	0
A2	20	75	67	13	0	0	0	0
B1	5	50	108	52	29	4	0	0
B2	0	7	49	37	43	17	10	0
B3	0	3	6	9	45	35	29	0
B4	0	0	1	48	12	2	71	7
B5	0	0	1	0	1	90	102	22
C	0	0	0	0	0	0	21	1

Fonte: Elaborado pelos autores (2017).

Tabela G.6: Representação tabular da distribuição percentual (%) por classificação Qualis/CAPES versus classificação proposta pela TRI, considerando as proporções do ano de 2012, $N = 1189$ periódicos.

Classificação CAPES	Proposta de classificação via TRI							
	A1 TRI	A2 TRI	B1 TRI	B2 TRI	B3 TRI	B4 TRI	B5 TRI	C TRI
A1	0.62	0.32	0.06	0.00	0.00	0.00	0.00	0.00
A2	0.11	0.43	0.38	0.07	0.00	0.00	0.00	0.00
B1	0.02	0.20	0.44	0.21	0.12	0.02	0.00	0.00
B2	0.00	0.04	0.30	0.23	0.26	0.10	0.06	0.00
B3	0.00	0.02	0.05	0.07	0.35	0.28	0.23	0.00
B4	0.00	0.00	0.01	0.34	0.09	0.01	0.50	0.05
B5	0.00	0.00	0.00	0.00	0.00	0.42	0.47	0.10
C	0.00	0.00	0.00	0.00	0.00	0.00	0.95	0.05

Fonte: Elaborado pelos autores (2017).

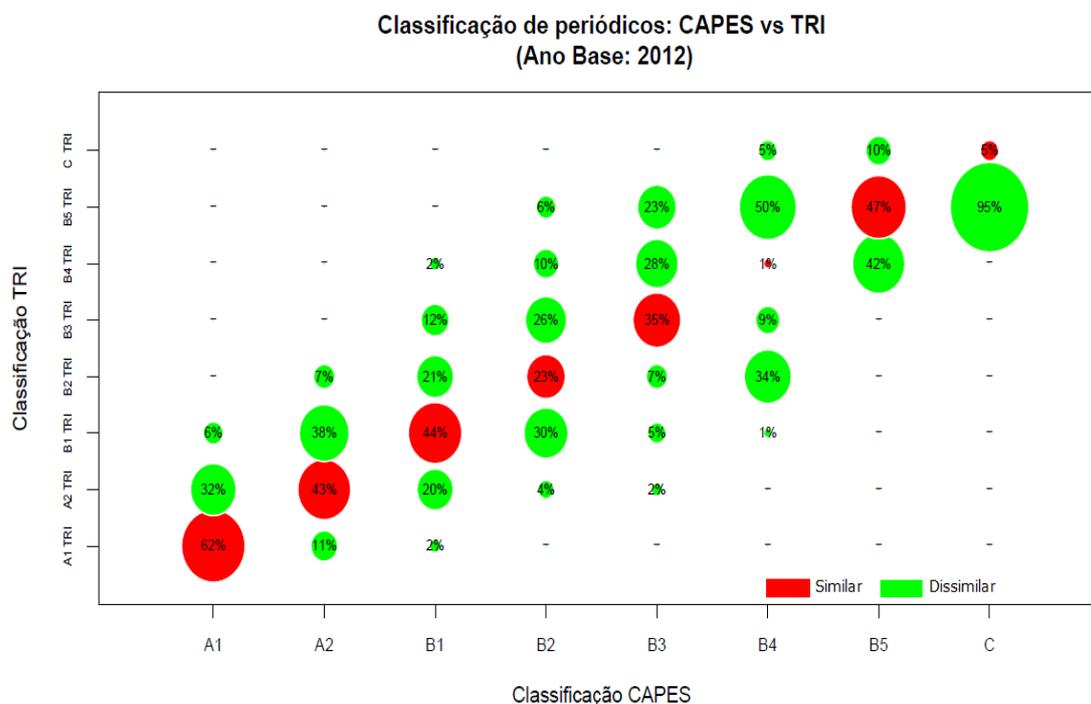


Figura G.1: Classificação de periódicos CAPES versus TRI, ano base média de 2012 a 2015, $N = 1189$ periódicos. Elaborado pelos autores (2017).

Tabela G.7: *Representação tabular parcial da classificação dos periódicos científicos segundo a proposta de classificação, desta pesquisa, via TRI para a Estatística versus classificação proposta pelo Qualis/CAPES, proporções do ano de 2012, N = 1189 periódicos.*

Ranking	Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
1	Biometrika	A1 TRI	A1
2	Annals of Statistics	A1 TRI	A1
3	Annals of Probability	A1 TRI	A1
4	Journal of the American Stat. Association	A1 TRI	A1
5	Probability Theory and Related Fields	A1 TRI	A1
6	Journal of Econometrics	A1 TRI	A2
7	Statistical Science	A1 TRI	A1
8	Statistics and Computing	A1 TRI	A1
9	Journal of the Royal Statistical Society	A1 TRI	A1
10	Finance and Stochastics	A2 TRI	A2

Fonte: Elaborado pelos autores (2017).

Tabela G.8: *Representação tabular parcial da classificação dos periódicos científicos segundo a proposta de classificação, desta pesquisa, via TRI para a Matemática versus classificação proposta pelo Qualis/CAPES, proporções do ano de 2012, N = 1189 periódicos.*

Ranking	Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
1	Acta Mathematica	A1 TRI	A1
2	Advances in Mathematics	A1 TRI	A1
3	American Journal of Mathematics	A1 TRI	A1
4	Transactions of the American Mat. S.	A1 TRI	A1
5	Annals of Mathematics	A1 TRI	A1
6	Archive for Rational Mechanics and An.	A1 TRI	A1
7	Commentarii Mathematici Helvetici	A1 TRI	A1
8	Compositio Mathematica	A1 TRI	A1
9	Duke Mathematical Journal	A1 TRI	A1
10	Annales Scientifiques	A1 TRI	A1

Fonte: Elaborado pelos autores (2017).

Tabela G.9: *Representação tabular parcial da classificação dos periódicos científicos segundo a proposta de classificação, desta pesquisa, via TRI para a Matemática Aplicada versus classificação proposta pelo Qualis/CAPES, proporções do ano de 2012, N = 1189 periódicos.*

Ranking	Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
1	Communications in Mathematical Ph.	A1 TRI	A1
2	Communications on Pure and Ap. Mat.	A1 TRI	A1
3	Journal of Computational Physics	A1 TRI	A1
4	Mathematical Programming	A1 TRI	A1
5	Numerische Mathematik	A1 TRI	A1
6	SIAM Journal on Mathematical Analysis	A1 TRI	A1
7	SIAM Journal on Numerical An.	A1 TRI	A1
8	SIAM Review	A1 TRI	A1
9	ACM Transactions on Mathematical S.	A1 TRI	A1
10	Mathematics of Operations Research	A1 TRI	A1

Fonte: Elaborado pelos autores (2017).

G.2 Algumas considerações

Visualizando as Tabelas G.7, G.8 e G.9 - comparação do desempenho estimado no modelo da TRI com o desempenho real - evidencia-se que o primeiro modelo apresenta alta similaridade ao classificar os periódicos em relação ao seu desempenho real. Ou seja, o modelo politômico está classificando bem.

Vale ressaltar que o modelo proposto possibilita o desenvolvimento e o aprimoramento constante em função da possibilidade de inserção de novas variáveis no modelo, com a manutenção da comparabilidade entre os periódicos, pela utilização da técnica de equalização. Esta refere-se à possibilidade de inserção de novas variáveis no modelo, que permite a evolução constante do modelo e a criação do conhecimento quanto ao construto da classificação dos periódicos.

Apêndice H - Comandos desenvolvidas em Visual *B*asic for Applications (VBA)

H.1 Programação em VBA

```
1 Sub QuickSortDesc(arr, Lo As Long, Hi As Long)
2 ' Regards, Hans Vogelaar (http://www.eileenslounge.com)
3
4     Dim varPivot As Variant
5     Dim varTmp As Variant
6     Dim tmpLow As Long
7     Dim tmpHi As Long
8     tmpLow = Lo
9     tmpHi = Hi
10    varPivot = arr((Lo + Hi) \ 2)
11    Do While tmpLow <= tmpHi
12        Do While arr(tmpLow) > varPivot And tmpLow < Hi
13            tmpLow = tmpLow + 1
14        Loop
15        Do While varPivot > arr(tmpHi) And tmpHi > Lo
16            tmpHi = tmpHi - 1
17        Loop
18        If tmpLow <= tmpHi Then
19            varTmp = arr(tmpLow)
20            arr(tmpLow) = arr(tmpHi)
21            arr(tmpHi) = varTmp
22            tmpLow = tmpLow + 1
23            tmpHi = tmpHi - 1
24        End If
25    Loop
26    If Lo < tmpHi Then QuickSortDesc arr, Lo, tmpHi
27    If tmpLow < Hi Then QuickSortDesc arr, tmpLow, Hi
28 End Sub
29
30 Function Pos_Transicao(vet_trans As Variant, str_trans As String) As
31     Integer
32     Dim tam, ret As Integer
33     Dim str_aux() As String
34
35     tam = UBound(vet_trans)
36     ret = -1
37
38     For ID = 1 To tam
39         vet_str_aux = Split(vet_trans(ID), ";")
40         If vet_str_aux(0) = str_trans Then
```

```

41         ret = ID
42         Exit For
43     End If
44 Next
45
46     Pos_Transicao = ret
47
48 End Function
49
50 Function LTrimEx(str)
51     Dim re
52     Set re = New RegExp
53     re.Pattern = "^\\s*"
54     re.Multiline = False
55     LTrimEx = re.Replace(str, "")
56 End Function
57
58 Function Acerta_PDF()
59
60 Dim fileXLS As Workbook
61 Set fileXLS = ActiveWorkbook
62 Dim Cabec() As String
63 Dim Transicao() As String
64 Dim Quebra As Variant
65 Dim Total_Registros, Total_ISSN, Total_ISSN_Sem_Info, Total_Transicao As
    Integer
66
67 Dim str_trans As String
68
69 'File_Externo = "C:\Users\Particular\Dropbox\SAM_JCGF\DADOS_PDF_EXCEL\
    Qualis_PDF_2012_Prob_Mat_Est.csv"
70 File_Externo = "C:\Users\SAMSUMG\Desktop\JCGF\Qualis_PDF_2012_Prob_Mat_Est
    .csv"
71 'File_Resumo = "C:\Users\Particular\Dropbox\SAM_JCGF\DADOS_PDF_EXCEL\
    Qualis_PDF_Resumo_2012_Prob_Mat_Est.txt"
72 File_Resumo = "C:\Users\SAMSUMG\Desktop\JCGF\
    Qualis_PDF_Resumo_2012_Prob_Mat_Est.txt"
73
74 Open File_Externo For Output As #1
75 Open File_Resumo For Output As #2
76
77 Total_ISSN = 0
78 Total_Registros = 0
79 Total_ISSN_Sem_Info = 0
80 Total_Transicao = 0
81 For id_plan = 1 To fileXLS.Sheets.Count
82     If Sheets(id_plan).Name <> "page 16" Then
83         Lin = 1
84         Col = 1
85         Do While (Len(Trim(Sheets(id_plan).Cells(Lin, 2)))) > 0)
86             Total_Registros = Total_Registros + 1
87             linha_out = ""
88             id_col = 1
89             If (id_plan = 1) And (Lin = 1) Then
90                 Do While (Len(Trim(Sheets(id_plan).Cells(Lin, id_col)))) > 0)
91                     ReDim Preserve Cabec(1 To id_col) As String
92                     Cabec(id_col) = Trim(Sheets(id_plan).Cells(Lin, id_col))
93                     If (id_col = 1) Or (id_col >= 5) Then
94                         If (id_col > 1) Then

```



```
148             End If
149         End If
150     End If
151     Lin = Lin + 1
152 Loop
153 End If
154
155 Next
156 Close #1
157 Print #2, "Total de registros: " & str(Total_Registros)
158 Print #2, "Total de ISSN: " & str(Total_ISSN)
159 Print #2, "Total de ISSNs sem informação: " & str(Total_ISSN_Sem_Info)
160 Print #2, "* _____ *"
161 Print #2, "*                TRANSIÇÕES                *"
162 Print #2, "* _____ *"
163 Print #2, "Qualis 2010" & Space(16 - Len("Qualis 2010")) & "Qualis 2012" &
    Space(15 - Len("Qualis 2012")) & "Ocorrências"
164 Dim Max As Long
165 Max = Int(UBound(Transicao))
166 QuickSortDesc Transicao, 1, Max
167 For ID = 1 To Max
168     Quebra = Split(Transicao(ID), ";")
169     ocorr = Int(Quebra(1))
170     Quebra = Split(Quebra(0), "/")
171     Print #2, Format(Quebra(0), "@") & Space(16 - Len(Quebra(0))) & Quebra
        (1) & Space(15 - Len(Quebra(1))) & Trim(str(ocorr))
172 Next
173 Close #2
174
175 End Function
```

Apêndice I - Sintaxes desenvolvidas no *software* Parscale, versão 4.1

I.1 Sintaxe Parscale - testando o efeito de DIF

```
1 CASO 1: TESTAR O EFEITO DE DIF
2
3 >FILES  DFNAME='DADOS.txt', SAVE;
4 >SAVE   PARM='GrupF1.PAR', SCORE='GrupF1.SCO';
5 >INPUT  NIDCHAR=9, MGROUP=4, NTOTAL=3;
6 (9A1,1X,1A1,1X,3A1)
7 >TEST   TNAME='grupos', ITEM=(1,2,3),
8         INAME=('I001', 'I002', 'I003'), NBLOCK=3;
9
10 TENTATIVA COM OS 3 COMANDOS ABAIXO NÃO FUNCIONA COM MODELO GRADED
11 BLOCK1  NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
        =(1,1,2,2,2,3,3,3,4), SCORE=(1,2,3,4);
12 BLOCK2  NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
        =(1,1,2,2,2,3,3,3,3), SCORE=(1,2,3);
13 BLOCK3  NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
        =(1,1,1,2,2,2,3,3,3), SCORE=(1,2,3);
14
15 TAMBEM NÃO FUNCIONA COM MODELO GRADED
16 BLOCK1  REPEAT=3, NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9);
17
18 NÃO FUNCIONA, MESMO QUANDO USO MODELO PARTIAL:
19 BLOCK1  REPEAT=3, NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9),MODIFIED
        =(1,1,2,2,3,3,3,4,4), SCORE=(1,2,3,4);
20
21 NÃO FUNCIONA, MESMO QUANDO USO MODELO PARTIAL:
22 BLOCK1  REPEAT=3, NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9),SCORE
        =(1,2,3,4,5,6,7,8,9);
23
24
25 NÃO FUNCIONA, MESMO QUANDO USO MODELO PARTIAL:
26 BLOCK1  REPEAT=3, NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
        =(1,1,2,2,2,3,3,3,3), SCORE=(1,2,3);
27
28 FUNCIONA QUANDO USO MODELO PARTIAL:
29 >BLOCK1  NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
        =(1,1,2,2,2,3,3,3,4), SCORE=(1,2,3,4);
30 >BLOCK2  NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
        =(1,1,2,2,2,3,3,3,3), SCORE=(1,2,3);
31 >BLOCK3  NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
        =(1,1,1,2,2,2,3,3,3), SCORE=(1,2,3);
32
```

```

33
34 >MGROUP DIF=(0,1,0,0), GNAME=('ESTATIST', 'MAT_APLC', 'MATEMATI', 'OUT-AREA'
    ),
35          GCODE=('1', '2', '3', '4');
36 >CALIB LOGISTIC, PARTIAL, NQPT=30, CYCLES=(100,1,1,1,1,1), NEWTON=50,
37          CRIT=0.01, POSTERIOR;
38 >SCORE ;

```

```

1 CASO 2: TESTAR O EFEITO DE DIF – MODELO PARTIAL CONSIDERARANDO GRUPOS

```

```

2
3 >FILES DFNAME='DADOS.txt', SAVE;
4 >SAVE PARM='GrupF1.PAR', SCORE='GrupF1.SCO';
5 >INPUT NIDCHAR=9, MGROUP=4, NTOTAL=3;
6 (9A1,1X,1A1,1X,3A1)
7 >TEST TNAME='grupos', ITEM=(1,2,3),
8       INAME=('I001', 'I002', 'I003'), NBLOCK=3;
9
10 TENTATIVA COM OS 3 COMANDOS ABAIXO NÃO FUNCIONA COM MODELO GRADED
11 BLOCK1 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
    =(1,1,2,2,2,3,3,3,4), SCORE=(1,2,3,4);
12 BLOCK2 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
    =(1,1,2,2,2,3,3,3,3), SCORE=(1,2,3);
13 BLOCK3 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
    =(1,1,1,2,2,2,3,3,3), SCORE=(1,2,3);

```

```

14
15 TAMBEM NÃO FUNCIONA COM MODELO GRADED
16 BLOCK1 REPEAT=3, NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9);

```

```

17
18 NÃO FUNCIONA, MESMO QUANDO USO MODELO PARTIAL:
19 BLOCK1 REPEAT=3, NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9),MODIFIED
    =(1,1,2,2,3,3,3,4,4), SCORE=(1,2,3,4);

```

```

20
21 NÃO FUNCIONA, MESMO QUANDO USO MODELO PARTIAL:
22 BLOCK1 REPEAT=3, NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9),SCORE
    =(1,2,3,4,5,6,7,8,9);

```

```

23
24
25 NÃO FUNCIONA, MESMO QUANDO USO MODELO PARTIAL:
26 BLOCK1 REPEAT=3, NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
    =(1,1,2,2,2,3,3,3,3), SCORE=(1,2,3);

```

```

27
28 FUNCIONA QUANDO USO MODELO PARTIAL:
29 >BLOCK1 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
    =(1,1,2,2,2,3,3,3,4), SCORE=(1,2,3,4);
30 >BLOCK2 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
    =(1,1,2,2,2,3,3,3,3), SCORE=(1,2,3);
31 >BLOCK3 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
    =(1,1,1,2,2,2,3,3,3), SCORE=(1,2,3);

```

```

32
33
34 >MGROUP DIF=(0,1,0,0), GNAME=('ESTATIST', 'MAT_APLC', 'MATEMATI', 'OUT-AREA'
    ),
35          GCODE=('1', '2', '3', '4');
36 >CALIB LOGISTIC, PARTIAL, NQPT=30, CYCLES=(100,1,1,1,1,1), NEWTON=50,
37          CRIT=0.01, POSTERIOR;
38 >SCORE ;

```

```

1 CASO 3: TESTAR O EFEITO DE DIF – MODELO PARTIAL SEM CONSIDERAR GRUPOS

```

```

2

```

```
3 >FILES  DFNAME='DADOS.txt', SAVE;
4 >SAVE   PARM='SEM-GRUPO.PAR', SCORE='SEM-GRUPO.SCO';
5 >INPUT  NIDCHAR=9, NTOTAL=3, NTEST=1, LENGTH=3;
6 (9A1,3X,3A1)
7 >TEST   TNAME='Sem grupos', ITEM=(1,2,3),
8         INAME=('I001', 'I002', 'I003'), NBLOCK=3;
9
10 >BLOCK1 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
      =(1,1,2,2,2,3,3,3,4), SCORE=(1,2,3,4);
11 >BLOCK2 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
      =(1,1,2,2,2,3,3,3,3), SCORE=(1,2,3);
12 >BLOCK3 NIT=1,NCAT=9, ORIGINAL=(1,2,3,4,5,6,7,8,9), MODIFIED
      =(1,1,1,2,2,2,3,3,3), SCORE=(1,2,3);
13
14 >CALIB  LOGISTIC, PARTIAL, NQPT=30, CYCLES=(100,1,1,1,1,1), NEWTON=50,
15        CRIT=0.01, POSTERIOR;
16 >SCORE  SMEAN=0.0, SSD=1.0;
```


Apêndice J - Apresentação de algumas árvores de decisão

J.1 Árvores de decisão e suas respectivas representações no espaço

J.1.1 Modelo $FI \sim AIS + HL + \varepsilon$

Os resultados apresentados neste apêndice são análogos aos apresentados no Capítulo 7, Seção 6.6 diferenciando apenas aos tipos de modelos considerados, a saber: $FI \sim AIS + HL + \varepsilon$ e $AIS \sim FI + \varepsilon$.

Na Figura J.1, a seguir, tem-se a representação de uma árvore de decisão onde o modelo tem como variável resposta o Fator de impacto (FI) e como covariáveis, o Fator AIS e o Fator de Meia Vida (HL). Matematicamente, este modelo é dado pela seguinte expressão:

$$AIS \sim FI + HL + \varepsilon. \tag{1}$$

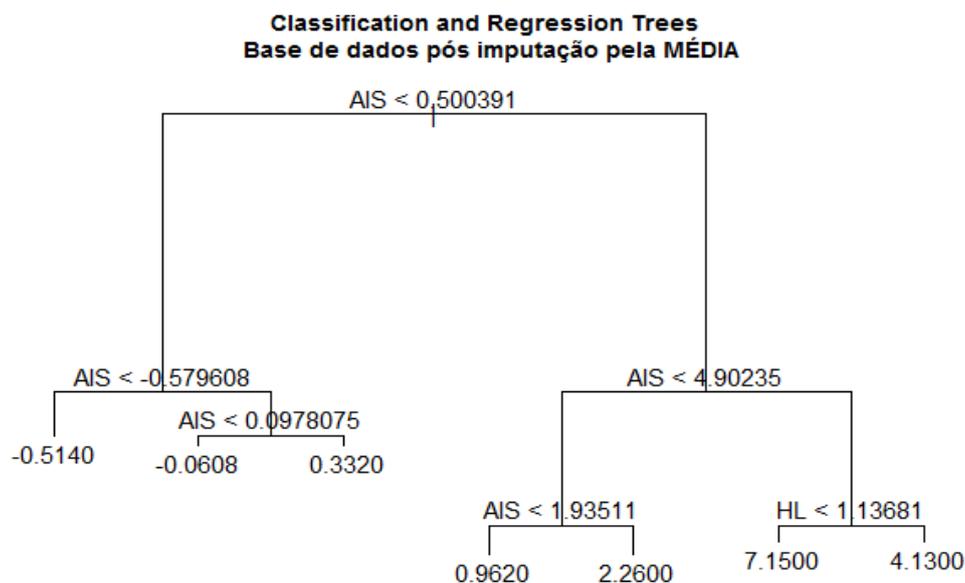


Figura J.1: Representação gráfica do CART para a base de dados após imputação pela média.

Na Tabela J.1 tem-se as regras que foram derivadas a partir da árvore de decisão/classificação apresentada na Figura J.1.

Antes da apresentação da Tabela J.1, apresenta-se na Figura J.2 a representação no espaço para a árvore de decisão/classificação apresentada em J.1:

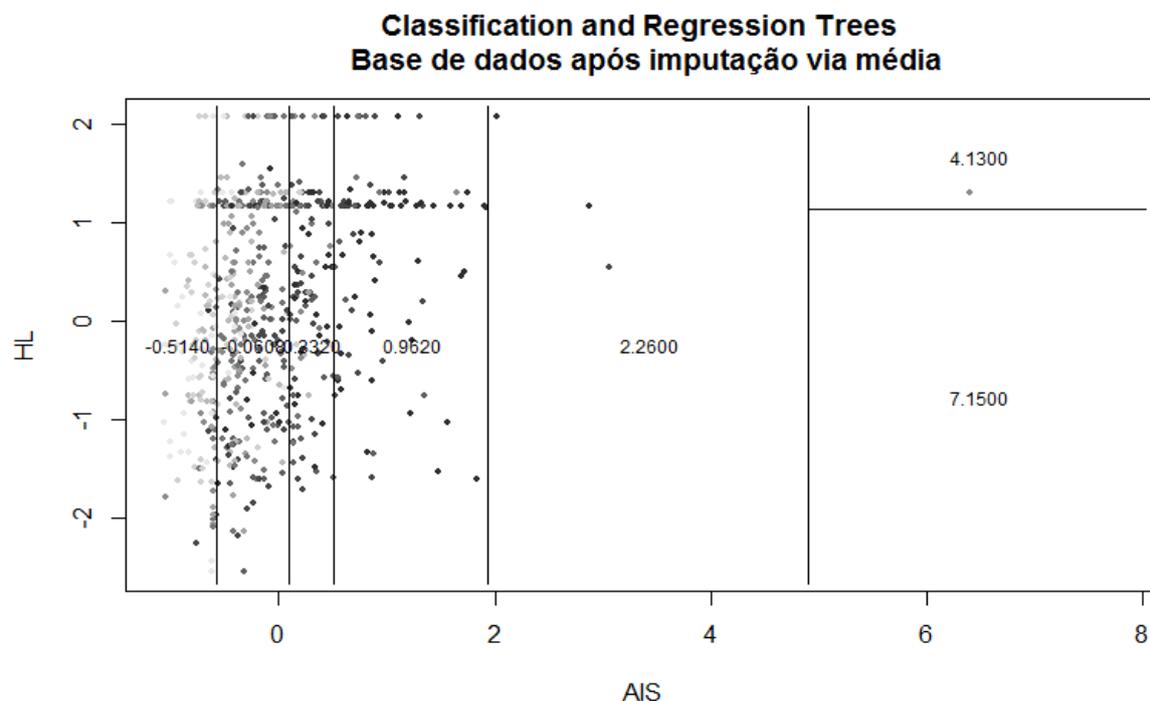


Figura J.2: Representação gráfica das partições do CART após imputação pela média.

Apresenta-se, na Tabela J.1, a classificação dos periódicos científicos, em valores absolutos, nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo o método “Classification and Regression Trees” com suas respectivas derivação de regras:

Tabela J.1: Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo o CART.

Categoria	Critério	FI médio	Número de periódicos por área			
			Mat.	Mat. Apl.	Est.	Out. área
1	$AIS \geq 4.900$ e $HL \leq 1.130$	7.150	2	0	0	3
2	$AIS \geq 4.900$ e $HL > 1.130$	4.130	2	2	0	1
3	$1.930 \leq AIS < 4.900$	2.260	6	3	5	16
4	$0.500 \leq AIS < 1.930$	0.962	40	19	17	72
5	$0.097 \leq AIS < 0.500$	0.332	36	23	12	54
6	$-0.570 \leq AIS < 0.097$	-0.060	139	70	36	148
7	$AIS < -0.570$	-0.514	98	36	29	172
Total			323	153	99	466

Fonte: Elaborado pelos autores (2017).

Na sequência, apresenta-se as Tabelas J.2 e J.3 com informações mais detalhas sobre a classificação dos periódicos científicos segundo o método “Classification and Regression Trees”.

Tabela KJ.2: *Representação tabular da distribuição percentual (%) por classificação dos periódicos científicos segundo o método “Classification and Regression Trees” apresentado na Tabela J.1.*

Área dos periódicos	Categorias de classificação						
	1	2	3	4	5	6	7
Estatística	0.0	0.0	0.17	0.11	0.10	0.09	0.09
Matemática	0.4	0.4	0.20	0.27	0.29	0.35	0.29
Matemática Aplicada	0.0	0.4	0.10	0.13	0.18	0.18	0.11
Outra área	0.6	0.2	0.53	0.49	0.43	0.38	0.51

Fonte: Elaborado pelos autores (2017).

Tabela J.3: *Representação tabular da distribuição percentual (%) por área dos periódicos científicos segundo o método “Classification and Regression Trees” apresentado na Tabela J.1.*

Área dos periódicos	Categorias de classificação						
	1	2	3	4	5	6	7
Estatística	0.00	0.00	0.05	0.17	0.12	0.36	0.29
Matemática	0.01	0.01	0.02	0.12	0.11	0.43	0.30
Matemática Aplicada	0.00	0.01	0.02	0.12	0.15	0.46	0.24
Outra área	0.01	0.00	0.03	0.15	0.12	0.32	0.37

Fonte: Elaborado pelos autores (2017).

Enfatiza-se nas Tabelas J.4, J.5 e J.6, a seguir, em ordem decrescente, apenas os 10 primeiros periódicos de maior categoria entre os 1189 que compõem a base de dados a partir da classificação proposta na Tabela K.1 para as áreas da Matemática, Matemática Aplicada e Estatística, respectivamente:

Tabela J.4: *Representação tabular parcial da classificação dos periódicos científicos segundo o método “Classification and Regression Trees” para a Matemática versus classificação proposta pelo Qualis/CAPES.*

Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
Anais da Academia Bra. de Ciências	6	B2
Acta Mathematica	2	A1
Advances in Mathematics	4	A1
Aequationes Mathematicae	5	A2
Algebra and Logic	6	B2
American Journal of Mathematics	4	A1
Proceedings of the Am. Mat. Soc.	6	A2
Transactions of the Am. Mat. Soc.	4	A1
Annals of Mathematics	2	A1
Applicable Analysis	6	B1

Fonte: Elaborado pelos autores (2017).

Tabela J.5: Representação tabular parcial da classificação dos periódicos científicos segundo o método “Classification and Regression Trees” para a Matemática Aplicada versus classificação proposta pelo Qualis/CAPES.

Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
Journal of Computational Phy.	4	A1
Communications on Pure App. Mat.	3	A1
Computer Aided Design	6	A2
Functional Analysis and App.	6	B1
International Journal of Comp. Mat.	7	B3
Journal of Computational Phy.	4	A1
Journal of Engineering Mat.	6	B1
Studies in Applied Mathematics	6	A2
Journal of Optimization Th. App.	6	A2
Mathematical Biosciences	6	A2

Fonte: Elaborado pelos autores (2017).

Tabela J.6: Representação tabular parcial da classificação dos periódicos científicos segundo o método “Classification and Regression Trees” para a Estatística versus classificação proposta pelo Qualis/CAPES.

Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
Advances in Applied Probability	6	A2
The American Statistician	6	A1
Biometrics (Washington)	4	A1
Biometrika (London)	4	A1
nnals of the Institute of Statis. Mat.	6	A2
International Economic Review)	6	B1
Journal of Applied Probability	6	B1
Journal of Statistical Physics	6	A2
Applied Statistics	6	A2
Statistica Neerlandica	7	B2

Fonte: Elaborado pelos autores (2017).

A classificação dos periódicos no modelo CART, conforme as Tabelas J.4, J.5 e J.6, segue padrões distintos a medida que altera as acategorias de classificação em uma análise feita nas linhas da tabela.

J.1.2 Modelo $AIS \sim FI + \varepsilon$

Analogamente ao apresentado na Seção J.1.1, considera-se nesta seção o modelo $AIS \sim FI + \varepsilon$.

Na Figura J.3, a seguir, tem-se a representação de uma árvore de decisão onde o modelo tem como variável resposta o Fator AIS e como covariável, apenas, o Fator de Impacto (FI). Matematicamente, este modelo é dado pela seguinte expressão:

$$AIS \sim FI + \varepsilon. \quad (2)$$

Na Tabela J.7 tem-se as regras que foram derivadas a partir da árvore de decisão/classificação apresentada na Figura J.3.

Portanto, apresenta-se, na Tabela J.7, a classificação dos periódicos científicos, em valores absolutos, nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo o método “Classification and Regression Trees” com suas respectivas derivação de regras:

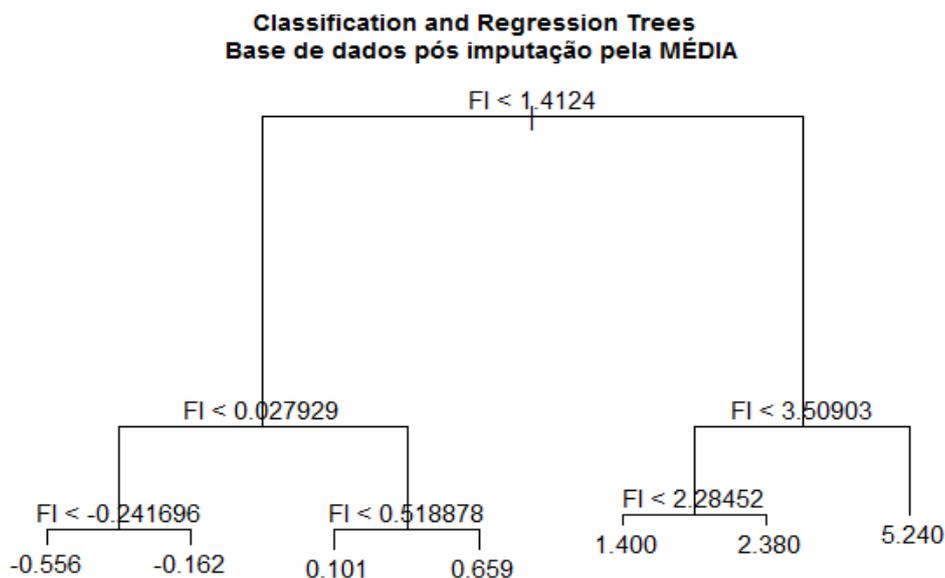


Figura J.3: Representação gráfica do CART para a base de dados após imputação pela média.

Tabela J.7: Classificação dos periódicos científicos nas áreas da Matemática, Matemática Aplicada, Estatística e Outra área segundo o método “Classification and Regression Trees”.

Categoria	Critério	AIS médio	Número de periódicos por área			
			Mat.	Mat. Apl.	Est.	Out. área
1	$FI \geq 3.509$	5.240	5	2	1	4
2	$2.284 \leq FI < 3.509$	2.380	5	1	4	12
3	$1.420 \leq FI < 2.284$	1.400	5	5	8	18
4	$0.518 \leq FI < 1.412$	0.659	47	23	16	66
5	$0.027 \leq FI < 0.518$	0.101	71	27	20	66
6	$-0.241 \leq FI < 0.027$	-0.162	79	22	17	56
7	$FI < 0.027$	-0.556	111	73	33	244
Total			323	153	99	466

Fonte: Elaborado pelos autores (2017).

Na sequência, apresenta-se as Tabelas J.8 e J.9 com informações mais detalhas sobre a classificação dos periódicos científicos segundo o método “Classification and Regression Trees”.

Tabela J.8: Representação tabular da distribuição percentual (%) por classificação dos periódicos científicos segundo o método “Classification and Regression Trees” apresentado na Tabela J.7.

Área dos periódicos	Categorias de classificação						
	1	2	3	4	5	6	7
Estatística	0.08	0.18	0.22	0.11	0.11	0.10	0.07
Matemática	0.42	0.23	0.14	0.31	0.39	0.45	0.24
Matemática Aplicada	0.17	0.05	0.14	0.15	0.15	0.13	0.16
Outra área	0.33	0.55	0.50	0.43	0.36	0.32	0.53

Fonte: Elaborado pelos autores (2017).

Enfatiza-se nas Tabelas J.10, J.11 e J.12, a seguir, segundo a proposta de classificação/-

Tabela J.9: Representação tabular da distribuição percentual (%) por área dos periódicos científicos segundo o método “Classification and Regression Trees” apresentado na Tabela J.7.

Área dos periódicos	Categorias de classificação						
	1	2	3	4	5	6	7
Estatística	0.01	0.04	0.08	0.16	0.20	0.17	0.33
Matemática	0.02	0.02	0.02	0.15	0.22	0.24	0.34
Matemática Aplicada	0.01	0.01	0.03	0.15	0.18	0.14	0.48
Outra área	0.01	0.03	0.04	0.14	0.14	0.12	0.52

Fonte: Elaborado pelos autores (2017).

categorias via CART, apenas os 10 primeiros periódicos de maior categoria entre os 1189 que compõem a base de dados a partir da classificação proposta na Tabela J.7 para as áreas da Matemática, Matemática Aplicada e Estatística, respectivamente:

Tabela J.10: Representação tabular parcial da classificação dos periódicos científicos segundo o método “Classification and Regression Trees” para a Matemática versus classificação proposta pelo Qualis/CAPES.

Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
Anais da Academia Bra. de Ciências	5	B2
Acta Mathematica	1	A1
Advances in Mathematics	4	A1
Aequationes Mathematicae	5	A2
Algebra and Logic	6	B2
American Journal of Mathematics	4	A1
Proceedings of the Am. Mat. Soc.	6	A2
Transactions of the Am. Mat. Soc.	4	A1
Annals of Mathematics	1	A1
Applicable Analysis	5	B1

Fonte: Elaborado pelos autores (2017).

Tabela J.11: Representação tabular parcial da classificação dos periódicos científicos segundo o método “Classification and Regression Trees” para a Matemática Aplicada versus classificação proposta pelo Qualis/CAPES.

Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
Journal of Computational Phy.	4	A1
Communications on Pure App. Mat.	3	A1
Computer Aided Design	5	A2
Functional Analysis and App.	7	B1
International Journal of Comp. Mat.	7	B3
Journal of Computational Phy.	3	A1
Journal of Engineering Mat.	7	B1
Studies in Applied Mathematics	7	A2
Journal of Optimization Th. App.	6	A2
Mathematical Biosciences	4	A2

Fonte: Elaborado pelos autores (2017).

Tabela J.12: Representação tabular parcial da classificação dos periódicos científicos segundo o método “Classification and Regression Trees” para a Estatística versus classificação proposta pelo Qualis/CAPES.

Periódico	Proposta de classificação/categorias	Classificação Qualis/CAPES
Advances in Applied Probability	6	A2
The American Statistician	5	A1
Biometrics (Washington)	4	A1
Biometrika (London)	3	A1
Annals of the Institute of Statist. Mat.	5	A2
International Economic Review)	4	B1
Journal of Applied Probability	6	B1
Journal of Statistical Physics	4	A2
Applied Statistics	7	A2
Statistica Neerlandica	7	B2

Fonte: Elaborado pelos autores (2017).

Analogamente ao apresentado nas Tabelas J.4, J.5 e J.6, complementa-se que a classificação dos periódicos via modelo CART, conforme as Tabelas J.10, J.11 e J.12, segue padrões distintos a medida que altera as acategorias de classificação em uma análise feita nas linhas da tabela.

Na sequência, apresenta-se algumas medidas resumo para os modelos de árvore de decisão considerados nesta pesquisa.

J.2 Algumas medidas resumo para os modelos: $AIS \sim FI + HL + \varepsilon$, $FI \sim AIS + HL + \varepsilon$ e $AIS \sim FI + \varepsilon$

Na sequência, discorre algumas análises descritivas destas para os modelos: $AIS \sim FI + HL + \varepsilon$, $FI \sim AIS + HL + \varepsilon$ e $AIS \sim FI + \varepsilon$:

Tabela J.13: *Algumas medidas resumo relativa aos modelos gerados a partir da árvore de decisão/classificação (CART).*

Modelo	Medidas resumo						
	Min.	1st. Qu.	Mediana	Média	3rd. Qu.	Max.	D_i
$AIS \sim FI + HL + \varepsilon$	-3.95500	-0.21130	-0.05464	0.00000	0.14850	6.56000	0.2873
$FI \sim AIS + HL + \varepsilon$	-4.37200	-0.25530	-0.04135	0.00000	0.24520	5.41800	0.3264
$AIS \sim FI + \varepsilon$	-3.95500	-0.21070	-0.05464	0.00000	0.14850	6.56000	0.2874

Fonte: Elaborado pelos autores (2017).

Como já apresentado no Capítulo 6.6, em um contexto de classificação, tem-se que o CART é utilizado como uma forma de alocar n indivíduos em classes arbitrárias. Logo, as árvores fornecem modelos de probabilidade, de forma individual, para cada membro, no caso em estudo, membros equivalem as categorias.

De acordo com Sutton (2005), D_i , dado na Tabela J.13, representa o desvio médio residual obtido pelo CART. Assim, o autor enfatiza que:

$$D_i = -2 \sum_k n_{ik} \ln p_{ik}, \quad (3)$$

onde i representa cada nó da árvore, p_{ik} a distribuição de probabilidade em cada classe, isto é, categoria e n_{ik} as folhas da árvore. Vale ressaltar que essas folhas fornecem uma amostra aleatória de uma distribuição multinomial especificada por p_{ik} .

Assim, o desvio de uma árvore, D_i , expressado matematicamente em (3), corresponde a soma sobre todas as folhas de uma árvore.

Visualizando a Tabela J.13 em que apresenta as medidas resumo relacionadas aos modelos obtidos a partir do CART e seus respectivos desvio médio residual. Nota-se um pequeno desempenho superior do modelo $AIS \sim FI + HL + \varepsilon$ quando comparado com o modelo $AIS \sim FI + \varepsilon$, $D_1 = 0.2873$ e $D_3 = 0.2874$, respectivamente. Todos os modelos apresentam diferenças mínimas em suas medidas. No entanto, ainda comparandos os desvios médios, nota-se que o modelo $FI \sim AIS + HL + \varepsilon$ apresenta um alto desvio, $D_2 = 0.3264$ quando comprado com os demais modelos.

Apêndice K - Pareamento das bases de dados desta pesquisa

Às bases de dados desta pesquisa, construídas, a partir do Journal Indicators CWTS, o Scimago e a Web Of Science, disponibilizam indicadores, que utilizam diferentes métricas, para a classificação de periódicos, nas mais diversas áreas. Tais métricas são importantes na avaliação da qualidade de publicações científicas.

Como as referidas três bases de dados disponibilizam distintos indicadores da qualidade dos periódicos, com cada um destes indicadores explorando algum aspecto distinto utilizado para a ordenação dos jornais, objetiva-se propor um novo indicador da qualidade de periódicos, baseado em uma análise conjunta, via Teoria da Resposta ao Item (TRI), dos referidos indicadores, nas diferentes bases. Para tanto, necessita-se criar uma base de dados comum que congregue todos os indicadores disponíveis. Desta forma, faz-se necessário criar mecanismos de pareamento dos registros das informações de um mesmo periódico nas diferentes bases. No entanto, tal tarefa é bastante desafiadora devido a inúmeras inconsistências das bases e limitações dos softwares disponíveis. Desta feita, o uso da linguagem *Python* (vide **Apêndice A**) na elaboração dos códigos computacionais necessários revelou-se uma ferramenta extremamente eficaz para esse tipo de problema nos pareamentos empreendidos fêz-se necessário o uso de medidas de similaridade entre strings.

Neste trabalho, técnicas de pareamentos determinísticos foram aplicadas no intuito de compor uma base consistente, utilizando-se como chave de pareamento entre as bases, ora o código identificador de periódicos ISSN (pareamento exato), ora o título dos periódicos (pareamento determinístico, utilizando medidas de similaridade).

O presente capítulo apresenta uma discussão referente às bases de dados desta pesquisa, o conceito de pareamento determinístico, as medidas de similaridade empregadas nos pareamentos.

Segundo Gill (2001), existem três tipos distintos de pareamentos de dados: pareamento determinístico, que é o foco desta pesquisa e que será apresentado em detalhes na Seção 9.1, pareamento probabilístico e o pareamento assistido.

No que concerne ao pareamento probabilístico, este é baseado no cálculo de pesos de similaridade entre os registros de duas bases de dados. São considerados iguais os registros que apresentam um peso ou soma de pesos maior que determinado ponto de corte. Já no pareamento assistido, conta-se com o auxílio de um indivíduo na inspeção visual dos registros, nas duas bases de dados.

Na sequência, discorre-se sobre algumas medidas de similaridade. Destacam-se os trabalhos de Ratcliff e Metzner (1988), que propuseram o algoritmo Ratcliff/Obershelp e o de Winkler (1999), que apresentou uma medida de similaridade entre duas strings, que estende a métrica de similaridade proposta por Jaro (1989). Esses algoritmos são bastante utilizados na área de Data Mining e, neste trabalho, fazemos extenso uso dos mesmos. Nas próximas seções apresentam-se vários conceitos de medidas de similaridade e uma aplicação às bases de dados em estudo.

K.1 Conceito de Pareamento Determinístico

No que se refere aos tipos de pareamentos, Gill (2001) destaca que o pareamento determinístico, também conhecido como pareamento exato ou pareamento “tudo ou nada”, consiste em selecionar variáveis comuns a duas bases de dados distintas e criar um modelo de comparação baseado em igualdades nessas variáveis para apontar os registros iguais que estão em ambas as bases de dados.

Logo, o termo *pareamento* decorre da formação de um par após a implementação do modelo de comparação, isto é, um par é formado contendo o registro da primeira base de dados e o registro apontado como igual a ele da segunda base de dados.

Portanto, é verificado que o resultado do modelo de comparação do pareamento determinístico é claro, isto é, ou os registros são classificados como iguais, gerando um par, ou são classificados como diferentes e o pareamento não é estabelecido. Esta clareza simplifica o processo e torna-o mais prático e mais rápido na sua execução.

O desafio do pareamento determinístico é criar o modelo de comparação adequado para realizar esta classificação de registros iguais e diferentes. Para isso, a escolha das variáveis deve ser feita com cuidado. O melhor modelo de comparação é aquele que pareia o maior número possível de pares verdadeiros com o menor número de pareamentos incorretos. Quando ocorre o pareamento de dois registros que não são iguais, ocorre um erro denominado *falso positivo* e quando não ocorre o pareamento de dois registros iguais ocorre um *falso negativo* (Zealand, 2013).

Em se tratando do pareamento determinístico, vale ressaltar que essa técnica não aponta como iguais somente registros que possuem concordância perfeita nos valores das variáveis. Winkler (1999) descreve a construção de medidas de similaridade que incluem uma ponderação, de modo a refletir a plausibilidade do estabelecimento de um par, entre registros com pequenas diferenças de grafia. Tal procedimento visa diminuir a ocorrência de falsos negativos (pareamento correto descartado como tal). Essas medidas de similaridade, podem ter associados um coeficiente que varia entre 0 (zero) e 1 (um). Quanto mais próximo de 1 (um) mais similares são considerados os registros em termos de grafia.

Dentre os estudos de similaridade, a verificação ortográfica é uma característica importante a ser considerada pelos *softwares* modernos. A verificação ortográfica é amplamente utilizada em programas de computador, tais como processadores de texto, dicionários e mecanismos de busca em geral.

Após a discussão do conceito de pareamento determinístico, baseando-se na literatura, apresenta-se, na seção seguinte, algumas das medidas de similaridades mais usuais.

K.2 Medidas de similaridade

Nesta seção, apresenta-se algumas noções e definições de medidas de similaridade sintática entre strings ou documentos. Neste contexto, estas medidas são utilizadas no pareamento de bases de dados. Dentre estas, a medida de similaridade, D_{ro} , elaborada por Ratcliff e Metzner (1988). Para esta medida, utiliza-se a biblioteca **diffib** presente no código *Python*, que foi desenvolvido nesta dissertação, com a finalidade de estabelecer pareamentos entre uma ou mais bases de dados.

Complementando o tema em questão, Baeza-Yates e Ribeiro-Neto (2013, p. 209) ressaltam que a medida de similaridade é medida por uma *função de distância*. Por exemplo, para strings de mesmo tamanho, pode-se definir a distância entre elas como sendo o número de posições que possuem caracteres diferentes. Então, a distância é zero se tais strings forem

iguais. Isso é chamado de distância de Hamming. Uma medida de distância precisa ser simétrica (ou seja, a ordem dos argumentos não importa) e deve satisfazer à desigualdade triangular, isto é:

$$distance(a, c) \leq distance(a, b) + distance(b, c). \quad (4)$$

Os autores supra citados complementam que, qualquer função de similaridade na faixa $[0,1]$ pode ser facilmente transformada em uma função de distância $D(d_i, d_j)$ por (Baeza-Yates e Ribeiro-Neto, 2013, p. 210):

$$D(d_i, d_j) = 1 - Sim(d_i, d_j). \quad (5)$$

Neste contexto, vale ressaltar que, por a linguagem de programação *Python* possuir código fonte aberto, é possível a inclusão de métodos de pareamento e medidas de similaridade pelos usuários.

Na sequência, apresenta-se as medidas de similaridades utilizadas nesta pesquisa.

K.2.1 Distância de Ratcliff/Obershelp

A medida de similaridade de Ratcliff/Obershelp foi introduzida por Ratcliffe, em 1983. Neste contexto, segundo Ratcliff e Metzener (1988), o algoritmo referente a esta medida, teve um forte impacto positivo sobre a indústria de *softwares* educativos. Assim, como as demais medidas de similaridade, Ratcliff/Obershelp calcula a semelhança entre duas cadeias ou bases de dados, como o número de *caracteres concordantes*, dividido pelo número total de caracteres nos dois strings. Os *caracteres concordantes* são aqueles com a maior subsequência comum.

O algoritmo Ratcliff/Obershelp (D_{ro}) é expresso pela fórmula por (Ratcliff e Metzener, 1988):

$$D_{ro} = \frac{2K_m}{|S_1| + |S_2|}, \quad (6)$$

em que

- K_m é um número de caracteres concordantes;
- $|S_1|$ e $|S_2|$ são comprimentos das strings, respectivamente.

Inicialmente, determina-se a substring mais longa que as strings S_1 e S_2 têm em comum. Tal substring é denominada de âncora.

O valor de K_m cresce de acordo com o tamanho da âncora. Na sequência, as demais partes string a direita e a esquerda, devem ser examinadas como se fossem novas strings (em outras palavras, o passo inicial do algoritmo é repetido). O processo é repetido até que todos os caracteres das strings S_1 e S_2 sejam analisados.

Exemplo: Cálculo de score via algoritmo de Ratcliff/Obershelp (exemplo extraído de Ilyankou (2014))

Inicialmente, considere as strings MATHEMATICS e MATEMATICA

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
S_1	M	A	T	H	E	M	A	T	I	C	S
S_2	M	A	T	E	M	A	T	I	C	A	

- (a) Os comprimentos das string S_1 e S_2 , são dados por $|S_1| = 11$ e $|S_2| = 10$, respectivamente.
- (b) A substring mais longa de caracteres em comum é *EMATIC*. Portanto, *EMATIC* é uma âncora, e $K_m = |EMATIC| = 6$.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
S_1	M	A	T	H	E	M	A	T	I	C	S
S_2	M	A	T	E	M	A	T	I	C	A	

- (c) À esquerda da âncora existem os conjuntos de símbolos *MATH* e *MAT* restantes. A maior subsequência comum delas é *MAT*. Portanto, $K_m = 6 + |MAT| = 9$.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
S_1	M	A	T	H	E	M	A	T	I	C	S
S_2	M	A	T	E	M	A	T	I	C	A	

- (d) Como a substring *MAT* está no início de ambas as strings S_1 e S_2 , não há símbolos para a esquerda da mesma. Na figura do item (c), verifica-se que à direita de *MAT*, na string S_1 , tem-se *H*, já na string S_2 , observa-se que não existem caracteres. Portanto, K_m permanece. Por fim, interessa-se verificar os caracteres à direita da âncora.

- (e) No que concerne aos caracteres à direita da âncora, tem-se *S* e *A*. Percebe-se que, por esses caracteres serem distintos, eles não deturparão os resultados. Assim, o valor de K_m permanece igual a 9 e todos os caracteres, em ambas as strings, são considerados. Desta feita, tem-se todos os dados necessários para calcular o score via algoritmo de Ratcliff/Obershelp.

Portanto, o score de similaridade Ratcliff/Obershelp para as strings MATHEMATICS e MATEMATICA é:

$$D_{ro} = \frac{2 \times 9}{10 + 11} = \frac{18}{21} = 0.857.$$

Na sequência, define-se a medida de similaridade de Jaro-Winkler.

K.2.2 Distância de Jaro-Winkler

Na ciência da computação e estatística, a distância Jaro-Winkler (Winkler, 1990) é uma medida da similaridade entre duas strings. Ela é uma variante da métrica de distância Jaro (Jaro, 1989, 1995; Winkler, 1990). Quanto menor a distância Jaro-Winkler entre duas strings, mais semelhantes elas são.

A distância de Jaro-Winkler (D_j) entre duas strings, S_1 e S_2 é expressa por (Winkler, 1990):

$$D_j = \begin{cases} 0, & \text{se } m = 0 \\ \frac{1}{3} \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right), & \text{caso contrário,} \end{cases} \quad (7)$$

em que

- m é o número de caracteres concordantes;
- t é metade do número de transposições.
- $|S_1|$ e $|S_2|$ correspondem ao comprimentos das strings.

Segundo Rodichevski (2013), dois caracteres de S_1 e S_2 são considerados concordantes (pares) somente se estes são iguais ou não se distanciam além do valor limite $\left\lfloor \frac{\max(|S_1|, |S_2|)}{2} \right\rfloor - 1$. Assim, para cada par de caracteres concordantes com ordens distintas nas strings S_1 e S_2 , o número de transposições, t , é acrescido de 1 (um).

Esclarece o conceito de transposição: o exemplo descrito por Ilyankou (2014) considere-se as palavras *HOUSE* e *HOME*. Neste caso, tem-se três símbolos concordantes, H , O e E . Uma vez que estes símbolos aparecem na mesma ordem, em ambas as strings, o número de transposições para tais strings é dada por $t = \frac{0}{2} = 0$.

Em contraste, o número de símbolos concordantes para as palavras *HOUSE* e *HOUES* são H , O , U , S , E . Mas, os caracteres S e E aparecem em ordens diferentes nas strings. Logo, o número de transposições para tais strings é dado por $t = \frac{2}{2} = 1$.

Finalmente, se o número de símbolos concordantes, m , é igual a 0 (zero), a distância de Jaro-Winkler deve ser considerada como 0 (zero), visto que, matematicamente, a divisão por 0 (zero) não pode ser realizada.

Exemplo: Cálculo da distância entre duas palavras usando o método de Jaro-Winkler (exemplo extraído de Ilyankou (2014))

Considere as strings MATHEMATICS e MATEMATICA

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
S_1	M	A	T	H	E	M	A	T	I	C	S
S_2	M	A	T	E	M	A	T	I	C	A	

- (a) Os comprimentos das string S_1 e S_2 , são dados por $|S_1| = 11$ e $|S_2| = 10$, respectivamente.
- (b) Inicialmente, os três primeiros símbolos M , A , T , de cada string coincidem (portanto, são concordantes). Assim, o número de caracteres concordantes, nesta etapa, é igual a $m = 3$.
- (c) O símbolo [4] da string S_1 - H - não é o mesmo símbolo da string S_2 (E). A distância d_m , de Rodichevski (2013) é dada por:

$$d_m = \left\lfloor \frac{\max(11, 10)}{2} \right\rfloor - 1 = \left\lfloor \frac{11}{2} \right\rfloor - 1 = 5 - 1 = 4.$$

- (d) Na string S_2 , a partir do símbolo [4], em que $d_m = 4$, conta-se 4 símbolos para à direita e quatro para à esquerda, a fim de encontrar o símbolo H , mas, para à direita, tem-se $M-A-T-I$ e à esquerda $M-A-T$. Portanto, não existe símbolos concordantes para H .
- (e) Agora, na string S_1 , tem-se que E está na posição [5]. Apesar do símbolo [5] na string S_2 ser M e não E . À esquerda deste, tem-se o símbolo [4], “ E ”, que é concordante. Assim, por definição, o número m deve ser acrescido de 1. Desta feita, o número de símbolos concordantes é $m = 4$.

(f) O processo descrito em (e) será repetido com todos os caracteres da string S_2 com os índices de [6] a [10]. Assim, depois de processar o símbolo [10], o número de símbolos concordantes é igual a $m = 9$. O string S_2 não contém um símbolo concordante para $S(\text{string } S_1, [11])$.

(g) Na sequência, após a identificação dos símbolos concordantes, deve-se calcular o número de transposições. Para este exemplo, todos os símbolos concordantes aparecem na mesma ordem em ambas as strings. Assim, por estas apresentarem diferentes índices nas duas strings, tal fato não influencia no número de transposições, isto é, $t = \frac{0}{2} = 0$.

(h) Desta feita, tem-se todos os dados necessários para calcular-se a distância de Jaro-Winkler, D_j . Assim, a distância de Jaro-Winkler, D_j , para as palavras MATHEMATICS e MATEMATICA é dada por:

$$D_j = \frac{1}{3} \left(\frac{9}{11} + \frac{9}{10} + \frac{9-0}{9} \right) = \frac{1}{3} \left(\frac{9}{11} + \frac{9}{10} + \frac{1}{1} \right) = 0.906.$$

K.2.3 Distância de Jaro

Alguns aperfeiçoamentos no processo de pareamento de Fellegi e Sunter (1969), foram propostos por Jaro e Winkler, em 1995, os quais foram elaborados durante a aplicação dessa teoria em processos de pareamento realizados em Censos e na Pesquisa de Avaliação da Cobertura da Coleta (PA) nos Estados Unidos.

Neste contexto, Jaro (1989), propôs pela primeira vez o uso do algoritmo **EM** para o processo de pareamento probabilístico do Censo e da PA de 1985 em Tampa, Flórida. No que concerne o uso deste algoritmo, vale ressaltar que este possibilitou estimar os parâmetros de Fellegi e Sunter, acima descritos, para cada par de registros, tornando possível o desenvolvimento de *softwares* para a realização de pareamento de dados.

De acordo com Cohen *et al.* (2003), dadas as strings $s = a_1 \dots a_K$ e $t = b_1, \dots b_L$, define-se o caracter a_i , comum entre s e t , em que $b_j = a_i$ de tal modo que $i - H \leq j \leq i + H$, em que $H = \frac{\min(|s|, |t|)}{2}$. Assim, seja $s' = a'_1 \dots a'_K$, os caracteres distintos a s e que são comuns com t (na mesma ordem em que aparecem os s) e, de maneira análoga, seja $t' = b'_1, \dots b'_L$, tem-se a transposição para s' , t' , tal que na posição i , verifica-se $a'_i \neq b'_i$. Desta feita, obtem-se a metade do número das transposições, isto é, $T_{s',t'}$.

Corroborando com o exposto, tem-se que, outro aperfeiçoamento proposto por Jaro foi o uso de uma medida de comparação entre as variáveis dos registros, tomadas enquanto duas strings, calculado pela fórmula por (Jaro (1995) e Cohen *et al.* (2003)):

$$Jaro(s, t) = \frac{1}{3} \left(\frac{|s'|}{|s|} + \frac{|t'|}{|t|} + \frac{|s'| - T_{s',t'}}{|s'|} \right), \quad (8)$$

em que

s é o número de caracteres da variável na string 1;

t é o número de caracteres da variável na string 2;

s' e t' são os números de caracteres iguais e na mesma posição das duas strings; e

$T_{s',t'}$ é o número de transposições necessárias para a string 1 ficar igual a 2.

Winkler (1999) aplicou a teoria de Fellegi e Sunter (1969) dos pareamentos nos dados do Censo e da PA estadunidense de 1990 nas cidades de Saint Louis, Columbia e Washington e propôs uma medida de comparação que é uma extensão daquela proposta por (Jaro, 1995):

$$Jaro - Winkler(s, t) = Jaro(s, t) + \frac{P'}{10} \cdot (1 - Jaro(s, t)), \tag{9}$$

em que P' é o número de caracteres comuns entre os quatro primeiros caracteres de ambas as strings, variando, dessa maneira, entre 0 e 4.

Em Da-Silva *et al.* (2010), enfatizam-se que as medidas de Jaro e Winkler objetivam parear registros cujos valores das variáveis não estão escritos de forma absolutamente igual. Como erros de grafia são comuns na composição de bases de dados, se os registros pertencem ao mesmo indivíduo ou unidade o par deve ser formado mesmo que haja divergências de grafia entre eles

Por fim, existem evidências que as medidas de similaridade de Jaro e Jaro-Winkler são destinadas para seqüências curtas (por exemplo, sobrenomes pessoais).

K.2.4 Distância de Jaccard

Cohen *et al.* (2003) propõem que duas string s e t , também, podem ser considerados como multisets, isto é, conjunto de palavras compostas por símbolos. Os autores, completam, consideram que existe várias métricas de distância baseados em símbolos.

Neste contexto, a similaridade de Jaccard é um índice comum para variáveis binárias. É definida como o quociente entre o cruzamento e a união das variáveis comparadas aos pares entre os dois objetos.

O índice de similaridade de (Jaccard, 1901) é definido por:

$$Jaccard(S, T) = \frac{|S \cap T|}{|S \cup T|}. \tag{10}$$

No que concerne a expressão (10), tem-se que o coeficiente de similaridade de Jaccard, entre dois conjuntos de dados, é o resultado da divisão entre o número de características que são comuns a todos, dividido pelo número de propriedades. Vale ressaltar que, se S e T são ambos vazios, tem-se que $Jaccard(S, T) = 0$, em que $0 \leq Jaccard(S, T) \leq 1$. Além disso, $Jaccard(S, S) = 1$, $Jaccard(S, T) = 0$ se $S \cap T = \emptyset$.

Exemplo: Deseja-se calcular a distância de Jaccard entre X e Y em que

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
X	1	0	0	0	0	0	0	0	0	0
Y	0	0	0	0	0	0	1	0	0	1

Sejam as contagens f_{ij} , com $i, j \in \{0, 1\}$ representando o número de atributos onde $X = i$ e $Y = j$. Desta forma,

$$f_{11} = 0, \quad f_{01} = 2, \quad f_{10} = 1, \quad e \quad f_{00} = 7.$$

Logo, a distância de Jaccard é dada por

$$Jaccard(X, Y) = \frac{f_{11}}{f_{01} + f_{10} + f_{11}} = \frac{0}{2 + 1} = 0.$$

K.2.5 Distância de Levenshtein

O nome da distância de Levenshtein advém do cientista russo Vladimir Levenshtein, que considerou esta distância já em 1965. Tal distância é útil para aplicações em que é necessário determinar quão semelhantes duas strings são, como, por exemplo é o caso com os verificadores ortográficos. Para mais detalhes, ver (Levenshtein, 1966).

A distância Levenshtein, também conhecida como distância de edição entre dois strings, é dada pelo número mínimo de operações necessárias para transformar uma string na outra. Entende-se por “operações” a substituição de um caracter.

Segundo Navarro (2001), este algoritmo permite inserções, remoções e substituições. E pode ser parafraseado como “o menor número de inserções, remoções e substituições para igualar duas strings”.

A similaridade de Levenshtein entre duas strings a e b , de tamanho $|a|$ e $|b|$, respectivamente, é dada por (Levenshtein, 1966):

$$Lev_{a,b}(i, j) = \begin{cases} \max(i, j) & \text{se } \min(i, j) = 0, \\ \min \begin{cases} Lev_{a,b}(i-1, j) + 1, \\ Lev_{a,b}(i, j-1) + 1, \\ Lev_{a,b}(i-1, j-1) + 1_{a_i \neq b_j} \end{cases} & \text{caso contrário,} \end{cases} \quad (11)$$

em que $1_{a_i \neq b_j}$ é a função indicadora igual a zero quando $a_i = b_j$ e igual a 1 caso contrário e $Lev_{a,b}(i, j)$ corresponde a distância entre o primeiro i -ésimo caracter de a e o primeiro j -ésimo caracter de b . Ou seja, i e j são respectivamente as posições dos caracteres destas duas strings a e b .

Note-se que o primeiro elemento no mínimo corresponde à anulação (de a para b), a segunda a inserção e a terceira são para igualar ou para gerar a incompatibilidade, dependendo se os respectivos símbolos são os mesmos.

Complementarmente ao exposto, verifica-se que a distância de Jaro como a de Levenshtein, também utiliza uma matriz, embora esta só seja utilizada no modelo conceitual e não precise ser implementada na prática, já que não é necessário gravar todos os dados da matriz.

Exemplo: Cálculo de score via algoritmo de Levenshtein

A distância de Levenshtein entre as strings “kitten” e “sitting” é 3 (três), uma vez que as três seguintes edições mudam de um para o outro, e não há maneira de fazê-lo com menos de três edições:

1. kitten → sitten (substituindo “k” por “s”);
2. sitten → sittin (substituindo “e” por “i”);
3. sittin → sitting (inserindo “g” no final).

K.2.6 Distância de Sørensen-Dice

(Sørensen, 1948) propôs uma medida de similaridade entre duas strings que tem a vantagem de ser muito simples de calcular.

A medida de similaridade de Sørensen ($Sorensen(A, B)$) é dada pela seguinte expressão (Sørensen, 1948):

$$Sorensen(A, B) = \frac{2C}{A + B}, \tag{12}$$

em que A representa o tamanho do primeiro conjunto de dados, B o tamanho do segundo conjunto de dados e C corresponde ao tamanho da intersecção entre os dois conjuntos de dados.

Exemplo: Cálculo de score via algoritmo de Sørensen-Dice

Considere as strings NIGHT e NACHT

	[1]	[2]	[3]	[4]	[5]
String 1	N	I	G	H	T
String 2	N	A	C	H	T

(a) Agora, deseja-se encontrar o conjunto de cruzamentos em cada palavra, string

	[1]	[2]	[3]	[4]
String (A)	NI	IG	GH	HT
String (B)	NA	AC	CH	HT

(b) Verifica-se que cada conjunto tem quatro elementos, e o cruzamento desses dois conjuntos tem apenas um elemento em comum, HT.

(c) Por fim, utilizando-se a expressão 12, tem-se

$$Sorensen(A, B) = \frac{2 \times 1}{(4 + 4)} = 0.25.$$

Completando o exposto, vale ressaltar que, nos anexos desta pesquisa, pode-se encontrar, dentre outros, todos os algoritmos destas medidas de similaridade, aqui citadas, devidamente implementadas em linguagem *Python*.

K.3 Resultado dos pareamentos

Conforme visto na Seção 9.1, o melhor modelo de pareamento é aquele que resulta na maior quantidade de pares verdadeiros com o menor número de pares falsos. Assim, nesta seção será apresentado alguns resultados relacionados aos bancos de dados apresentados nas Seções 2.1, 2.2, 2.3 e 2.4.

A Tabela K.1 apresenta algumas informações gerais sobre os bancos de dados utilizados na pesquisa.

Observa-se que a Web Of Science é um banco de dados bastante extenso, mas que possui algumas deficiências, como, por exemplo, a duplicidade de registros. A base da CWTS tem como característica a ausência do ISSN dos periódicos, e a limitação de obter-se, através do site, o limite máximo 1.000 periódicos, dificultando, assim, o estabelecimento de pareamentos.

Tabela K.1: *Informações sobre os periódicos das bases de dados: Classificação de Periódicos no Qualis/CAPEs 2012, Web Of Science, Scimago Journal & Country Rank e Journal Indicators da University's Centre for Science and Technology Studies (CWTS).*

Base de dados	Banco de dados	Total de observações	Sem ISSN	ISSN duplicados	ISSN distintos	Ano base
1	Qualis	1.188	6	0	1.182	2012
2	Scimago	1.284	16	1	1.267	2014
3	Web Of Science	11.813	3	619	11.191	2014
4	CWTS	1.000	1.000	-	-	2014

Fonte: Elaborado pelos autores (2016).

Complementando o exposto na Tabela K.1, destacam-se as seguintes informações sobre os referidos bancos de dados

- 1 - Qualis ano base 2012: o banco de dados é formado por 1.188 periódicos, dos quais 6 não possuem ISSN, resultando, assim, em 1.182 chaves distintas;
- 2 - Scimago ano base 2014: a base é formada por 1.284 periódicos, dos quais 16 não possuem ISSN e 1 periódico está repetido. Logo, a base de dados a ser analisada possui 1.267 chaves distintas. A chave repetida tem ISSN 1025-8973. Esta chave não foi encontrada nos demais bancos de dados;
- 3 - Web Of Science ano base 2014: esta base é formado por 11.813 periódicos dos quais 3 não possuem ISSN e 619 periódicos estão duplicados. Desta forma, a base de dados é formada por 11.191 chaves distintas;
- 4 - CWTS ano base 2014: esta base possui uma amostra de 1.000 periódicos, ordenados segundo o valor de seu SNIP, embora não tenha sido obtido o ISSN para nenhum destes;
- 5 - o total de registros comuns, segundo o ISSN, nas bases 1, 2 e 3 é igual a 215, dos quais 5 são repetidos duas vezes. Os ISSNs que apresentam duplicidades são: 0304-4068 / 0167-6687 / 0282-423X / 0714-0045 / 0304-4076.

K.3.1 Pareamentos estabelecidos através do ISSN

Como já mencionado, observou-se um total de apenas 215 registros (ISSNs) comuns às bases de dados Qualis, Web Of Science e Scimago (vide Tabela K.2). Considerando o subconjunto formado por estes 215 registros, procedeu-se um novo pareamento, agora com a base CWTS, utilizando os títulos dos periódicos, resultando em uma base de 206 registros comuns. Esses 206 registros teoricamente incorporam o menor número de pares falsos, visto que, de acordo com a Tabela K.1, a base de dados CWTS é a única que não apresenta o ISSN, chave usada para realizar os pareamentos entre as bases 1, 2 e 3.

A Tabela K.2 apresenta os resultados dos pareamentos que foram realizados dois a dois e três a três, levando em conta o ISSN como chave comum entre as bases de dados.

Tabela K.2: *Pareamentos dois a dois e três a três entre as bases de dados, classificação de periódicos no Qualis/CAPES, Web Of Science e Scimago, que apresentam o ISSN, chave a ser utilizada, como variável em comum.*

Pareamentos	Descrição	ISSN distintos	Número de registros
1	Qualis vs. Scimago	297	297
2	Qualis vs. Web Of Science	766	799
3	Scimago vs. Web Of Science	416	427
4	Qualis vs. Scimago vs. Web Of Science	210	215

Fonte: Elaborado pelos autores (2016).

De acordo com a Tabela K.2, observa-se que, com mais de duas bases de dados, os valores das interseções diminuem, quando se aumenta o número de elementos envolvidos, tal como esperado. O pareamento 4 é constituído por todos os elementos que pertencem, simultaneamente, às bases de dados 1, 2 e 3, vistas anteriormente na Tabela K.1.

Nos Capítulos 7 e 8 apresenta-se as bases produtos para esta pesquisa. Ressalta-se que no Capítulo 8 para a construção da base final não considerou nos pareamentos a base do CWTS, devido a tal fato os valores são diferentes dos apresentados na Tabela K.2.

K.3.2 Pareamentos estabelecidos através do título do periódico

Como visto na Seção 9.1 deste trabalho, existem medidas de similaridade que permitem concluir que os registros são iguais, mesmo que haja divergências de grafia nos títulos dos periódicos.

Nos pareamentos feitos a partir do “título do periódico” utilizou-se algumas medidas de similaridade para as quais os coeficientes mais próximos de 1 indicam maior precisão, isto é, permitem menores divergências de grafia. Para a variável, ISSN, apenas os pareamentos exatos foram considerados.

No decorrer da construção dos códigos de programação usando a linguagem *Python*, (vide **Apêndice A**), foi necessária a realização de algumas alterações nos bancos de dados, pois identificou-se problemas na construção da base de dados, tais como o deslocamento de campos que ficaram descobertos após ter-se observado registros com o campo ISSN em branco.

Na construção da base de dados final para esta pesquisa realizou-se, também, algumas modificações na programação dos códigos em *Python*, objetivando testar a inclusão de registros que, nas bases originais, não contemplavam o ISSN, tal como o caso da base CWTS. Desta forma, criou-se um algoritmo para testar qual seria o ponto de corte ideal para se considerar um pareamento correto, aplicando-se medidas de similaridade entre os títulos dos periódicos. Tal procedimento resultou no ponto de corte de 0.7636, que descreve o valor mínimo para o qual se obteve um desempenho aceitável de pareamento utilizando-se cada uma das medidas citadas no capítulo anterior. Entretanto, objetivando um melhor desempenho dos pareamentos por meio dos títulos dos periódicos, utilizou-se como ponto de corte padrão o valor de 0.866 que, segundo os testes realizados, (vide as figuras ?? e ??) confirmou a superioridade do método de Jaro-Winkler.

Apesar do procedimento realizado não ser capaz de apontar com precisão o percentual de falsos positivos e/ou falsos negativos, foi possível estabelecer pareamento de quase 100% dos títulos procurados. Ainda que este resultado seja bastante promissor, o mesmo motivou o estudo de procedimentos para a obtenção dos valores reais dos ISSNs dos periódicos pareados por similaridade. Neste sentido, o desenvolvimento de um *web crawler* (código de programação que funciona como um rastreador web, ou seja, um programa de computador

que navega pela World Wide Web de forma metódica e automatizada) foi a solução para descobrir os códigos que correspondessem à amostra dos 1.000 periódicos obtidos da base de dados da CWTS. Maiores detalhes sobre o Web Crawler desenvolvido são descritos na Seção 9.1.

No que concerne ao modelo determinístico, aplicado sem a utilização de uma chave exata, necessária, em especial, no caso da base de dados CWTS, novos pareamentos foram testados utilizando-se o título dos periódicos como chave de pareamento, considerando-se diferentes pontos de corte (vide Figuras L.1 e L.2). Estes pareamentos foram estabelecidos entre as bases CWTS e a base 4 (vide Tabela L.2), resultante do pareamento Qualis *vs.* Scimago *vs.* Web Of Science. Assim, um algoritmo foi utilizado para comparar a eficiência dos métodos de similaridade, na busca do que melhor caracterizasse os pareamentos realizados. Este procedimento foi avaliado utilizando-se dois intervalos distintos de percentuais de similaridade, a saber: [70%, 95.5%] e [86%, 95.5%]. Os resultados podem ser observados, respectivamente, nas Figuras L.1 e L.2.

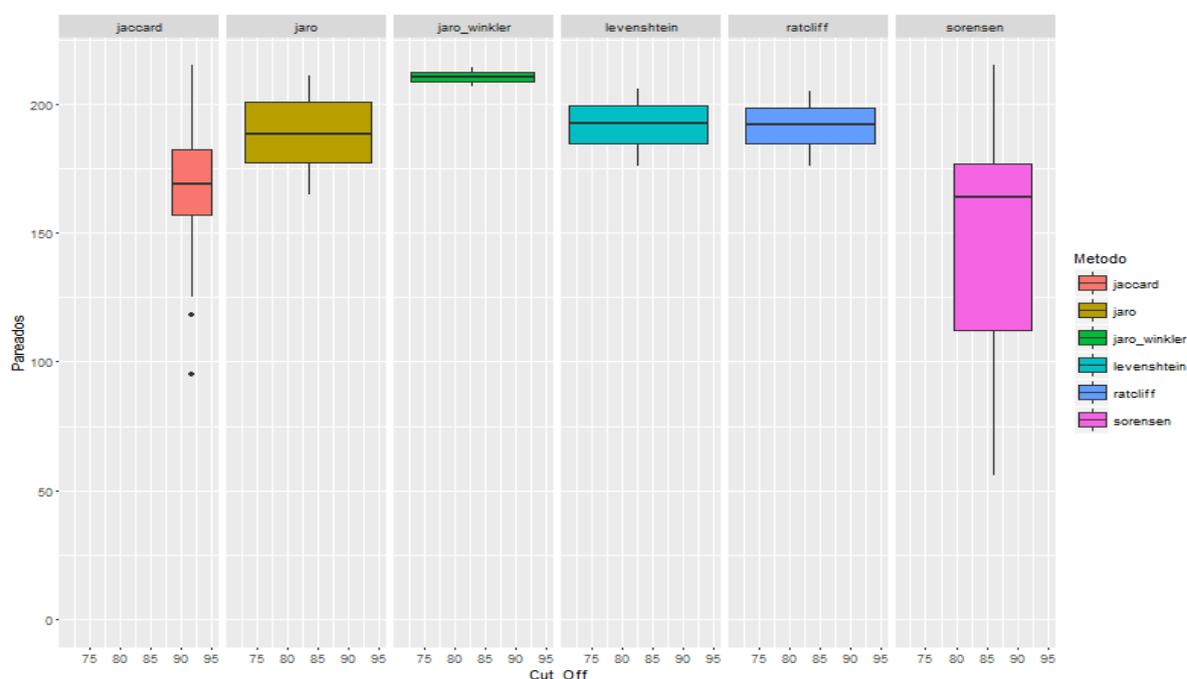


Figura K.1: Avaliação das medidas de similaridade no intervalo de [70%, 95.5%].

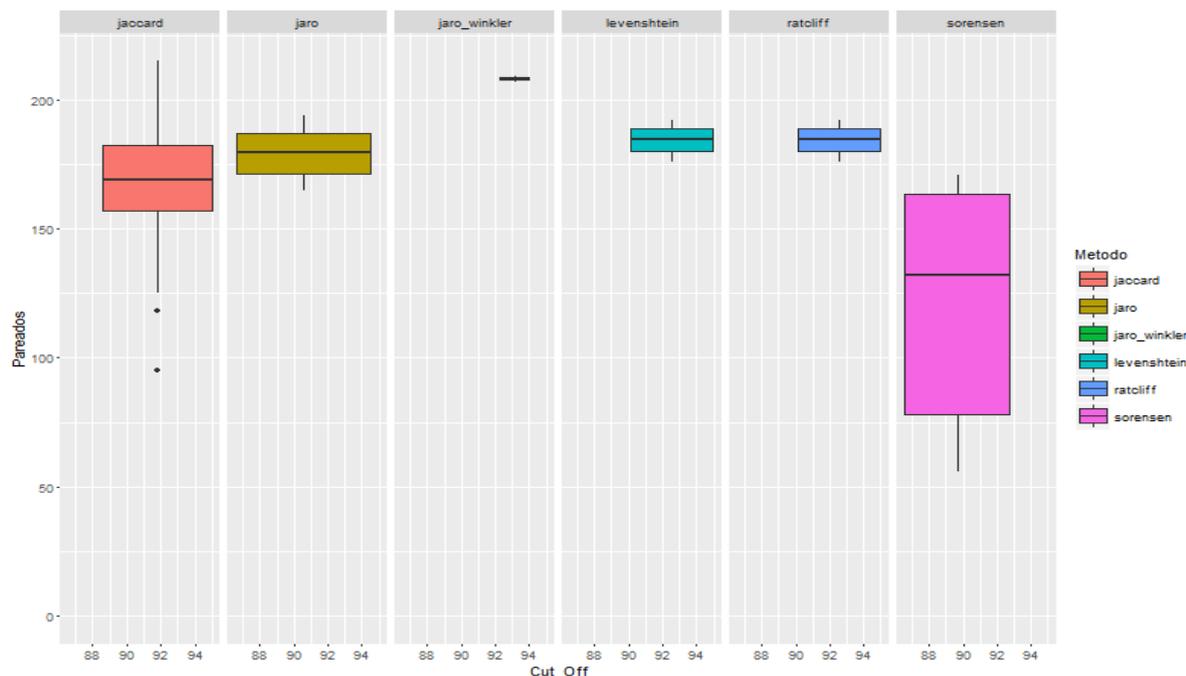


Figura K.2: Avaliação das medidas de similaridade no intervalo de [86%, 95.5%].

Acerca dos pares formados, Da-Silva *et al.* (2010) propôs que, num processo de pareamento, dois registros são considerados um par verdadeiro quando eles referem-se à mesma unidade física, neste caso título dos periódicos. De forma análoga, os registros são considerados falsos pares quando eles não estão associados à mesma unidade.

De acordo com os box-plots apresentados, à medida que o percentual de um coeficiente de similaridade aumenta, o número de pares verdadeiros tendem a diminuir.

Visto que a base do CWTS não apresenta o ISSN em seus registros, considerou-se como ponto de corte o percentual de 86.66%. No que concerne esse ponto de corte, vale ressaltar que o mesmo é baseado em conhecimento experimental. A Tabela K.3 apresenta os resultados obtidos nestes pareamentos. Tais resultados foram avaliados manualmente para verificar se os pares formados estavam corretos.

Tabela K.3: Medidas de similaridade usadas nesta análise.

Regra	Par Verdadeiro	Repetidos	Par Falso
Exato	210	5	0
Jaccard	0	0	215
Jaro	197	5	13
Jaro-Winkler	110	2	103
Levenshtein	196	5	14
Ratcliff	196	5	14
Sörensen-Dice	0	0	215

Fonte: Elaborado pelos autores (2016).

De acordo com a Tabela K.3, verifica-se que cada uma das medidas de similaridade apresentam diferentes particularidades a serem consideradas conforme o caso em que se estiver tratando. De um modo geral, o que se observa conjuntamente com os resultados das Figuras ?? e ?? é que as medidas de similaridade de Levenshtein, Jaro e Ratcliff têm comportamento semelhante para o caso estudado, uma vez que formaram um elevado número de pares verdadeiros, com baixa variabilidade.

Utilizando-se as medidas de similaridades descritas na Tabela K.3, foram pareados periódicos na variável *título* escritos de maneira distinta como “Mathematics” e “Matema-

tics”. Claramente, trata-se do mesmo periódico, entretanto, em algumas das bases de dados encontrou-se o termo “Mathematics”, onde em outras somente “Matemáticas”. Sendo assim, periódicos com um score igual ou superior a 0.866 (ponte de corte considerado nas análises) nesta variável foram pareados.

Desta forma, o resultado obtido utilizando as medidas de similaridade foi satisfatório, pois proporcionou um bom percentual de pares verdadeiros comparados aos pares obtidos na primeira fase do pareamento, no qual obteve-se 215 registros pareados, sendo 5 de mesmo ISSN. Além disso, considerando que o procedimento executado em *Python* foi desenvolvido de forma a permitir pareamentos de grandes bases de dados, vale ressaltar que, o tempo de processamento de um pareamento efetuado com as bases de dados utilizadas, foi de aproximadamente, trinta segundos. Este tempo, se considerado o tamanho das bases utilizada comparado com o tempo de processamento no R, não é muito diferente. Porém, à medida em que o tamanho das bases aumenta, o tempo de resposta do R é, provavelmente, maior do que o tempo executado em *Python*. Ressalta-se que testes desta natureza não foram realizados, por não serem foco deste trabalho. Porém, sabe-se que o R ao tratar com grandes bases de dados apresenta problemas de memória (ex.: <http://www.r-bloggers.com/reduce-memory-use-for-large-datasets>), fato este de muito menor impacto no uso das rotinas criadas em *Python* em função da indexação de arquivos mantidos em disco.

K.3.3 Busca sistemática do ISSN via *Web Crawler*

O contexto no qual realizou-se o pareamento entre as quatro bases de dados apresentadas na Tabela K.1 foi meramente determinístico. Com exceção da base de dados CWTS, para todas as demais utilizou-se o valor do ISSN como chave direta para parear periódicos. No caso da CWTS, como já apresentado anteriormente, a variável de interesse utilizada na similaridade das bases de dados, passou a ser o *título do periódico*.

A título de investigação inicial do problema, tomou-se uma pequena amostra de registros de cada uma das bases em questão, ordenados por título, para reduzir o esforço de verificação manual. Com as bases ordenadas desta forma, foi possível observar pequenas divergências na grafia entre alguns dos títulos, o que certamente inviabilizava o uso direto dos títulos como chave direta.

Para os periódicos com títulos exatamente iguais em mais de uma base de dados, registrou-se o resultado da busca por título no site da CWTS (<http://www.journalindicators.com/indicators>), afim de verificar se o ISSN resultante da busca no site era o mesmo em qualquer uma das outras bases de dados que continham o periódico. Este foi o exercício inicial para se entender a dinâmica de funcionamento da pesquisa manual. O desafio era automatizar este processo e repeti-lo para cada um dos 1.000 periódicos presentes na base obtida da CWTS.

Além de encontrar uma forma sistematizada de obter os ISSNs de cada periódico, fêz-se necessário repetir os pareamentos já estabelecidos, utilizando-se, como chave, o título de cada periódico, com o valor de ISSN obtido do site, para, então, avaliar quantos periódicos haviam sido corretamente identificados, em ambas as bases, com o uso das medidas de similaridade.

O primeiro passo foi estudar a estrutura do site da CWTS que fornece uma forma de buscar informações dos periódicos por consulta do título. Assim, um estudo minucioso das *tags* da linguagem HTML (HyperText Markup Language) utilizadas no código fonte da página de busca e sua dinâmica de construção sob requisição, foi realizado com a ajuda das ferramentas de desenvolvimento disponíveis no web browser Mozilla Firefox. Estas ferramentas indicaram que, para se obter o ISSN de cada periódico para o qual se tinha o título, eram necessárias duas etapas:

- 1º - Avaliar a lista de *links* disponibilizados após o resultado da busca, considerando apenas os registros que tenham o título exatamente igual ao solicitado e redirecionar o site corrente para a *url* (*Uniform Resource Locator* - endereço de um recurso disponível em uma rede, seja a rede internet ou intranet) indicada (vide Figura L.3), exemplificando o caso do periódico *Advances i Applied Mathematics*.

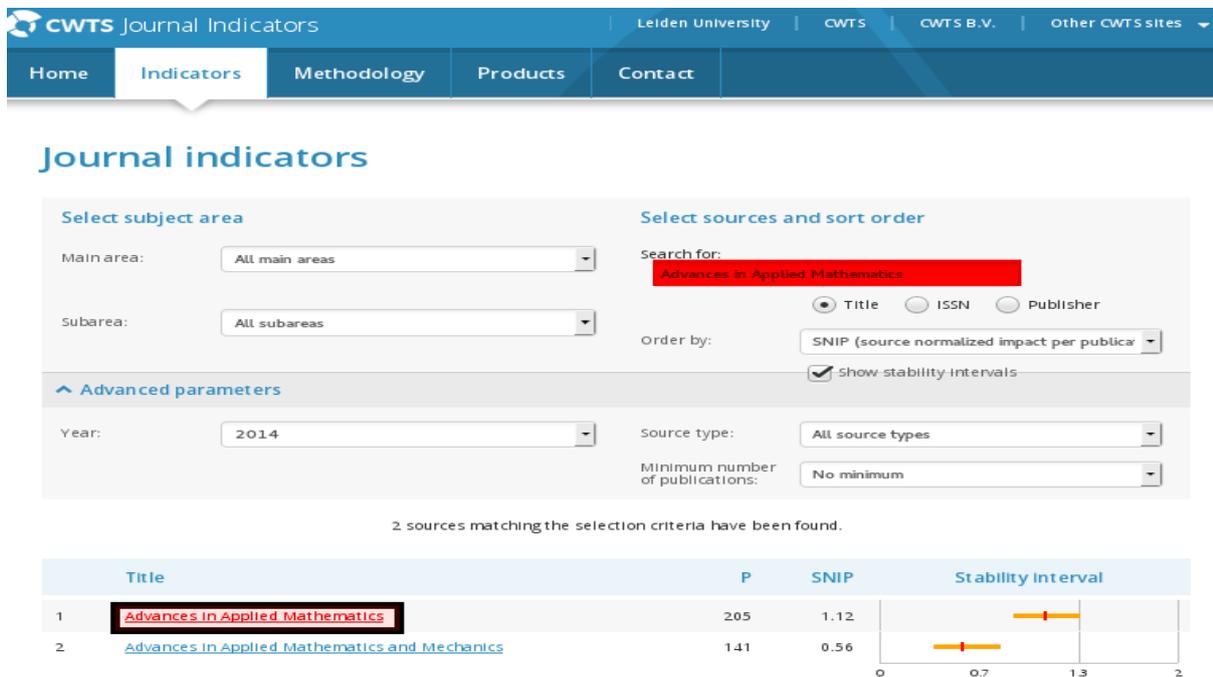


Figura K.3: Pesquisa pelo título no CWTS.

2º - Verificar a existência do valor do ISSN na url redirecionada. Neste passo considerou-se apenas o valor do ISSN apresentado no rótulo “Print ISSN” (vide Figura K.4).

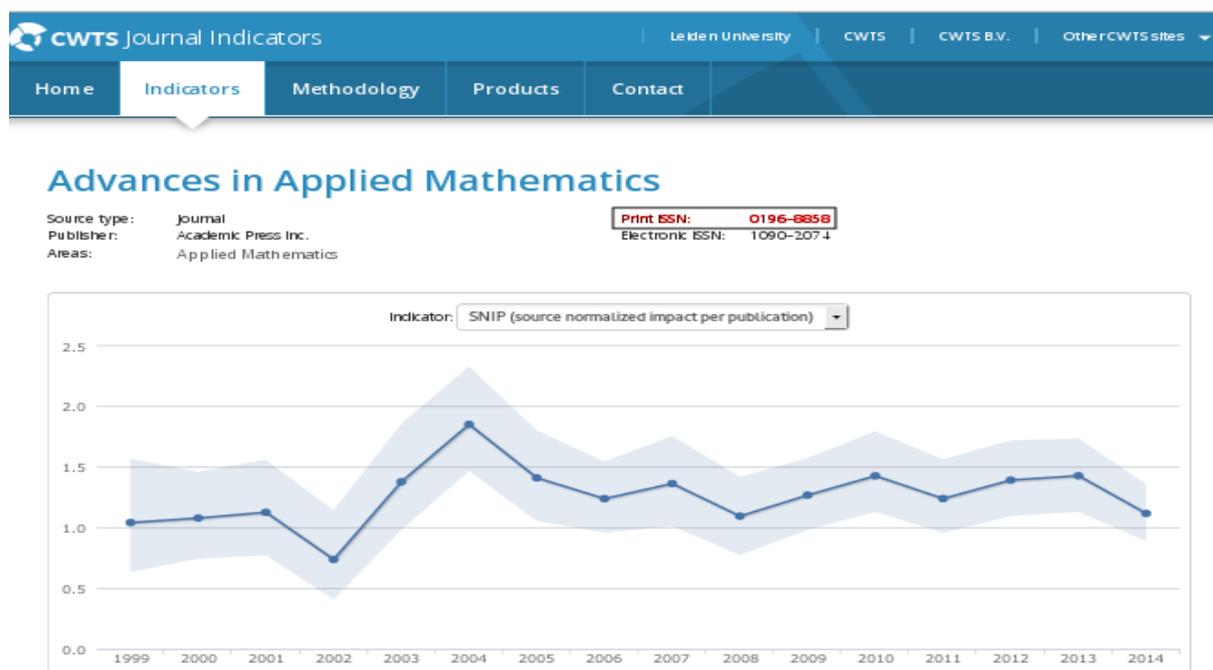


Figura K.4: Valor do ISSN no CWTS.

Compreendida a dinâmica de acesso dos dados no site de busca, utilizou-se um framework de desenvolvimento em *Python* para facilitar a criação do crawler, conhecido por *Scrapy* (<http://scrapy.org>). De fácil compreensão e com documentação detalhada de uso de cada funcionalidade, o Scrapy permitiu maior eficiência na solução do problema em questão, de modo a possibilitar que mais tempo fosse gasto no planejamento, do que na criação do código. Apesar da falsa aparência de simplicidade, o procedimento para sistematizar estes passos necessitavam considerar todas as diferentes formas que poderiam ocorrer como resultado da busca no site. Neste contexto, Nascimento (2011) enfatiza que, devido à grande quantidade de artigos existentes em sites de consulta, tratar de artigos relacionados a um trabalho tornou-se impraticável sem a utilização de algum arcabouço estruturado de pesquisa. Em Scheidt *et al.* (2008) a bibliometria (uma subárea da Cientometria, derivada da Bibliografia, que está diretamente relacionada a estudo de informações baseado em técnicas quantitativas e estatísticas de medição dos índices e também da produção e disseminação do conhecimento científico) é utilizada em conjunto com técnicas de mineração de dados para desenvolver uma nova metodologia de busca em trabalhos científicos aplicando um algoritmo de ranqueamento. Assim, faz uso de um *web crawler* para a busca otimizada de informações.

Para além do uso de buscas sistematizadas no que se conhece na World Wide Web tradicional, trabalhos têm sido desenvolvidos para encontrar referências de periódicos científicos utilizando-se crawlers até mesmo na parte oculta da web. Em Karishma e Singh (2012) apresenta-se uma discussão sobre o uso de crawlers, desenvolvidos para a *deep web* (também conhecida por Deepnet, Web Invisível, Undernet ou Web oculta, que se refere ao conteúdo da World Wide Web que não é indexado pelos mecanismos de busca padrão), em um cenário onde os volumes de informação são realmente muito extenso, reforçando, assim, a importância do desenvolvimento de ferramentas que sistematizem buscas, não apenas de periódicos científicos, como de quaisquer elementos que necessitem de uma estruturação detalhada de

pesquisa. Na mesma linha, [Vieira et al. \(2008\)](#) tratam de termos e questões técnicas no uso de crawlers para a mineração de dados, na chamada deep web.

K.4 Resumo de procedimentos adotados

Desde a primeira etapa, executada com o download das bases de dados, vários procedimentos foram realizados para a obtenção de uma base de dados final, contando com muitos tratamentos, vários destes realizados via pequenas instruções em *Shell Script* (linguagem de script usada em vários sistemas operacionais, com diferentes dialetos, dependendo do interpretador de comandos utilizado. Um exemplo de interpretador de comandos é o *bash*, usado na grande maioria das distribuições GNU/Linux).

De uma maneira geral, os procedimentos, passo a passo, executados com o intuito de obter-se uma base de dados final para o estudo da ordenação dos periódicos, podem ser assim resumidos:

- **Etapa I:** Download das bases de dados;
- **Etapa II:** Padronização da variável ISSN, de forma a garantir que em todas as bases de dados tivessem o mesmo formato;
- **Etapa III:** Verificação de registros deslocados (alguns dos arquivos que foram baixados só estavam disponíveis em formato PDF. Neste casos, os mesmos foram posteriormente, convertidos para CSV, acarretando falhas posicionais de campos, corrigidos manualmente com o auxílio de ferramentas como *Awk* - linguagem de programação interpretada que é, geralmente, usada para deixar os scripts de shell mais poderosos e com mais recursos);
- **Etapa IV:** Conversão de arquivos para o padrão UTF-8, permitindo padrão único de encoding nos arquivos;
- **Etapa V:** Adaptação de scripts em *Python* para realizar pareamentos conforme arquivos de configuração previamente estabelecidos;
- **Etapa VI:** Criação de arquivos indexados (gerados arquivos do tipo **Berkely DB** - biblioteca de software que fornece um banco de dados integrado de alto desempenho para dados de chave/valor) pela chave ISSN, com utilização destes para realizar e registrar os pareamentos iniciais segundo Tabela 9.1;
- **Etapa VII:** Avaliação dos arquivos indexados, verificando-se número de chaves distintas em cada base de dados, assim como o total de pareamentos realizados, verificando-se também o total de registros duplicados;
- **Etapa VIII:** Estudo e implementação das medidas de similaridade, com adaptação dos scripts em *Python* para realizar pareamentos por título, conforme arquivos de configuração previamente estabelecidos;
- **Etapa IX:** Execução do pareamento por título, contemplando o cálculo da similaridade entre cada um 210 periódicos encontrados, em comum, nas três bases que continham ISSN, com cada um dos 1.000 periódicos da amostra retirada da base do CWTS, considerando o valor de 86,6% de similaridade, como ponto de corte experimental, para se assumir correto o pareamento estabelecido. Este procedimento resultou na

geração de arquivos CSV para cada uma das 6 medidas de similaridade (vide Seção 9.1) consideradas no estudo, além dos respectivos arquivos de índice, contendo informações com o valor máximo de similaridade resultante do cruzamento entre os 210 periódicos resultantes do pareamento “4” da Tabela 9.1 com os 1.000 periódicos da amostra obtida da CWTS.

- **Etapa X:** Estudo do ponto de corte mínimo e máximo para cada uma das medidas de similaridade, apresentados nas Figuras ?? e ?? (este procedimento ainda não garantia, com precisão, o percentual de falsos positivos e/ou falsos negativos, uma vez que os valores de periódicos pareados se alteravam de acordom com o ponto de corte utilizado);
- **Etapa XI:** Estudo da dinâmica de acesso ao site da CWTS para obtenção do ISSN (testes foram realizados com alguns registros de forma manual, com o intuito de se entender quais as tags HTML continham os dados de *input* e *output* para obtenção do ISSN relativo o título pesquisado);
- **Etapa XII:** Estudo e teste dos aspectos a serem considerados no desenvolvimento do crawler para busca sistemática do ISSN no site da CWTS, considerando a utilização de bibliotecas de pareamento de tags em *Python*;
- **Etapa XIII:** Implementação e testes do crawler com a utilização do framework *Scrapy*.
- **Etapa XIV:** Avaliação dos resultados dos 1.000 ISSNs retornados pelo crawler, permitindo novo pareamento exato, assim como a confirmação de 95,24% (200 em 210) dos pareamentos por similaridade corretos (este procedimento permitiu concluir que o pareamento por similaridade é uma excelente opção, na falta do ISSN. Observou-se que os 10 periódicos não encontrados, não estavam incorretos, apenas não constavam na amostra inicial do CWTS, em razão desta ter sido retirada utilizando-se o ordenamento por SNIP, que à época resultou em uma amostra com resultados diferentes da ordenação pelo mesmo item, alguns meses depois, quando foi implementado e testado o crawler;
- **Etapa XV:** Preparação e categorização de variáveis dos 200 periódicos finais resultantes do pareamento realizado entre as 4 bases de dados estudadas (vide Tabela 9.1).

K.5 Base de dados final

As bases de dados finais que serão utilizadas nesta pesquisa, vale ressaltar sem a base CWTS, será apresentada em detalhes no Capítulo 6 deste estudo de dissertação.

Após a discussão da introdução da pesquisa, da apresentação dos bancos de dados e das técnicas de pareamento discutidas na literatura, apresenta-se, no capítulo seguinte, a teoria da resposta ao item - TRI.

Apêndice L - Processo de imputação de dados

Nesta apêndice, apresenta-se um resumo sobre a técnica de imputação de dados, a qual foi utilizada para o desenvolvimento do modelo relacionado ao estudo empírico contido neste trabalho. Trata-se de uma técnica amplamente aceita e utilizada pelo mercado na construção de bases nas quais se tem dados faltantes.

L.1 Imputação de dados

Uma complicação comum na pesquisa científica é a ocorrência de dados faltantes ou dados perdidos (*missing data*). Determinar a abordagem analítica adequada para bancos de dados com observações incompletas é uma questão que pode ser bastante delicada, pois a utilização de métodos inadequados pode levar a conclusões erradas sobre o conjunto de dados.

Para contornar este problema, existem diversas técnicas que envolvem a substituição dos dados faltantes por estimativas de valores plausíveis a serem imputados aos dados faltantes, tendo por objetivo “completar” os bancos de dados e possibilitar a análise com todos os dados em estudo.

Os métodos ou procedimentos para imputação são aqueles que se preocupam em substituir os valores ausentes, de uma unidade ou de um item, por estimativas dos mesmos. As primeiras técnicas de imputação desenvolvidas envolviam métodos relativamente simples, tais como substituição dos dados faltantes pela média ou pela mediana da variável, por interpolação ou até por regressão linear. Todos esses mecanismos permitem preencher dados faltantes através do que se chama “imputação simples”. Para se definir o melhor método de imputação a ser utilizado é importante conhecer o mecanismo de não resposta, pois a partir dele pode se ter a idéia a respeito do relacionamento entre a perda da informação e os valores das variáveis presentes na matriz de dados (Dias e Albieri, 1992).

Complementando o exposto, Nunes *et al.* (2009) enfatizam que os mecanismos de resposta são:

- Dado faltante completamente aleatório (MCAR): se a distribuição dos dados faltantes não depende dos valores observados ou perdidos tem-se o chamado mecanismo de perda completamente ao acaso;
- Dados faltantes aleatório (MAR); este mecanismo considerado mecanismo de perda previsível, em que os dados são perdidos por um processo aleatório, quando a probabilidade de não resposta depende dos dados presentes mas não dos ausentes;
- Dados faltantes não aleatório (NMAR): o mecanismo de não resposta chamado de perda não ao acaso é aquele em que a distribuição da matriz “M” depende dos valores faltantes da matriz “X”.

Vale ressaltar que os métodos de imputação se dividem em imputação simples ou única e imputação múltipla. Para mais detalhes ver [Rubin \(1987\)](#), [Dias e Albieri \(1992\)](#) e [Nunes *et al.* \(2009\)](#).

Na sequência será apresentado o método de tratamento de dados faltantes sobre a base de dados real obtida, a qual foi descrita no Capítulo 2.