



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

Suélio Alves de Moura

sob orientação da

Prof<sup>a</sup>. Dr<sup>a</sup>. 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<sup>a</sup>. Dr<sup>a</sup>. 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 amigos. 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

<b>Lista de Abreviaturas</b>	<b>xxi</b>
<b>Lista de Símbolos</b>	<b>xxiii</b>
<b>Lista de Figuras</b>	<b>xxv</b>
<b>Lista de Tabelas</b>	<b>xxix</b>
<b>1 Introdução</b>	<b>1</b>
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
<b>2 Apresentação dos bancos de dados</b>	<b>9</b>
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
<b>3 Teoria da Resposta ao Item (TRI)</b>	<b>23</b>
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
<b>4</b>	<b>Processo de estimação dos parâmetros na TRI para dados binários</b>	<b>41</b>
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
<b>5</b>	<b>Processo de estimação dos parâmetros na TRI - modelos politômicos</b>	<b>63</b>
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
<b>6</b>	<b>Pré-processamento dos dados</b>	<b>79</b>
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
<b>7</b>	<b>Ordenação de periódicos científicos via Teoria da Resposta ao Item</b>	<b>99</b>
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
<b>8</b>	<b>Análises via TRI para uma base de dados com mais indicadores</b>	<b>125</b>
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
<b>9</b>	<b>Considerações finais</b>	<b>143</b>
9.1	Conclusões . . . . .	143
	<b>Referências Bibliográficas</b>	<b>147</b>
	<b>Apêndice A - Apresentação da linguagem <i>Python</i></b>	<b>155</b>
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
<b>Apêndice B - Comandos utilizando o <i>Python</i></b>		<b>159</b>
B.1	Programação . . . . .	159
<b>Apêndice C - Comandos utilizando o <i>Software R</i></b>		<b>163</b>
C.1	Programação . . . . .	163
<b>Apêndice D - Bases de dados desta pesquisa</b>		<b>165</b>
D.1	Bases de dados . . . . .	165
<b>Apêndice E - Algumas análises descritivas das bases de dados desta pesquisa</b>		<b>167</b>
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
<b>Apêndice F - Resultado das simulações para a calibração dos parâmetros dos itens</b>		<b>175</b>
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
<b>Apêndice G - Dados históricos das distribuições dos periódicos por categoria de classificação da CAPES</b>		<b>181</b>
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

<b>Apêndice H - Comandos desenvolvidas em <i>Visual Basic for Applications</i> (VBA)</b>	<b>187</b>
H.1 Programação em VBA . . . . .	187
<b>Apêndice I - Sintaxes desenvolvidas no <i>software Parscale</i>, versão 4.1</b>	<b>191</b>
I.1 Sintaxe Parscale - testando o efeito de DIF . . . . .	191
<b>Apêndice J - Apresentação de algumas árvores de decisão</b>	<b>195</b>
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
<b>Apêndice K - Pareamento das bases de dados desta pesquisa</b>	<b>203</b>
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
<b>Apêndice L - Processo de imputação de dados</b>	<b>221</b>
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  $i$
- $\theta_j$  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
- $\zeta_i$  conjunto de parâmetros dos itens, em que  $i = 1, \dots, I$
- $\mu$  média, medida de posição para o traço latente
- $\sigma$  desvio-padrão, medida de dispersão para o traço latente
- $\tau$  desvio-padrão na escala transformada
- $\pi$  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 <a href="#">Turik (2010)</a> . . . . .	27
3.2	Exemplo de uma Curva Característica do Item - CCI. Adaptado de <a href="#">Andrade et al. (2000)</a> . . . . .	31
3.3	Exemplos de curvas características e de informação de vários itens. Adaptado de <a href="#">Andrade et al. (2000)</a> . . . . .	32
3.4	Curva de informação e erro padrão dos testes A e B. Adaptado de <a href="#">Embretson e Reise (2000, p. 168)</a> . . . . .	35
3.5	Representação gráfica dos Modelos de Escala Gradual e de Resposta Gradual. Adaptado de <a href="#">Andrade et al. (2000)</a> . . . . .	38
4.1	Proporção observada da resposta correta em função da habilidade. Adaptado de <a href="#">Baker e Kim (2004)</a> . . . . .	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) <b>após</b> 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) <b>após</b> o processo de imputação via média. . . . .	94
6.7	Análise de Componentes Principais (PCA) <b>após</b> 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 <b>versus</b> 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 <b>antes</b> 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, <b>após</b> o processo de imputação. . . . .	85
6.8	Análise descritiva dos indicadores bibliométricos <b>após</b> a imputação via média. . . . .	85
6.9	Análise descritiva dos indicadores bibliométricos <b>após</b> 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 ( $z$ ) <b>após</b> a imputação via média. . . . .	89
6.12	Análise descritiva dos indicadores bibliométricos padronizados ( $z$ ) <b>após</b> 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 ( $\chi^2$ ): 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 <b>após</b> 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;

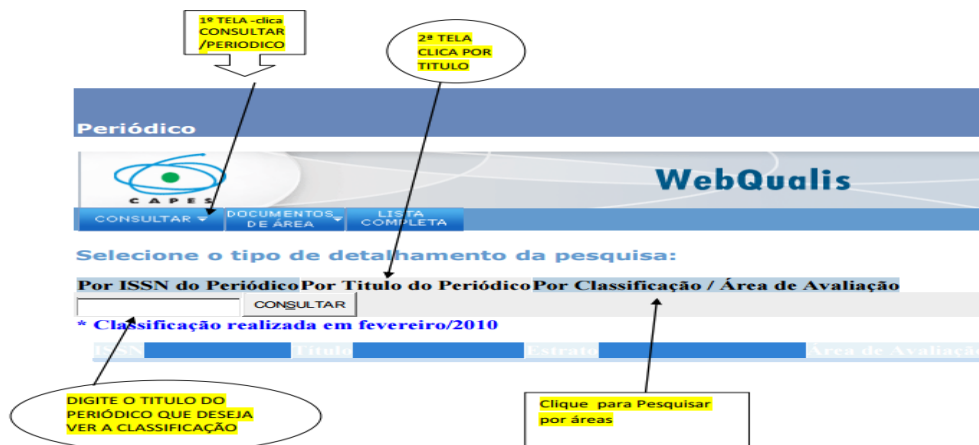
**Meia vida:** O principal aspecto da obsolescência da literatura é a meia-vida, do inglês Half-Life, das citações. O termo “meia-vida das citações” teve origem no conceito físico de meia-vida dos materiais radioativos. O conceito foi transposto para a área de ciência da informação, pela primeira vez, 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;



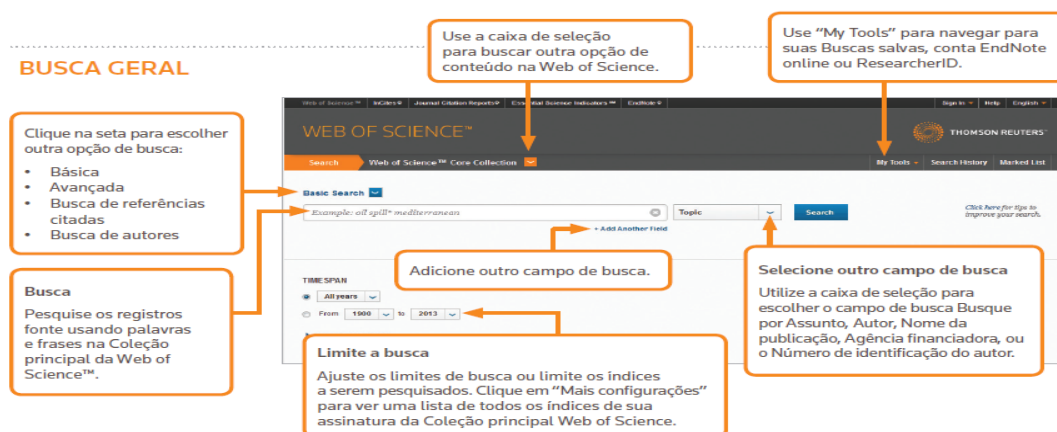
**Average Journal Impact Factor Percentile:** Corresponde ao cálculo da média percentual do fator de impacto sobre a publicação de uma revista;

**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](http://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>

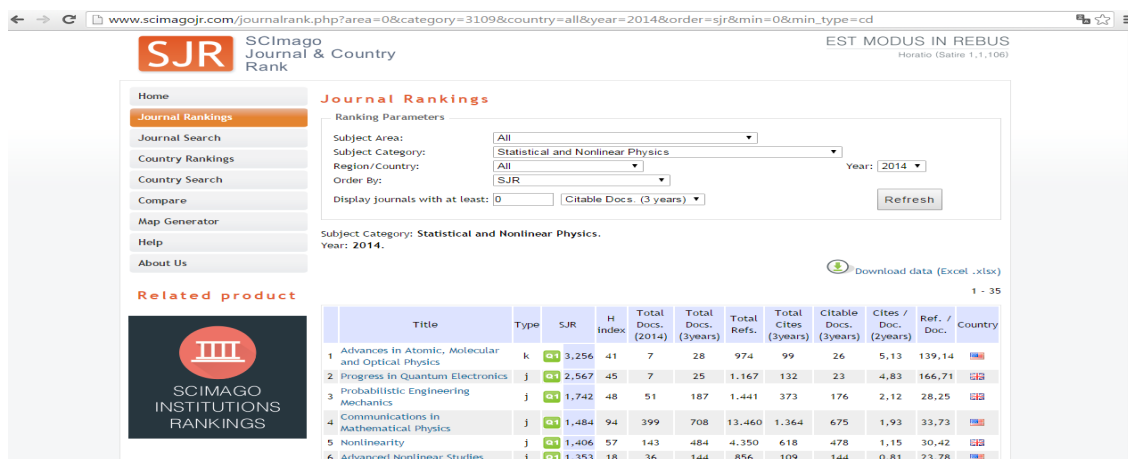


Figura 2.3: Visualização da página na internet ao acessar o endereço eletrônico da Scimago Journal & Country Rank (SJR). Online; último acesso em: 28 de fevereiro de 2016.

## 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](http://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](http://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](http://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  $j$ ,

$\epsilon_{ij}$  : 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).

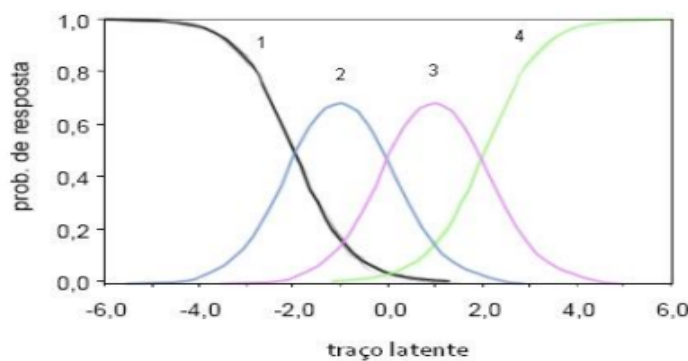
A partir das respostas obtidas de respondentes a um conjunto de itens, a TRI possibilita a estimação dos parâmetros dos itens e dos indivíduos em uma escala de medida. Por exemplo, considere-se que a variável latente de interesse é o nível de satisfação: uma análise feita por meio da TRI pode estimar 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  $i$ ;

$\theta_j$  : representa a habilidade (traço latente) do  $j$ -ésimo indivíduo;

$P(U_{ij} = 1|\theta_j)$  : é a probabilidade de um indivíduo  $j$  com habilidade  $\theta_j$  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;

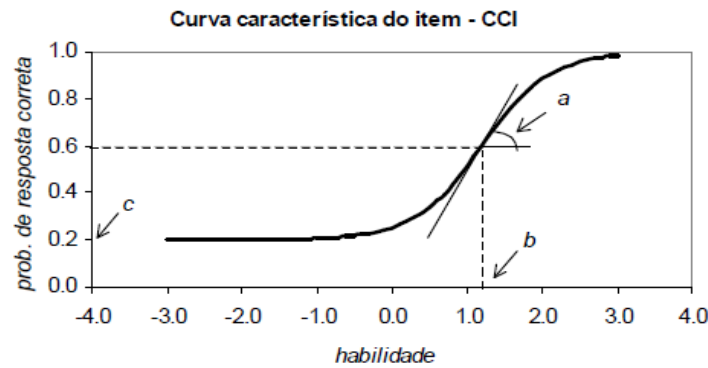
$a_i$  : é o parâmetro de discriminação (ou de inclinação) do item  $i$ , com valor proporcional à inclinação da Curva Característica do Item - CCI no ponto  $b_i$ ;

$c_i$  : 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,  $\theta_j$ .

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  $\theta$ , 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  $\theta$ . 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  $\theta$ , 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

$\theta^*$  : consiste a habilidade (traço latente) na escala transformada;

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

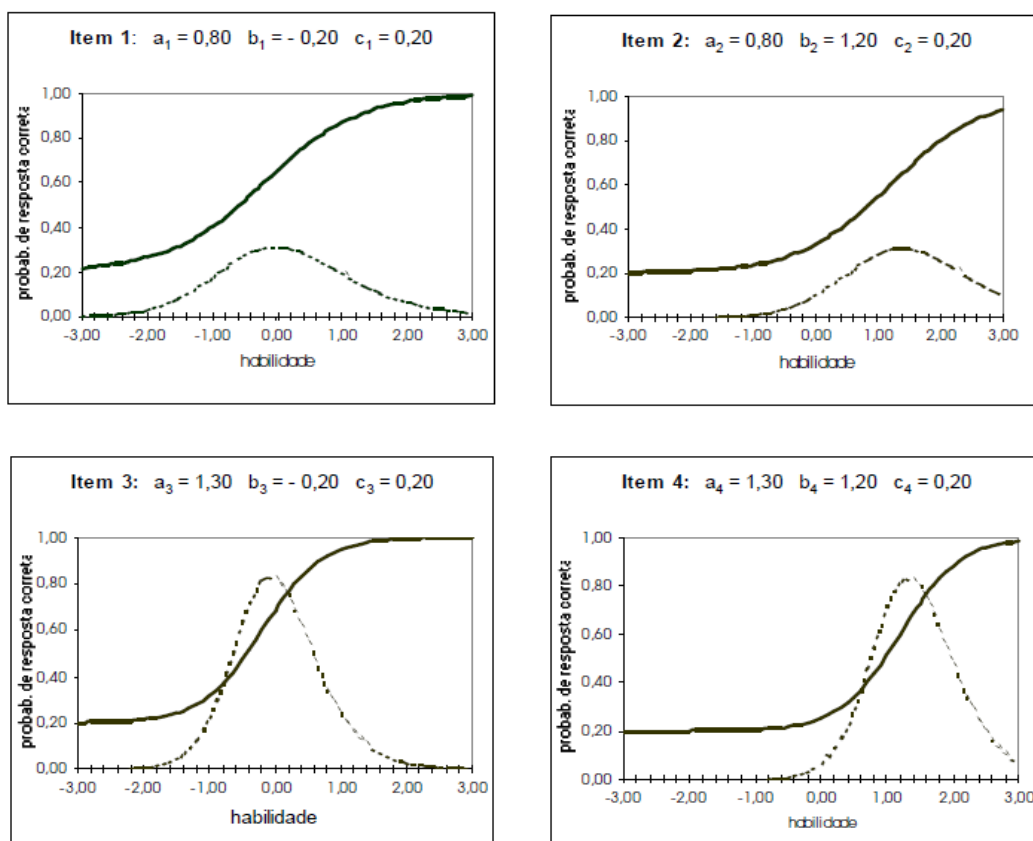
$b^*$  : é o parâmetro de dificuldade na escala transformada;

$\mu$  : é a média desejada na escala transformada;

$\tau$  : 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  $\theta$ ;

$P_i(\theta) = P(X_{ij} = 1|\theta)$  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  $\theta$ ; *(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:

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

É importante notar que as medidas 3.5 e 3.6 de informação dependem do valor do parâmetro  $\theta$ . Desta feita, a amplitude do intervalo de confiança para  $\theta$  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.

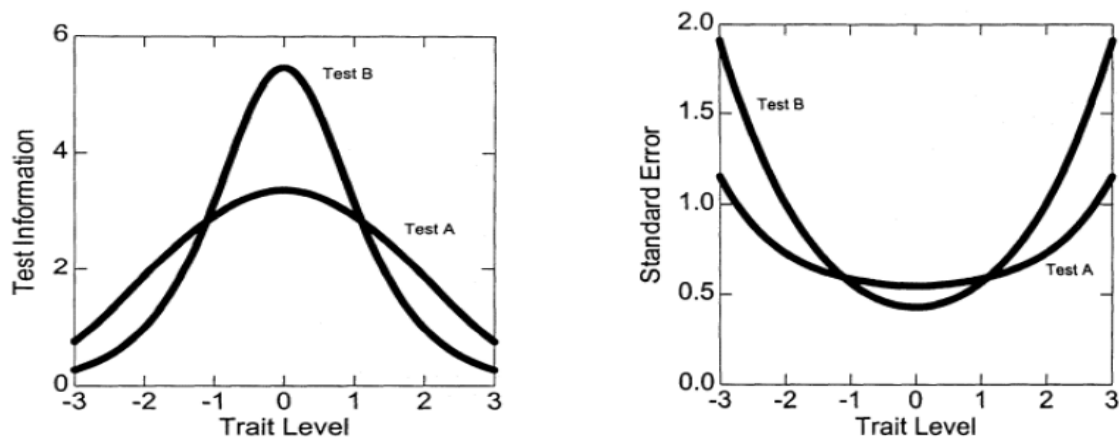
### 3.5.1.3 O modelo logístico de 1 parâmetros (ML1)

**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 Likert<sup>1</sup>.

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

<sup>1</sup>A *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  $\theta_j$ , 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.

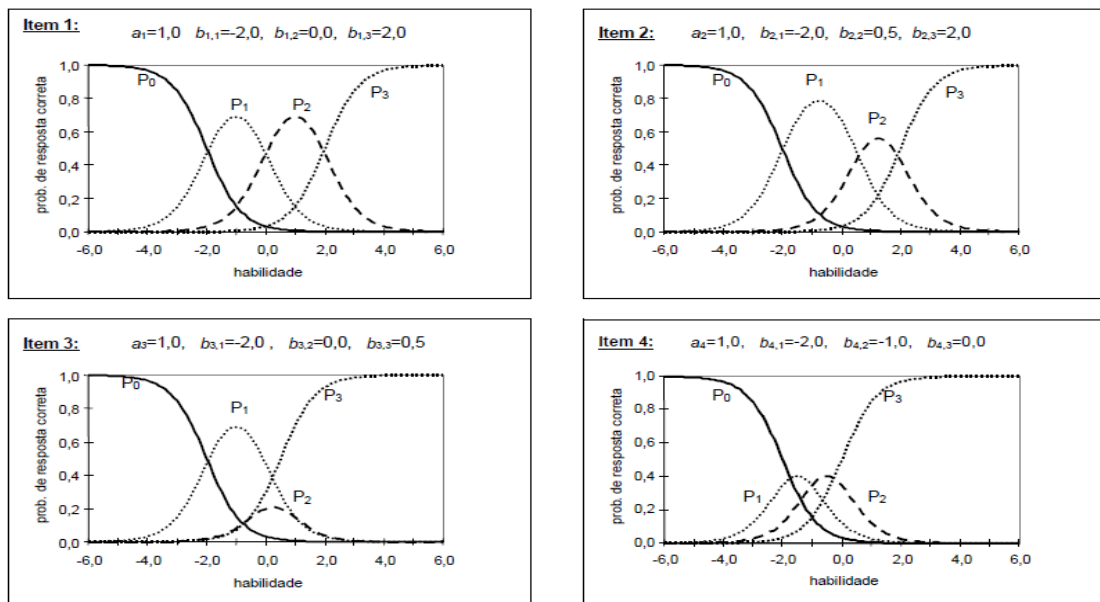
Como  $P_{i,k}^+(\theta_j) - P_{i,k+1}^+(\theta_j) \geq 0$ , então  $d_k - d_{k+1} \geq 0$  e  $d_1 \geq d_2 \geq \dots \geq d_m$ . Sob a suposição de escores equidistantes das categorias de resposta, o parâmetro  $b_{i,k}$  no modelo de escala gradual fica

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

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  $\theta_j$  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
<b>Dicotômico (um grupo ou mais)</b>	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
<b>Politômico (não - dicotômico)</b>		
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

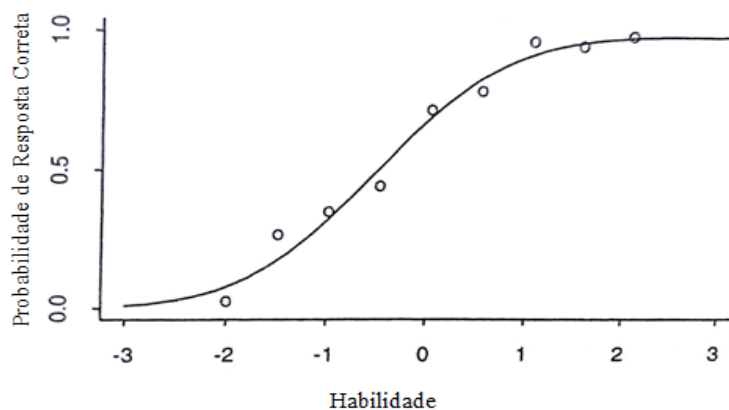
## Processo de estimação dos parâmetros na TRI para dados binários

### 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  $\theta_j$  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  $\mathbf{U}_{..} = (\mathbf{U}_{1.}, \mathbf{U}_{2.}, \dots, \mathbf{U}_{n.})$  o conjunto integral de respostas. De forma similar, representa-se as observações por  $u_{ji}$ ,  $\mathbf{u}_j$ , e  $\mathbf{u}_{..}$ . 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  $\theta$ , ao tomar-se um conjunto de indivíduos com habilidade  $\theta$ , 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  $(\mu, \sigma)$  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  $\zeta_i$ ,  $i = 1, \dots, I$ .

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  $\zeta_i$ . 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  $\zeta_i$  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 obtém-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  $\zeta_i$  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  $\theta$  e não de  $\zeta$ . 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  $\theta_j$  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  $\theta_j$ ,  $j = 1, \dots, n$ , 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  $\theta_j$  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  $\alpha$  e  $\beta$  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,  $\mu$ , e desvio-padrão,  $\sigma$ , 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  $\pi$  composta por  $n$  indivíduos com habilidades  $\theta_j$ ,  $j = 1, \dots, n$ , 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  $\theta_j$  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  $\theta$  no experimento de retirar um indivíduo ao acaso da população  $\pi$  e observar seu parâmetro  $\theta$ . Nesta mesma linha, vale ressaltar que, quando atribui-se uma distribuição de probabilidade para  $\theta$  *não se está aplicando nenhum argumento bayesiano*. A distribuição de  $\theta$  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,  $\theta_j$ ,  $j = 1, \dots, n$ , são realizações de uma variável aleatória  $\theta$ , 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  $\theta$  tem distribuição normal, tem-se que  $\boldsymbol{\eta} = (\mu, \sigma^2)$ , em que  $\mu$  é a média e  $\sigma^2$  a variância das habilidades dos indivíduos de  $\pi$ . 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  $\mathbf{I}$  itens ao total, com 2 possíveis repostas para cada item (0 ou 1), há  $S = 2^{\mathbf{I}}$  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  $\zeta$  na presença das variáveis latentes  $\theta$ . 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  $\theta$  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"> <li>• Adequado para o método de estimação da máxima verossimilhança conjunta com um, dois e três parâmetros.</li> <li>• Capacidade de lidar com respostas omitidas e itens não-alcançados (neste caso, a função de verossimilhança é modificada).</li> </ul>	Wingersky <i>et al.</i> (1982)
BICAL	<ul style="list-style-type: none"> <li>• Adequado para métodos de estimação com um parâmetro.</li> <li>• Fornece testes para o ajuste de itens.</li> </ul>	Wright e Mead (1976)
ANCILLES; OGIVA	<ul style="list-style-type: none"> <li>• Adequado para estimativa conjunta dos parâmetros dos itens e da habilidade em modelo de três parâmetros.</li> <li>• Baseado, parcialmente, no critério do qui-quadrado mínimo.</li> <li>• Fornece boas estimativas para um grande número de examinandos e itens.</li> </ul>	Urry (1976)
PML	<ul style="list-style-type: none"> <li>• Adequado para o método de estimação da máxima verossimilhança conjunta com um parâmetro</li> <li>• Fornece boas estimativas para modelos de multi categorias.</li> </ul>	Gustafsson (1980)
BilogMG de 1997	<ul style="list-style-type: none"> <li>• Adequado para o método de estimação da máxima verossimilhança marginal com um e dois parâmetros.</li> </ul>	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:

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

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:

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

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

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

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

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



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

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

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

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

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

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

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

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  $\zeta^*$ , de modo a contemplar a restrição descrita em (5.22) faz com que o processo iterativo usando o novo vetor de parâmetros  $\Gamma$ , seja computacionalmente mais eficiente do que, por exemplo, a aplicação de técnicas como a dos **Multiplicadores de Lagrange**, usando-se o vetor  $\zeta^*$ .

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-agrupamento<sup>1</sup>, 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  $\theta_j$  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

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

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

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

---

<sup>1</sup>Um 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 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

