



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA E ESTATÍSTICA

Dissertação de Mestrado

MODELOS DE EQUAÇÕES ESTRUTURAIS VERSUS TEORIA DA RESPOSTA AO ITEM

Raul Barbosa Eluan

Orientação: **Profa. Dra. Maria Regina Madruga Tavares**
Co-orientação: **Prof. Dr. Héilton Ribeiro Tavares**

Durante a elaboração deste trabalho o autor recebeu apoio financeiro da CAPES

Belém
2018

Raul Barbosa Eluan

MODELOS DE EQUAÇÕES ESTRUTURAIS VERSUS TEORIA DA RESPOSTA AO ITEM

Dissertação apresentada ao Curso de Mestrado em Matemática e Estatística da Universidade Federal do Pará, como pré-requisito para a obtenção do título de Mestre em Estatística.

Orientação: **Profa. Dra. Maria Regina Madruga Tavares**

Co-orientação: **Prof. Dr. Héilton Ribeiro Tavares**

Belém

2018

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

E37m

Eluan, Raul Barbosa
MODELOS DE EQUAÇÕES ESTRUTURAIS VERSUS TEORIA DA RESPOSTA AO ITEM /
Raul Barbosa Eluan. — 2018
63 f. : il. color

Dissertação (Mestrado) - Programa de Pós-graduação em Matemática e Estatística (PPGME),
Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Belém, 2018.
Orientação: Profa. Dra. Maria Regina Madruga Tavares
Coorientação: Prof. Dr. Héilton Ribeiro Tavares.

1. Variável Latente. 2. Teoria de Resposta ao Item. 3. Modelo de Equação Estrutural. I.
Tavares, Maria Regina Madruga, *orient.* II. Título

CDD 310

Raul Barbosa Eluan

MODELOS DE EQUAÇÕES ESTRUTURAIS VERSUS TEORIA DA
RESPOSTA AO ITEM

Esta Dissertação foi julgada e aprovada para a obtenção do grau de Mestre em Estatística, no Programa de Pós-Graduação em Matemática e Estatística da Universidade Federal do Pará.

Belém, 03 de abril de 2018

APROVADO

Banca Examinadora

Maria Regina Madruga Tavares

Profa. Dra. Maria Regina Madruga Tavares
PPGME/UFPA.
Orientador(a)

Heliton Ribeiro Tavares

Prof. Dr. Heliton Ribeiro Tavares
PPGME/UFPA.
Co-orientador(a)

Paulo Cerqueira dos Santos Junior

Prof. Dr. Paulo Cerqueira dos Santos Junior
Universidade Federal Rural da Amazônia - UFRA.
Examinador(a) Externo

Aos meus pais, meus amigos, minha namorada e meus professores.

Agradecimentos

À Deus por ter me mostrado o caminho e ter me dado força e saúde para sempre persistir e nunca desistir.

A minha orientadora, meu co-orientador e todos os professores que estavam sempre dispostos a me aconselhar e compartilhar ideias.

Aos meus pais que sempre me apoiaram e estavam do meu lado em todos os momentos.

Aos colegas e amigos do PPGME, demais colegas da UFPA e todos meus amigos que me ajudaram com conselhos, experiências e compartilhamento de ideias

À Amanda Martins por todo o apoio, amor e paciência durante o período deste trabalho.

Finalmente, gostaria de agradecer à UFPA pelo ensino gratuito de qualidade, ao PPGME e à CAPES, sem os quais essa dissertação dificilmente poderia ter sido realizada e a todos mais que eu não tenha citado nesta lista de agradecimentos, mas que de uma forma ou de outra contribuíram não apenas para a minha dissertação, mas também para meu amadurecimento.

"Para conseguir algo neste mundo, é preciso sacrificar algo de valor equivalente."

Edward Elric

"A vitória perde todo significado se você não puder compartilhar com quem ama, só assim somos vencedores."

Yugi Muto

”Quando alguém te machuca, você sente ódio e quando você machuca alguém, você fica amargo, mas se sente culpado também. Conhecer a dor nos ajuda a crescer, a madurecer e crescer significa ser capaz de pensar e tomar decisões próprias.”

Naruto Uzumaki

Resumo

ELUAN, Raul Barbosa. MODELOS DE EQUAÇÕES ESTRUTURAIIS VERSUS TEORIA DA RESPOSTA AO ITEM. 2018. Dissertação para obtenção do grau de mestre em Estatística, UFPA, Belém - Pará, Brasil.

Neste trabalho pretende-se comparar as metodologias *Teoria da Resposta ao Item (TRI)* e *Modelos de Equações Estruturais (MEE)* com o intuito identificar as condições de equivalência. Almeja-se ainda verificar qual o melhor método para realizar estimações de características de itens e de variáveis latentes em função do número de itens e do tamanho amostral. Para demonstrar a equivalência dos métodos, foi considerado o Modelo Logístico de 2 Parâmetros da TRI. O critério de determinação do melhor método foi menor Erro Quadrático Médio e o menor Tempo Computacional. Todas as etapas computacionais foram desenvolvidas no *software R*.

Palavras-chave: Variável Latente. Teoria da Resposta ao Item. Modelos de Equações Estruturais.

Abstract

ELUAN, Raul Barbosa. STRUCTURAL EQUATION MODELING VERSUS ITEM RESPONSE THEORY. 2018. Dissertation to obtain master's degree in Statistics, UFPA, Belém - Pará, Brazil.

This work intends to compare the methodologies of *Item Response Theory (IRT)* and *Structural Equation Modeling (SEM)* in order to identify equivalence conditions. It is also aims to verify which method is better to perform estimation of item characteristics and latent variables in function of the number of items and the sample size. To demonstrate the equivalence of methods, it was considered for the IRT Two-Parameters Logistic Model. The criteria to define the better method was less Mean Squared Error and lower Computational Time. All computational steps were developed in *R software*.

Keywords: Latent Variable. Item Response Theory. Models Equation Modeling.

Sumário

Agradecimentos	vi
Resumo	ix
Abstract	x
Lista de Tabelas	xiii
Lista de Figuras	xiv
1 Introdução	1
1.1 Aspectos gerais	1
1.2 Justificativa e Importância da Dissertação	3
1.3 Objetivos	4
1.3.1 Objetivo Geral	4
1.3.2 Objetivos Específicos	4
1.4 Estrutura da Dissertação	4
2 A Teoria da Resposta ao Item	6
2.1 Introdução	6
2.2 Estimacão: Parâmetros da TRI	8
2.2.1 Introdução	8
2.2.2 Estimacão dos parâmetros dos itens	10
2.2.3 Estimacão das Habilidades	11
3 Modelos de Equacões Estruturais	13
3.1 Introdução	13
3.1.1 Contexto Histórico	13
3.1.2 Conceitos Básicos	14
3.2 A relacão entre Análise Fatorial e TRI	16
3.2.1 Uma Interpretacão TRI do Modelo de Um Fator	16
3.2.2 Modelos Fatoriais para Dados Dicotômicos: o Limiar e o Modelo de Ogiva Normal	17
3.2.3 Modelos de Resposta ao Item Logístico e Ogiva Normal	21
4 Aspectos Computacionais: o pacote MIRT e o pacote LAVAAN do programa R	22

4.1	Pacote MIRT	22
4.1.1	Introdução	22
4.1.2	O Modelo Logístico de 2 parâmetros da TRI	23
4.2	Pacote LAVAAN	24
4.2.1	Introdução	24
4.2.2	Por que precisamos do Lavaan?	25
4.2.3	Especificando Modelos em Lavaan	26
4.3	Resultados Obtidos Com Uma Amostra	29
4.4	Resultados do método da TRI - Amostra 1	30
4.5	Resultados do método da MEE - Amostra 1	33
5	Estudo de Simulação	37
5.1	Introdução	37
5.2	Resultados dos EQMs da TRI e da MEE	40
6	Conclusões	46
6.1	Concluições Obtidas neste Trabalho	46
6.2	Pespectivas para Trabalhos Futuros	47
A	Desenvolvimento das questões do Capítulo 2	48
A.1	Expressões da seção 2.2.2	48
A.1.1	Algoritmo Newton-Raphson	50
A.2	Expressões da seção 2.2.3	53
A.2.1	Algoritmo de Newton-Raphson	54
B	Algoritmo usado na simulação	56
B.1	Geração do Bando de Dados	56
B.2	MIRT	57
B.3	LAVAAN	59
	Referências Bibliográficas	62

Lista de Tabelas

5.1 Erro Quadrático Médio	38
-------------------------------------	----

Lista de Figuras

2.1	Exemplo de uma comparação da Função de distribuição Normal Acumulada $N(0,1)$ com a Função Logística.	7
2.2	Exemplo de uma Curva Característica do Item.	7
3.1	Exemplo de Modelo de Equação Estrutural (Pilati and Laros, 2007).	15
4.1	Saída das estimativas de φ , utilizando o pacote <i>Lavaan</i> do <i>software R</i>	28
4.2	Saída das estimativas de μ , utilizando o pacote <i>Lavaan</i> do <i>software R</i>	29
4.3	Cruzamento dos Parâmetros de Discriminação estimados pelo Mirt com os verdadeiros.	31
4.4	Diferença dos Parâmetros de Discriminação Verdadeiros com os Estimados pelo Mirt.	31
4.5	Cruzamento dos Parâmetros de Dificuldade estimados pelo Mirt com os verdadeiros.	32
4.6	Diferença dos Parâmetros de Dificuldade Verdadeiros com os Estimados pelo Mirt.	32
4.7	Cruzamento das Habilidades estimadas pelo Mirt com as verdadeiras.	33
4.8	Diferença das Habilidades Verdadeiras com as Estimadas pelo Mirt.	33
4.9	Cruzamento dos Parâmetros de Discriminação estimados pelo Lavaan com os verdadeiros.	34
4.10	Diferença dos Parâmetros de Discriminação Verdadeiros com os Estimados pelo Lavaan.	34
4.11	Cruzamento dos Parâmetros de Dificuldade estimados pelo Lavaan com os verdadeiros.	35
4.12	Diferença dos Parâmetros de Dificuldade Verdadeiros com os Estimados pelo Lavaan.	35
4.13	Cruzamento das Habilidades estimadas pelo Lavaan com as verdadeiras.	36
4.14	Diferença das Habilidades Verdadeiras com as Estimadas pelo Lavaan.	36
5.1	EQMs do parâmetro de discriminação com tamanho de itens igual a 10.	40
5.2	EQMs do parâmetro de discriminação com tamanho de itens igual a 20.	40
5.3	EQMs do parâmetro de discriminação com tamanho de itens igual a 50.	41
5.4	EQMs do parâmetro de discriminação com tamanho de itens igual a 100.	41
5.5	EQMs do parâmetro de dificuldade com tamanho de itens igual a 10.	42
5.6	EQMs do parâmetro de dificuldade com tamanho de itens igual a 20.	42
5.7	EQMs do parâmetro de dificuldade com tamanho de itens igual a 50.	42

5.8	EQMs do parâmetro de dificuldade com tamanho de itens igual a 100. . . .	43
5.9	EQMs das habilidades com tamanho de itens igual a 10.	43
5.10	EQMs das habilidades com tamanho de itens igual a 20.	44
5.11	EQMs das habilidades com tamanho de itens igual a 50.	44
5.12	EQMs das habilidades com tamanho de itens igual a 100.	44

Capítulo 1

Introdução

1.1 Aspectos gerais

Na área de pesquisa estatística o termo variável latente está associado a quantidades que não podem ser medidas/observadas diretamente, ou seja, são variáveis que podem ser unicamente estimadas através de algum método estatístico apropriado (de Andrade et al., 2000). Exemplos de variáveis latentes são: nível socioeconômico, qualidade de vida, proficiência (conhecimento) em alguma área, intensidade de uma dor, satisfação de uma pessoa, dentre outros exemplos.

Os trabalhos científicos que abordam a estimação de variáveis latentes vêm sendo desenvolvidos desde o século XVII em diversas áreas do conhecimento. Dentre essas áreas podemos citar: Psicologia, Economia, Educação e Ciências Sociais (Marôco, 2010).

No cotidiano, as variáveis latentes se encontram em ampla aplicação na vida das pessoas. Uma simples expressão do tipo “como você vai?”, até a indicação de um aluno inteligente/habilidoso são exemplos disso. Esses conceitos abstratos não permitem sua medição direta. As variáveis latentes podem ser dadas como características de determinado evento.

Em resumo, o que esses exemplos ilustram é a prática comum entre os seres humanos de explicar, entender, e às vezes, prever eventos, com base no papel dos conceitos que não são diretamente observáveis. Além disso, as variáveis latentes fornecem um grau de abstração que nos permite descrever as relações entre uma classe de eventos ou variáveis que compartilham algo em comum (Bollen, 2002).

Na prática não existe uma única maneira de se trabalhar com variáveis latentes. Existem vários métodos estatísticos com o intuito de realizar estimações de variáveis latentes. Dentre eles podemos citar: Análise de Classe Latente, Teoria da Resposta ao Item, Modelos de Curva Latente, Análise Fatorial, Modelo de Equação Estrutural e Regressão com Variáveis Latentes.

A Teoria da Resposta ao Item (TRI) é um conjunto de modelos estatísticos com capacidade de realizar a estimação de variáveis latentes de indivíduos, através de itens elaborados para avaliá-las. As respostas aos itens são, em geral, categóricas com duas ou mais opções de resposta, podendo também ser contínuas. Um dos trabalhos primordiais em relação ao uso desse método pode ser visto em Lord (1952); e mais recentemente em Andrade et al. (2000), Pasquali and Primi (2003), Soares (2005), Glockner-Rist and Hoijsink (2003), dentre outros.

É comum encontramos aplicação desse método em áreas como avaliação educacional, pesquisa de opinião, marketing, entre outras. A TRI é utilizada em uma vasta gama de países. No Brasil, sua importância é facilmente vista em Programas de Avaliação como, o Sistema de Avaliação da Educação Básica (SAEB) e o Exame Nacional do Ensino Médio (ENEM). Para detalhes, ver Tavares et al. (2016).

Toda vez em que um teste de avaliação (aplicado para um grupo de indivíduos) é elaborado, deseja-se saber certas características dos itens, e/ou dos respondentes daquele teste. Em relação a um teste avaliativo, a TRI tem a capacidade de realizar estimação de características (parâmetros) de itens e das variáveis latentes associadas, tais como: discriminação do item, dificuldade do item, acerto ao acaso do item e habilidades dos indivíduos (Andrade et al., 2000). Estas características serão formalmente apresentadas nos próximos capítulos.

Um outro método que tem se proliferado na ciência psicológica, e é capaz de realizar a estimação de variáveis latentes, é o Modelo de Equações Estruturais (MEE). Esse método é entendido como uma mistura de análise fatorial e análise de regressão, que permite aos pesquisadores testar estruturas fatoriais de instrumentos de medida psicométrica, por meio da análise fatorial confirmatória (Pilati and Laros (2007)).

O MEE não apenas permite o teste confirmatório da estrutura psicométrica de escalas de medida, mas também pode ser utilizado para analisar relações explicativas entre múltiplas variáveis simultaneamente, sejam essas latentes ou observadas.

Podemos encontrar esse método nas áreas de Psicologia e Ciências Sociais e suas aplicações em alguns trabalhos, como: Glockner-Rist and Hoijsink (2003), Brei and Liberali Neto (2006), Pilati and Laros (2007), Rosseel (2012), Campana et al. (2009), dentre outros.

No contexto brasileiro, desde meados da década passada começou-se a notar um

aumento de interesse sobre o uso da técnica de MEE como estratégia de análise de dados e teste de modelos teóricos (Pilati and Laros (2007)).

O MEE utiliza de certas tarefas comportamentais com o intuito de estimar certas características latentes. Glockner-Rist and Hoijsink (2003) demonstraram haver uma certa equivalência entre o método da TRI e o método da MEE. Essa equivalência possibilitou a aplicação dos MEE a testes avaliativos, e com isso realizar estimações de características associadas ao teste. Nesta mesma equivalência o MEE foi capaz de realizar estimações das variáveis latentes e também características dos itens, tais como a discriminação do item, dificuldade do item, além das habilidades dos indivíduos. A característica conhecida como (parâmetros de) acerto ao acaso não está inserida nessa equivalência entre os modelos.

1.2 Justificativa e Importância da Dissertação

Em um procedimento cirúrgico, é desejável realizar o melhor método para uma determinada situação. No caso de métodos que realizam estimações de variáveis latentes, a história não é diferente. É desejável identificar o método que seja capaz de realizar a estimação de forma mais precisa, ou seja, que possua o menor erro em relação a variável de interesse a ser estimada. Mas como podemos saber qual o melhor método? Existe algum método mais preciso que outros? Ou cada situação tem um método com melhor precisão a ser utilizado?

São perguntas como essas que fazem surgir o estudo de vários métodos para diversas situações, com o intuito de verificar qual método que melhor se adapta à exigência de cada situação.

Utilizando a equivalência proposta por Glockner-Rist and Hoijsink (2003), foi necessário estudar, através de dados simulados, as modelagens dos métodos da TRI e MEE e realizar a aplicação de ambos para várias combinações de tamanhos amostrais e tamanhos de itens.

Como o modelo proposto por Glockner-Rist and Hoijsink (2003) não utiliza a característica do acerto ao acaso, o modelo da TRI utilizado foi o modelo logístico de dois parâmetros (ML2).

Outro ponto importante, que também é desejável em um método de estimação, é o seu custo computacional, ou seja, quanto tempo de processamento é necessário para que o processo de estimação seja executado. Esta preocupação é necessária pois não podemos

utilizar um tempo tão grande para executar este processamento, que inviabilize o sentido da pesquisa, fazendo com que um tempo excessivo, acabe atrapalhando o restante do estudo.

Sendo assim torna-se necessário a verificação do tempo dos processos de estimação dos métodos da TRI e da MEE. Em síntese, o objetivo deste trabalho será comparar as estimações propostas pelos dois métodos, TRI e MEE, observando o tempo de execução de cada um. Assim, iremos verificar qual o método que obteve a melhor estimativa e o menor tempo computacional, para diferentes tamanhos de amostra e tamanhos de itens.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo principal desta dissertação consiste em comparar os processos de estimação dos métodos da TRI e da MEE, e aplicar esses processos em dados simulados, com o intuito de apontar vantagens e desvantagens de cada um, para diferentes tamanhos de amostras, tamanhos de testes e suas combinações.

1.3.2 Objetivos Específicos

- Estudar e apresentar a modelagem dos métodos da TRI e MEE, demonstrando em que situação há equivalência no desempenho dos dois métodos;
- Aplicar o processo de estimação pelos métodos TRI e MEE em um banco de dados simulados e verificar os resultados das estimativas de cada parâmetro;
- Realizar a comparação dos resultados obtidos dos dois métodos através de um amplo estudo de simulação.

1.4 Estrutura da Dissertação

Este trabalho encontra-se dividido em 6 capítulos, a saber:

- No Capítulo 1 é feita uma introdução aos conceitos de TRI e MEE, são abordados os aspectos gerais, justificativa e importância do trabalho, os objetivos geral e os específicos, e o sumário da dissertação.

- No Capítulo 2 é realizada a descrição conceitual do método da TRI, apresentando, o processo de estimação de cada parâmetro de item e das habilidades.
- No Capítulo 3 é realizada a descrição conceitual do método da MEE, mostrando seu contexto histórico e conceitos básicos. Também é apresentada a demonstração da equivalência do método da MEE com o da método TRI, com os respectivos processos de estimação dos parâmetros.
- No Capítulo 4 serão apresentados os aspectos computacionais para o uso do método da TRI e do método da MEE, abordando o programa R-Project utilizado e os pacotes do mesmo. Também foi abordado uma prévia dos resultados, utilizando análise gráfica em apenas um tamanho de amostra.
- No Capítulo 5 serão apresentados os resultados para algumas combinações de tamanho de amostra e tamanho de testes, assim como a análise separada dos tamanhos de amostra dos tamanhos de testes, utilizando análise gráfica.
- No Capítulo 6 serão abordadas as conclusões obtidas com a apresentação deste trabalho, assim como as perspectivas pra trabalhos futuros.

Capítulo 2

A Teoria da Resposta ao Item

2.1 Introdução

Este capítulo dedica-se a uma descrição conceitual do método da Teoria da Resposta ao Item (TRI). Aqui será apresentado basicamente sobre alguns conceitos da TRI, o modelo utilizado no método, e parâmetros com suas estimativas. O modelo mais utilizado na TRI é o modelo logístico unidimensional de três parâmetros (ML3P, ou 3PL em inglês):

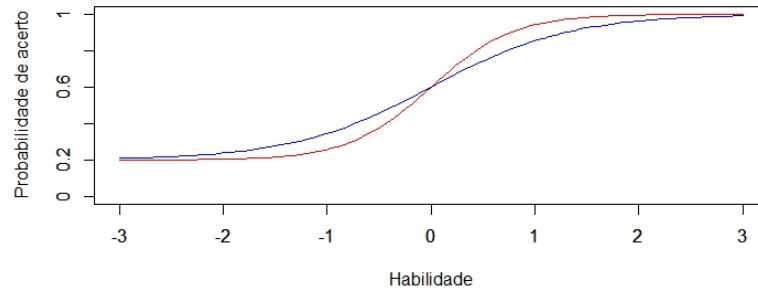
$$P(X_{ij} = 1|\theta_i, a_j, b_j, c_j) = c_j + (1 - c_j) \frac{1}{1 + e^{-Da_j(\theta_i - b_j)}}, \quad (2.1)$$

onde:

- $i = 1, \dots, n$ representa o número de indivíduos.
- $j = 1, \dots, J$ representa o número de itens.
- X_{ij} é uma variável dicotômica que assume os valores 1 quando o indivíduo i responde corretamente ao item j , ou 0 quando indivíduo i não responde corretamente ao item j .
- θ_i representa a habilidade do i -ésimo indivíduo.
- c_j é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item j .
- b_j é o parâmetro de dificuldade do item j , medido na mesma escala da habilidade.
- a_j é o parâmetro de discriminação do item j , com valor proporcional à inclinação da Curva Característica do Item (CCI) no ponto b_j .
- D é o fator de escala, constante igual a 1. Utiliza-se o valor 1.702 quando deseja-se que a função logística forneça resultados semelhantes ao da função ogiva normal.

A Figura (2.1) é um exemplo de comparação da Função de distribuição Normal Acumulada $N(0,1)$ com a Função Logística.

Figura 2.1 *Exemplo de uma comparação da Função de distribuição Normal Acumulada $N(0,1)$ com a Função Logística.*

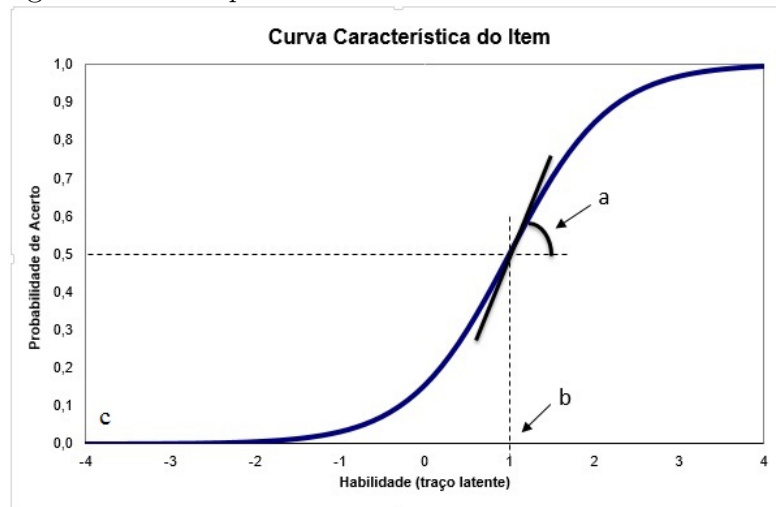


Este método é aplicado em testes com itens de múltipla escolha de duas ou mais opções de respostas (ABCDE por exemplo). Frequentemente as respostas a esses itens são dicotomizadas para 0 ou 1, certo ou errado, sim ou não, etc.

A expressão em (2.1) tem como significado a probabilidade do indivíduo i , com habilidade θ_i , responder corretamente o item j . Também pode ser vista como a proporção esperada de acertos ao item j dentre todos os indivíduos da população com habilidade θ_i . Isto é chamado de Função de Resposta do Item (FRI).

A relação existente entre $P(X_{ij} = 1|\theta_i)$ e os parâmetros do modelo, é apresentada na Figura (2.2), que é chamada de Curva Característica do Item (CCI).

Figura 2.2 *Exemplo de uma Curva Característica do Item.*



A CCI tem uma forma de “S” de modo que quanto maior o valor da habilidade θ_i ,

maior é a probabilidade de acertar o item. O parâmetro a_j é determinado pela inclinação da curva em torno do ponto b_j . Dessa forma, não são esperados valores negativos para a_j , pois isso indicaria que a probabilidade de responder corretamente o item diminui com o aumento da habilidade. Quanto maior o valor de a_j , maior será o nível de discriminação do item, fazendo que o item diferencie (discrimine) melhor os indivíduos que possuem habilidade abaixo do parâmetro b_j e daqueles que possuem habilidade acima do parâmetro b_j .

A diferença $\theta_i - b_j$ em (2.1) nos informa que o parâmetro b_j é medido na mesma escala da habilidade, e representa a habilidade necessária para uma probabilidade de acerto igual a $(1 + c_j)/2$. Quanto maior o valor do parâmetro b_j , mais difícil é o item, e vice-versa.

O parâmetro c_j representa a probabilidade de um indivíduo com baixa habilidade responder corretamente o item. Em certas ocasiões não faz sentido considerarmos um *acerto casual*, tal como em itens abertos ou questionários de escolas de alternativas, e neste caso adota-se $c_j = 0$, e o b_j representará o ponto na escala da habilidade onde a probabilidade de acertar o item é 0,5. Ainda, frequentemente representa-se por $\zeta_j = (a_j, b_j, c_j)$ o vetor de parâmetros do item j . Essas definições podem ser encontradas com maior detalhe em Andrade et al. (2000).

2.2 Estimação: Parâmetros da TRI

2.2.1 Introdução

A estimação dos parâmetros dos itens e das habilidades dos respondentes é uma das etapas mais importantes do método da TRI. A probabilidade de um respondente acertar um determinado item, depende somente da habilidade desse indivíduo e dos parâmetros que caracterizam o item. Porém, em geral, ambos são desconhecidos. Apenas as respostas dos indivíduos aos itens do teste são conhecidas.

A estimação dos parâmetros dos itens e/ou das habilidades dos repondentes pode ser dividido em três situações: a primeira situação ocorre quando os parâmetros dos itens são conhecidos, tendo apenas que estimar as habilidades. A segunda situação ocorre quando as habilidades dos respondentes são conhecidas, tendo que estimar os parâmetros dos itens. E a terceira situação ocorre quando tanto quando os parâmetros dos itens e as

habilidades dos respondentes são desconhecidas, sendo necessário realizar a estimação de ambas simultaneamente.

Nesse trabalho será abordada a situação mais comum, em que tanto os parâmetros dos itens quanto as habilidades dos respondentes são desconhecidas. Para isso, existem duas abordagens: estimação conjunta, parâmetros dos itens e habilidades; ou em duas etapas, primeiro a estimação dos parâmetros dos itens e, posteriormente, das habilidades Andrade et al. (2000). Aqui será utilizada a abordagem estimação conjunta em duas etapas através do método da Máxima Verossimilhança Marginal (MVM).

No método de estimação que será apresentado a seguir, algumas notações e suposições, serão necessárias para o desenvolvimento do modelo. Em particular, seja θ_i a habilidade do indivíduo i e X_{ij} a variável aleatória que representa a resposta do indivíduo i ao item j , com

$$X_{ij} = \begin{cases} 1, & \text{resposta correta} \\ 0, & \text{resposta incorreta} \end{cases}$$

Seja $\mathbf{X}_i = (\mathbf{X}_{i1}, \mathbf{X}_{i2}, \dots, \mathbf{X}_{iJ})$ o vetor aleatório de respostas do indivíduo i e $\mathbf{X}_{..} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$ o conjunto completo de respostas. De forma similar, os valores observados serão representados por \mathbf{x}_{ij} , \mathbf{x}_i e $\mathbf{x}_{..}$. $\theta = (\theta_1, \dots, \theta_n)$ representará o vetor de habilidade dos n indivíduos e $\zeta = (\zeta_1, \dots, \zeta_J)$ o conjunto de parâmetros dos itens.

As duas principais suposições que serão usadas em todo o restante deste texto, são as seguintes:

(S1) as respostas oriundas de indivíduos diferentes são independentes;

(S2) os itens são respondidos de forma independente por cada indivíduo, fixada sua habilidade.

A estimação realizada neste trabalho foi pelo Método da Máxima Verossimilhança Marginal. Como o método não possui solução analítica, então será necessário uma aplicação numérica para a realização desse método. A aplicação realizada nesse trabalho foi o algoritmo de *Newton-Raphson* (Isaac, 1966). No caso da estimação conjunta, o processo será composto por dois estágios:

Estágio 1: Como as habilidades não são conhecidas, foi preciso fazer uso de algum artifício de forma que a verossimilhança não seja mais função das habilidades. O artifício utilizado foi marginalizar a função de verossimilhança. Após realizado o artifício será possível obter as estimativas para $\zeta_j, j = 1, \dots, J$.

Estágio 2: Condicionado nas estimativas de ζ , obtidas no Estágio 1, e tratando esses parâmetros como conhecidos, obtém-se estimativas para as habilidades $\theta_i, i = 1, \dots, n$.

A seguir será mostrado a estimação dos parâmetros dos itens, levando em consideração a estratégia do Estágio 1.

2.2.2 Estimação dos parâmetros dos itens

Aqui será mostrado a estimação dos parâmetros dos itens pelo método da máxima verossimilhança marginal que pode ser encontrado em Andrade et al. (2000). Como já foi mencionado anteriormente, para realizar a estimação através desse método, será necessário utilizar um artifício de forma que a verossimilhança não seja mais função das habilidades, pois as habilidades não são conhecidas. O artifício utilizado consistiu em marginalizar a verossimilhança, integrando-a com relação a distribuição da habilidade.

De forma geral consideremos que as habilidades, $\theta_i, i = 1, \dots, n$, são realizações de uma variável aleatória θ com distribuição contínua e função densidade de probabilidade (*fdp*) $g(\theta|\eta)$, duplamente diferenciável, com as componentes de η conhecidas e finitas. Para o caso em que θ tenha uma distribuição Normal, temos $\eta = (\mu, \sigma^2)$, onde μ é a média e σ^2 é a variância das habilidades dos indivíduos. Portanto, se desejarmos que os itens sejam estimados na métrica (0,1), devemos adotar $\mu = 0$ e $\sigma = 1$.

Pela independência entre as respostas de diferentes indivíduos (S1) e a independência local (S2), podemos escrever a verossimilhança, $L(\zeta) = P(\mathbf{X}_{..} = \mathbf{x}_{..}|\theta, \zeta)$, como

$$L(\zeta) = \prod_{i=1}^n P(\mathbf{X}_i = \mathbf{x}_i|\theta_i, \zeta) = \prod_{i=1}^n \prod_{j=1}^J P(\mathbf{X}_{ij} = \mathbf{x}_{ij}|\theta_i, \zeta_j), \quad (2.2)$$

em que, na expressão (2.2) temos que a distribuição de \mathbf{X}_{ij} só depende de ζ através de ζ_j .

A probabilidade marginal de \mathbf{X}_i é dada por

$$\begin{aligned} P(\mathbf{x}_i|\zeta, \eta) &= \int_{\mathbb{R}} P(\mathbf{x}_i|\theta, \zeta, \eta) \mathbf{g}(\theta|\eta) \mathbf{d}\theta \\ &= \int_{\mathbb{R}} P(\mathbf{x}_i|\theta, \zeta) \mathbf{g}(\theta|\eta) \mathbf{d}\theta, \end{aligned} \quad (2.3)$$

onde na última igualdade usamos que a distribuição de \mathbf{X}_i não é função de η e \mathbb{R} representa o conjunto dos números reais. Usando a independência entre as respostas de diferentes indivíduos, podemos escrever a probabilidade associada ao vetor de respostas $\mathbf{X}_{..}$ como

$$P(\mathbf{x}_{..}|\zeta, \eta) = \prod_{i=1}^n P(\mathbf{x}_i|\zeta, \eta). \quad (2.4)$$

Embora a verossimilhança possa ser escrita como em (2.4), tem sido frequente utilizar a abordagem de *Padrões de Resposta*. Como temos J itens no total, com duas possíveis respostas para cada item (0 ou 1 por exemplo), há $S = 2^J$ possíveis respostas (*Padrões de Resposta*). Quando o número de indivíduos é grande com relação ao número de itens, podem haver vantagens computacionais em trabalhar com o número de ocorrências dos diferentes padrões de resposta. Neste sentido, será considerado este raciocínio. O índice i não representará um indivíduo, mas sim um padrão de resposta.

Seja r_i o número de ocorrências distintas do padrão de resposta i , e ainda $s \leq \min(n, S)$ o número de padrões de resposta com $r_i \geq 0$. Segue disso que

$$\sum_{i=1}^s r_i = n. \quad (2.5)$$

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

$$L(\zeta, \eta) = \frac{n!}{\prod_{i=1}^s r_i!} \prod_{i=1}^s [P(\mathbf{x}_i|\zeta, \eta)]^{r_i}, \quad (2.6)$$

e, portanto, a log-verossimilhança é

$$\log L(\zeta, \eta) = \log \left\{ \frac{n!}{\prod_{i=1}^s r_i!} \right\} + \sum_{i=1}^s r_i \log [P(\mathbf{x}_i|\zeta, \eta)]. \quad (2.7)$$

As equações de estimação para os parâmetros dos itens são dadas por

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_j} = 0; j = 1, \dots, J, \quad (2.8)$$

A próxima seção mostrará a estimação das habilidades.

2.2.3 Estimação das Habilidades

Nesta seção será tratado da estimação das habilidades quando os parâmetros dos itens são conhecidos.

Pela independência entre as respostas de diferentes indivíduos (S1) e a independência local (S2), podemos escrever a log-verossimilhança como função de θ ,

$$\log L(\theta) = \sum_{i=1}^n \sum_{j=1}^J \{ \mathbf{x}_{ij} \log P_{ij} + (1 - \mathbf{x}_{ij}) \log Q_{ij} \}. \quad (2.9)$$

O EMV de θ_i é o valor que maximiza a verossimilhança, ou equivalente, é a solução da equação

$$\frac{\partial \log L(\theta)}{\partial \theta} = 0, i = 1, \dots, n. \quad (2.10)$$

No próximo capítulo será abordado o método de Modelo de Equação Estrutural (MEE) já comparando sua estrutura com a TRI, a fim de, posteriormente, realizar a comparação dos resultados dos dois métodos.

Capítulo 3

Modelos de Equações Estruturais

3.1 Introdução

Este capítulo dedica-se a uma descrição do método da MEE. Aqui será comentado, brevemente, sobre o contexto histórico desse método; a relação da MEE com Análise de Caminho Clássico e Análise Fatorial; alguns conceitos básicos assim como alguns aspectos a serem considerados e a apresentação do modelo visto em Glockner-Rist and Hoijtink (2003), já modificado com o intuito de incluir o modelo da TRI.

3.1.1 Contexto Histórico

Como podemos ver em Pilati and Laros (2007) o MEE é um método de análise de dados que possui origem híbrida em ao menos três diferentes ciências, dentre elas a psicometria. O modelo da psicometria clássica, estabeleceu o princípio de mensuração de características latentes humanas por meio de tarefas comportamentais. Nas ciências comportamentais, há um crescente interesse em compreender e caracterizar mecanismos de desenvolvimento e de processos comportamentais (Van Montfort et al., 2004).

Os psicólogos procuravam mensurar tais capacidades humanas por meio de uma série de tarefas (como por exemplo, itens dos testes) e a identificação da estrutura. Essas medidas eram dada pelas intercorrelações entre esses itens.

O MEE é um método de equações estruturais, que realiza o ajuste dos modelos confirmatórios dos sistemas dinâmicos implícitos a dados de observações repetidas. Como visto em Mueller (1999) os modelos de equações estruturais não são nada mais, e nada menos, que Análise de Caminho Clássico (método multivariado baseado na Análise de Regressão) que envolvem variáveis latentes (medidas não diretamente observadas, mas sim derivadas de indicadores observáveis; como itens de uma escala de avaliação psicológica ou de comportamento organizacional).

Porém a Análise de Caminho Clássico possui certas desvantagens quando integrada

com variáveis latentes, sendo uma das principais desvantagens o método de estimação por mínimos quadrados ordinários (MQO), pois este método possui rigorosos pressupostos como por exemplo, medição perfeita de variáveis observadas e erros não correlacionados em todas as equações.

Na maioria das situações de modelagem da equação estrutural, as técnicas de estimação, como a máxima verossimilhança (ML) ou os mínimos quadrados generalizados (MQG), são preferidas ao invés da MQO, uma vez que permitem a análise de modelos que envolvam variáveis latentes e covariâncias de erros não nulos em equações estruturais.

Dessa forma, o método da MEE utiliza princípios teóricos e lógicos dos modelos psicométricos da análise fatorial e a lógica de teste de equações múltiplas da econometria.

Como a maioria das técnicas de análise de dados, o MEE é útil para o teste de modelos teóricos em diferentes áreas do conhecimento. No Brasil, o MEE é mais utilizado, em pesquisas de Psicologia Organizacional e do Trabalho, Psicologia Social, Marketing, gestão de pessoas, entre outras áreas. Assim, demonstra-se como técnica útil para o desenvolvimento teórico de diversas subáreas da psicologia, bem como para outras disciplinas que tenham a necessidade de mensuração e teste de relações entre fenômenos complexos e que exijam a análise de múltiplas relações simultâneas, entre as quais variáveis latentes.

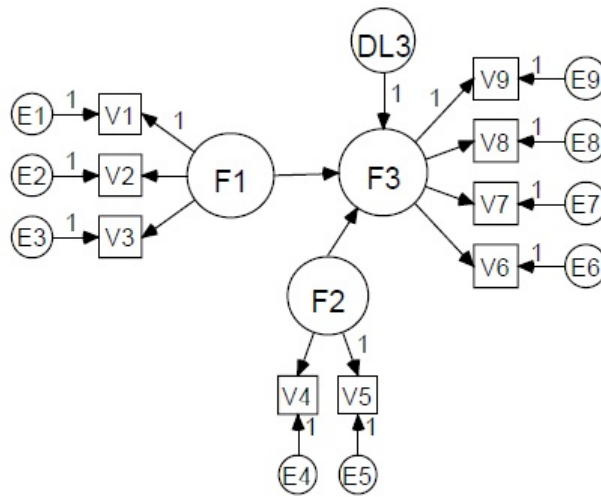
3.1.2 Conceitos Básicos

A MEE tem como objetivo a reprodução da matriz de covariâncias populacionais por meio das covariâncias amostrais associadas às imposições de parâmetros determinadas pelo pesquisador. Se o modelo imposto for plausível, isso significa que covariâncias provenientes de dados amostrais podem reproduzir os dados populacionais associados ao conjunto de parâmetros do modelo estrutural. Modelos plausíveis indicam que a reprodução gerou resíduos reduzidos. Do contrário, se essa reprodução não for plausível, os resíduos são elevados.

Com essa lógica, o modelo da MEE procura replicar um conjunto de dados observados por meio da imposição de parâmetros nas matrizes, que são as relações teóricas definidas pelo pesquisador. Essa característica é a principal diferença entre a MEE das outras técnicas de análise fatorial multivariada, pois a imposição dos parâmetros na matriz de relações entre as variáveis dá a ela um caráter confirmatório, pois exige do pesquisador uma pré-definição do tipo de relações existentes entre as variáveis do modelo em teste, que são operacionalizadas em termos de restrições nas matrizes.

Por esse motivo, a MEE necessita que: a) as medidas utilizadas pelo pesquisador sejam de boa qualidade psicométrica e b) modelos teóricos sólidos e fundamentados em pesquisas anteriores, permitam ao pesquisador estabelecer essas imposições (relações pré-definidas) com propriedade. Devido a essa última característica é que a MEE é entendida, não como uma Análise Fatorial Exploratória (AFE), mas como uma Análise Fatorial Confirmatória (AFC), pois a modelação teórica sobre o que está sob investigação deve ter ocorrido antes da análise dos dados (Pilati and Laros, 2007).

Figura 3.1 *Exemplo de Modelo de Equação Estrutural (Pilati and Laros, 2007).*



A Figura 3.1 apresenta um exemplo de representação pictórica de modelo de equações estruturais. Nessa figura, as variáveis latentes (*Fatores*) são representados pela letra ‘F’, as variáveis observadas pela letra “V” e os erros, associados a cada variável observada, pela letra “E”. O distúrbio, erro associado a uma variável dependente, é representado pela letra “D”.

Como pode ser observado em Pilati and Laros (2007), as relações explicativas entre as variáveis são descritas por setas unidirecionais. Cada uma dessas setas pode ser descrita matematicamente como uma reta de regressão. Essas relações também podem ser entendidas como cargas fatoriais, se relativas à análise fatorial confirmatória de instrumentos de medida. Já as setas curvas e bidirecionais representam correlação ou covariância entre variáveis.

Um outro aspecto relevante na nomenclatura da MEE diz respeito ao tipo de variáveis dos modelos, também comum em outras áreas, tais como *Economia*, indicando que uma variável determina ou é determinada por outras variáveis. Na MEE temos o que é chamado

de variáveis *Endógenas*, que se referem àquelas que recebem setas unidirecionais e são conhecidas também como variáveis dependentes. Temos também as variáveis *Exógenas* que são aquelas que originam setas unidirecionais e são conhecidas como variáveis independentes. Na MEE ocorrem casos nos quais uma mesma variável pode ser endógena e exógena, ou seja, ela pode “receber” e “enviar” setas, comportando-se como variável dependente e variável independente simultaneamente. Isso é possível devido a aspectos teóricos de mediação entre variáveis, tipo de relação que pode ser testado em MEE. Esse aspecto será visto com mais detalhes no Capítulo 4.

Após apresentado o contexto histórico e os conceitos básicos, será mostrado a seguir o modelo de análise fatorial linear modificado para incluir modelos TRI com a estrutura da MEE a fim de combinar os métodos e comparar as análises como pode ser visto em Glockner-Rist and Hoijsink (2003).

3.2 A relação entre Análise Fatorial e TRI

3.2.1 Uma Interpretação TRI do Modelo de Um Fator

O modelo de Um Fator de Spearman é apropriado quando as respostas de $i = 1, \dots, n$ pessoas, para $j = 1, \dots, J$ variáveis contínuas observadas y_{ij} pode ser explicado pelos escores dos respondentes θ_i em uma variável latente.

Tradicionalmente, uma decomposição dos autovalores da matriz de correlação de Pearson entre as variáveis observadas é usada para calcular as cargas fatoriais λ_j e comunalidades (o quanto da variância de determinada variável é medida).

Os software que ajustam o MEE, utilizam o seguinte modelo:

$$y_i \sim N(\mu, \Sigma), \quad (3.1)$$

com $i = 1, \dots, n$, com μ e Σ indicando a média populacional e a matriz de covariância das variáveis observadas, respectivamente. A decomposição do autovalor é usada para modelar Σ :

$$\Sigma = \lambda\lambda' + \Psi, \quad (3.2)$$

em que Ψ denota a matriz diagonal de variância residual. Posteriormente, as cargas fatoriais e as comunalidades podem ser estimadas usando, por exemplo, a máxima verossimilhança, assim como os interceptos μ . Pelas equações (3.1) e (3.2) podemos ver que:

$$y_{ij} = \mu_j + \lambda_j\theta_i + \varepsilon_{ij}, \quad (3.3)$$

com $\theta_i \sim (0, 1)$ e $\varepsilon_{ij} \sim N(0, \varphi_j)$. O parâmetro φ_j representa o j -ésimo elemento diagonal de Ψ (isto é, a variância residual quando as respostas à j -ésima variável observada são explicadas usando os escores θ_i das pessoas). O modelo de Um Fator de Spearman é obtido se as variáveis observadas forem padronizadas (isto é, se $\mu = 0$ e Σ for a matriz de correlação).

Para interpretar o modelo de um fator como modelo do método da TRI, a equação (3.3) deve ser reescrita. Fazendo $\mu_j = -\lambda_j \delta_j$ que é uma reparametrização válida se $\lambda_j \neq 0$ para $j = 1, \dots, J$, então

$$y_{ij} = -\lambda_j \delta_j + \lambda_j \theta_i + \varepsilon_{ij} = \lambda_j (\theta_i - \delta_j) + \varepsilon_{ij}, \quad (3.4)$$

onde θ_i é a habilidade da pessoa i e δ_j a dificuldade do item j . Portanto, y_{ij} é considerado como resultado de um confronto da habilidade da pessoa i com a dificuldade da variável j ponderada com o parâmetro de discriminação de item da variável j .

Na seção a seguir é introduzido um modelo que explica explicitamente a natureza dicotômica das respostas dos itens com uma estrutura subjacente ou multifatorial.

3.2.2 Modelos Fatoriais para Dados Dicotômicos: o Limiar e o Modelo de Ogiva Normal

No MEE e na Análise Fatorial convencional, a matriz de correlação tetracóricas é geralmente usada para uma análise de respostas dicotômicas. As correlações tetracóricas são estimativas das correlações entre J não-observadas, variáveis contínuas normais padrão (y) e J itens dicotômicos observados (x), utilizando a suposição,

$$X_{ij} = \begin{cases} 1, & y_{ij} > \mu_j \\ 0, & \text{caso contrário.} \end{cases} \quad (3.5)$$

O uso de correlações tetracóricas implica que

$$y_i \sim N(0, \mathbf{R}), \quad (3.6)$$

onde \mathbf{R} denota uma matriz de correlação populacional que pode ser decomposta em autovalores

$$\mathbf{R} = \mathbf{\Lambda} \mathbf{\Lambda}' + \mathbf{\Psi}, \quad (3.7)$$

onde $\mathbf{\Lambda} = [\lambda_1, \dots, \lambda_d]$ e d denota o número da dimensão. Neste trabalho a dimensão é unidimensional, ou seja $d = 1$.

Uma implicação da equação (3.7) é que $\lambda_j^2 + \varphi_j = 1$, pois os elementos diagonais de \mathbf{R}

são iguais a 1. Essa restrição não é válida para o modelo descrito na Seção 3.2.1. Como o limite de intercepção é contabilizado na Equação (3.5), a contrapartida multidimensional da Equação (3.3) torna-se

$$y_{ij} = \lambda_j \theta_i + \varepsilon_{ij}, \quad (3.8)$$

com $\varepsilon_{ij} \sim N(0, \varphi_j)$.

Takane and De Leeuw (1987) provaram que a equação (3.8) está relacionada à

$$P(X_{ij} = 1|\theta_i) = \Phi[(\lambda_j \theta_i - \mu_j)/\sqrt{\varphi_j}]. \quad (3.9)$$

Prova de Takane and De Leeuw (1987) da equação (3.9) sobre (3.8):

Seja $\mathbf{X}^i = (\mathbf{X}_{i1}, \dots, \mathbf{X}_{iJ})$ um vetor aleatório de padrões de resposta para J itens de teste dicotômicos, onde cada um deles é definido como

$$X_{ij} = \begin{cases} 1, & \text{se o item } j \text{ for respondido corretamente} \\ 0, & \text{caso contrário} \end{cases}$$

para $j = 1, \dots, J$. θ representa as habilidades dos respondentes com função de densidade denotada por $g(\theta)$. Esse parâmetro não é observado diretamente mas, segue que assume uma distribuição normal multivariada com média 0 e covariância I (matriz identidade), ou seja, $u \sim N(0, I)$. O domínio de θ ($\theta \in \mathbb{R}$) é a região multidimensional definida por $(-\infty, \infty)$.

O Modelo de Ogiva Normal de Dois Parâmetros na TRI especifica a probabilidade marginal de $X = x$ como

$$P(X = x) = \int_{\Theta} P(X = x|\theta)g(\theta)d\theta, \quad (3.10)$$

onde $P(X = x|\theta)$ é a probabilidade condicional das respostas padronizadas observadas x dado θ . $P(X = x|\theta)$ é assumido como

$$P(X = x|\theta) = \prod_{j=1}^J (p_j(\theta))^{x_j} (1 - p_j(\theta))^{(1-x_j)}, \quad (3.11)$$

com

$$p_i(\theta) = \int_{-\infty}^{\alpha'\theta + \gamma} \phi(z)dz = \Phi(\alpha'\theta + \gamma), \quad (3.12)$$

onde ϕ é a função de distribuição normal padrão e Φ a função de ogiva normal (ou seja, a função de distribuição acumulativa da distribuição normal padrão).

Em análise fatorial de variáveis dicotomizadas, por outro lado, a probabilidade marginal padronizada de resposta x é especificada como

$$P(X = x) = \int_R h(y) dy, \quad (3.13)$$

onde R é a região de integração multidimensional, e

$$y = \lambda\theta + \varepsilon. \quad (3.14)$$

O modelo (3.14) é o modelo usual comum de análise fatorial com λ sendo a matriz de cargas fatoriais, θ o vetor de scores (que representa as habilidades) e ε o vetor aleatório de singularidades. E é assumido que $\theta \sim N(0, I)$ e $\varepsilon \sim N(0, \varphi)$ onde φ denota o j -ésimo elemento diagonal de Ψ , e θ e ε são independentes. Isso segue que

$$y \sim N(0, \lambda\lambda' + \Psi), \quad (3.15)$$

(distribuição marginal de y) e

$$y|\theta \sim N(\lambda\theta, \Psi), \quad (3.16)$$

(distribuição condicional de y dado θ). As variáveis aleatórias contínuas, y , são dicotomizadas como na equação (3.5),

$$X_{ij} = \begin{cases} 1, & y_{ij} \geq \mu_j, \\ 0, & y_{ij} < \mu_j, \end{cases}$$

para $j = 1, \dots, J$, onde, μ_j é o parâmetro limiar para variável j . Assim, R é a região de integração multidimensional com $R_j = (r_j, \infty)$ se $X_{ij} = 1$ e $R_j = (-\infty, r_j)$ se $X_{ij} = 0$.

Agora (3.10) incluindo (3.11) e (3.12) é equivalente a (3.13) com θ definido em (3.14). Então, primeiramente, será provado (3.13) \rightarrow (3.10). De 3.13 nós temos

$$\begin{aligned} P(X = x) &= \int_R h(y) dy \\ &= \int_R \left(\int_{\Theta} f(y|\theta)g(\theta)d\theta \right) dy \end{aligned}$$

onde

$$P(X = x) = \int_{\Theta} g(\theta) \left(\int_R f(y|\theta) d(y) \right) d(\theta) \quad (3.17)$$

Observe que (3.17) é "completamente" geral, na medida em que não estão envolvidos pressupostos de distribuição. Mesmo a suposição de independência local, tão característica dos métodos variáveis latentes, não é necessária. Mas por causa de (3.16) temos

$$\int_R f(y|\theta) d(y) = \prod_j \int_{R_i} f_i(y_j|\theta) dy_j$$

e com isso, obtendo

$$\int_R f(y|\theta) d(y) = \prod_j \left(\int_{r_i}^{\infty} f_i(y_j|\theta) dy_j \right)^{x_i} \left(1 - \int_{r_i}^{\infty} f_i(y_j|\theta) dy_j \right)^{1-x_i} \quad (3.18)$$

onde

$$\int_{r_i}^{\infty} f_i(y_j|\theta) dy_j = \Phi \left(\frac{\lambda_j \theta - \mu_j}{\sqrt{\varphi_j}} \right) \quad (3.19)$$

para $j = 1, \dots, J$. A equação (3.18) é equivalente a (3.12) definindo,

$$\alpha_j = \frac{\lambda_j}{\sqrt{\varphi_j}} \quad (3.20)$$

$$\gamma_j = \frac{-\mu_j}{\sqrt{\varphi_j}} \quad (3.21)$$

e com isso comprova-se de que a partir da equação (3.8) é possível obter (3.9).

Este modelo tem parâmetros com interpretações semelhantes ao modelo apresentado na Equação (3.3). A principal diferença é que todas as variâncias residuais são fixadas em $1 - \lambda_j^2$ para $j = 1, \dots, J$, para identificação do modelo. Mais restrições de identificação são: (i) $\theta_i \sim N(0, \mathbf{I})$, ou (ii) $\theta_i \sim N(0, \mathbf{\Omega})$, e $\lambda_{1d} = 1$, para $d = 1, \dots, D$, onde $\mathbf{\Omega}$ denota a matriz de covariância dos escores dos fatores.

Usando

$$\mu_j = \delta_j \lambda_j \quad (3.22)$$

tendo μ como o parâmetro de dificuldade, e

$$\beta = \lambda_j / \sqrt{(1 - \lambda_j^2)} \quad (3.23)$$

como o parâmetro de discriminação.

Com isso a equação (3.9) pode ser reescrita como:

$$P(X_{ij} = 1|\theta_i) = \Phi[(\beta_j(\theta_i - \delta_j))] \quad (3.24)$$

Este é o modelo de ogiva normal multidimensional de dois parâmetros, com o qual Takane and De Leeuw (1987) estabeleceram a equivalência entre modelos fatoriais e TRI. Nesse modelo cada pessoa tem uma habilidade θ_i , e cada item tem uma discriminação β_j ou carga fatorial λ_j . Para identificar o modelo para cada item, apenas um item de dificuldade δ_j ou limiar μ_j está incluído.

3.2.3 Modelos de Resposta ao Item Logístico e Ogiva Normal

Paralelamente aos modelos de ogiva normais, os modelos logísticos da TRI foram desenvolvidos. Eles calculam a probabilidade condicional $P(x_{ij} = 1|\theta_i)$ integrando sobre a logística em vez da distribuição normal. Para o modelo de um fator, foi mostrado que

$$\Phi[\beta_j(\theta_i - \delta_j)] \approx \exp(1.7\beta_j(\theta_i - \delta_j)) / [1 + \exp(1.7\beta_j(\theta_i - \delta_j))]. \quad (3.25)$$

Então teremos a equação Modelo Logístico de Dois Parâmetros. O fator de escala 1.7 apenas expressa o fato de que os parâmetros de discriminação obtidos a partir de modelos logísticos da TRI são aproximadamente 1.7 vezes maiores que os parâmetros correspondentes dos modelos de ogiva normais.

No próximo capítulo será apresentado um pouco dos aspectos computacionais em relação aos métodos da TRI e MEE e suas estimativas, apresentados no Capítulo 2 e no Capítulo 3, respectivamente.

Capítulo 4

Aspectos Computacionais: o pacote MIRT e o pacote LAVAAN do programa R

Neste capítulo serão abordados os aspectos computacionais utilizados para o desenvolvimento deste trabalho, relativos às estimações dos parâmetros desenvolvidas pelo método da TRI, mostradas no Capítulo 2, assim como as estimações dos parâmetros desenvolvidas pelo método da MEE, mostradas no Capítulo 3. As sintaxes foram elaboradas no *software R* versão 3.4.1 e executadas em um computador com processador Intel(R) Core(TM)i3-4005U CPU @ 1.70GHz e memória (RAM): 4,00GB.

Nas seções a seguir serão mostrados detalhes sobre cada pacote adotado, ambos do programa R, utilizados para a obtenção das estimativas dos parâmetros. Será apresentada uma breve descrição dos métodos antes de abordar os pacotes. O pacote utilizado para estimar as habilidades e os parâmetros de discriminação e dificuldade pelo método da TRI foi o MIRT (Chalmers et al. (2012)). O pacote utilizado para o método da MEE foi o pacote LAVAAN (Rosseel (2012)).

4.1 Pacote MIRT

4.1.1 Introdução

O pacote MIRT (*mirt: A Multidimensional Item Response Theory Package for the R Environment*) foi apresentado em Chalmers et al. (2012). Como já informado, a TRI é amplamente utilizada na pesquisa educacional e psicológica para modelar como os participantes respondem a itens de teste isoladamente e em grupos. É uma estrutura geral para especificar a relação funcional entre o nível de traço latente de um respondente (habilidade do indivíduo) e um nível de estímulo de um item. A metodologia da TRI tenta modelar padrões de resposta individuais, especificando como os traços latentes subjacentes inte-

ragem com as características do item (como a dificuldade a discriminação de um item) para formar uma probabilidade associada ao padrão de resposta. Como tal, um objetivo principal da TRI é separar os parâmetros dos itens e as características da amostra de uma população para que ambos possam ser entendidos e estudados separadamente. Esta separação de parâmetros geralmente requer técnicas avançadas de análise numérica para obtenção das estimativas e pode exigir um custo computacional muito alto à medida que a complexidade do modelo aumenta.

Os modelos TRI mais simples e mais populares são aqueles que especificam uma característica latente única (ou seja, unidimensional). Os modelos TRI unidimensionais têm sido predominantes em todas as ciências sociais e pesquisas educacionais principalmente por causa das tradições históricas, mas também porque os procedimentos multidimensionais de estimação de parâmetros não foram totalmente desenvolvidos ou estudados (Baker and Kim 2004; Reckase 2009). Embora os modelos unidimensionais sejam geralmente mais simples e possam ter várias propriedades de medição interessantes e importantes, muitas construções psicológicas são inevitavelmente de natureza multidimensional. Um impedimento importante ao decidir utilizar esses modelos foi que a estimativa dos parâmetros do item no espaço dimensional maior (que ocorre devido ao aumento do número de fatores) é computacionalmente difícil para técnicas de integração numéricas padrão. No entanto, com os avanços recentes na teoria, a pesquisa da TRI multidimensional vem sendo desenvolvida e melhorada como uma metodologia de análise estatística Chalmers et al. (2012).

Vários pacotes R (R Development Core Team 2012) podem ser usados para modelagem em TRI, como: o pacote ltm (Latent Trait Models under IRT) ou o pacote irtoys (A Collection of Functions Related to Item Response Theory). Embora bastante úteis, a maioria dos pacotes têm limitações que só permitem adotar os modelos TRI unidimensionais efetivamente. O pacote MIRT (Multidimensional Item Response Theory) foi criado para estimar parâmetros de teoria de resposta ao item multidimensionais para modelos exploratórios e confirmatórios usando métodos de máxima verossimilhança. Embora neste trabalho só estejamos trabalhando com o caso uni-dimensional, a adoção do MIRT será útil para futuros trabalhos com mais de uma dimensão.

4.1.2 O Modelo Logístico de 2 parâmetros da TRI

A classe de Modelos Logísticos se estende até o modelo de 5 parâmetros. Um dos modelos de resposta mais usados para itens dicotômicos/dicotomizados é o modelo de três

parâmetros, apresentado no Capítulo 2, e que pode ser reduzido a um modelo 1PL ou 2PL. Sendo $i = 1, \dots, n$ representando o número de respondentes e $j = 1, \dots, J$ representando o total de itens (variáveis observadas) a probabilidade de responder corretamente a um item dicotômico é:

$$P(x_{ij} = 1 | \theta_i, a_j, d_j) = \frac{1}{1 + e^{-D(a_j\theta_i - d_j)}}, \quad (4.1)$$

onde, θ_i é a habilidade do respondente i , D é um ajuste (geralmente 1,7) usado para tornar a função logística mais próxima da função ogiva normal, a_j é o parâmetro de discriminação do item j e d_j é o intercepto do item sendo $d_j = -b_j * a_j$. As demais estimativas dos parâmetros são equivalentes às apresentadas no Capítulo 2.

Este modelo é um pouco diferente do apresentado no Capítulo 2, pois no *software R* a estimação dos parâmetros é calculada usando o intercepto d_j . Sendo assim é necessário, após obtidas as estimativas, realizar a divisão do intercepto do item d_j pelo parâmetro de discriminação do item $-a_j$ para obter o parâmetro de dificuldade do item b_j , ou seja, $b_j = d_j / (-a_j)$.

O comando da fórmula *mirt* é dado por:

```
mod1 = mirt(data, nfact = 1, itemtype = '2PL')
```

onde, *data* representa o banco de dados dicotomizados, *nfact* é representados pelo número de fatores (dimensões) e *itemtype* representa o tipo de modelo a ser seguido. O comando para a extração dos parâmetros é dado por

```
coef(mod1)
```

onde, *mod1* é o nome da variável definida da função *mirt*. A primeira coluna de *coef(mod1)* representa a variável a_j e a segunda coluna a variável d_j . Para termos a variável b_j basta dividir a segunda coluna pela primeira coluna e multiplicar por menos um. E para a extração das habilidades, basta utilizar o comando:

```
fscores(mod1)
```

4.2 Pacote LAVAAN

4.2.1 Introdução

O lavaan (Latent Variable Analysis) foi proposto por Rosseel (2012) como um pacote dedicado a modelagem de equações estruturais. Seu nome revela o objetivo de longo

prazo: fornece uma coleção de ferramentas que podem ser usadas para explorar, estimar e compreender uma ampla família de modelos de variáveis latentes, incluindo análise fatorial, equação estrutural, longitudinal, multinível, classe latente, teoria da resposta do item e modelos de dados missings.

No entanto, o desenvolvimento do lavaan só começou, e ainda há muito a ser feito para alcançar esse objetivo ambicioso. Até o momento, o desenvolvimento de lavaan se concentrou na modelagem de equações estruturais (MEE) com variáveis observáveis contínuas. Os modelos de equações estruturais abrangem uma ampla gama de técnicas estatísticas multivariadas.

Dentro do ambiente R, existem duas abordagens para estimar modelos de equações estruturais. A primeira abordagem é conectar R com programas MEE externos comerciais. Isso geralmente é útil em estudos de simulação em que o modelo com software MEE é uma parte da simulação. A segunda abordagem é usar um pacote R dedicado para modelagem de equações estruturais.

4.2.2 Por que precisamos do Lavaan?

Muitos pacotes de software MEE estão disponíveis, tanto gratuitos como comerciais, incluindo alguns pacotes que são executados no ambiente R. Por que, então, existe a necessidade de um novo pacote MEE? As respostas a esta questão são as seguintes:

1. O lavaan pretende atrair um grande grupo de pesquisadores aplicados que precisam do software MEE para responder suas questões de pesquisa. Pesquisadores aplicados geralmente valorizam software que é intuitivo e rico em características de modelagem, e a lavaan tenta cumprir esses dois objetivos.
2. O lavaan visa apelar para aqueles que ensinam MEE ou oficinas MEE; idealmente, os professores devem ter acesso a um programa de MEE fácil de usar, porém completo e fácil de instalar em uma sala de aula de informática.
3. O lavaan visa atrair os estatísticos que trabalham no campo da MEE.

O primeiro objetivo é indiscutivelmente o mais difícil de conseguir. Se quisermos convencer os usuários de programas de MEE comerciais a usar lavaan, deve haver razões convincentes para trocar. Que lavaan é livre, muitas vezes é irrelevante. O que mais interessa a muitos pesquisadores aplicados é que: (1) o software é fácil e intuitivo de usar, (2)

o software tem todos os recursos que eles querem, e (3) os resultados de lavaan são muito próximos, se não idênticos, a aqueles relatados pelo seu programa comercial atual.

O segundo objetivo visa aqueles que ensinam técnicas MEE em aulas ou workshops. Para os professores, o fato de a lavaan ser livre é importante. Se o software for gratuito, não há mais necessidade de instalar "versões estudantis" limitadas dos programas comerciais para acompanhar o curso MEE. Claro, os professores também apreciarão uma experiência de usuário fácil e intuitiva, para que possam passar mais tempo discutindo e interpretando os resultados substantivos de uma análise MEE, em vez de gastar tempo explicando a sintaxe do modelo estranho de um programa específico. Finalmente, a opção torna possível uma transição suave de lavaan para um dos principais programas comerciais e de volta. Portanto, os alunos que receberam instruções iniciais em SEM com lavaan devem ter pouca dificuldade usando outros programas (comerciais) MEE no futuro.

O terceiro objetivo visa os estatísticos profissionais que trabalham no campo da modelagem da equação estrutural. Por muito tempo, este campo foi dominado por softwares comerciais de código fechado. Já lavaan disponibiliza exemplos de aplicações da MEE para interessados em trabalhar na área estatística.

4.2.3 Especificando Modelos em Lavaan

No pacote lavaan, os modelos são especificados por meio de uma sintaxe baseada em texto que descreve o modelo, denominada "sintaxe do modelo lavaan". Considere um modelo de regressão simples com uma variável dependente contínua y e quatro variáveis independentes x_1 , x_2 , x_3 e x_4 . O modelo de regressão usual pode ser escrito da seguinte forma:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + e_i, \quad (4.2)$$

onde β_0 é chamado de intercepto, β_1 a β_4 são os coeficientes de regressão para cada uma das quatro variáveis, e e_i é o erro residual para a observação i . Uma das características atraentes do ambiente R é a forma compacta em que podemos expressar uma fórmula de regressão como a anterior:

$$y \sim x1 + x2 + x3 + x4.$$

Uma maneira de olhar para os modelos MEE é que eles são simplesmente uma

extensão da regressão linear. Uma primeira extensão é que você pode ter várias equações de regressão ao mesmo tempo. Uma segunda extensão é que uma variável que é uma variável independente (exógena) em uma equação pode ser uma variável dependente (endógena) em outra equação. Por exemplo, poderíamos ter um conjunto de três equações de regressão:

$$y1 \sim x1 + x2 + x3 + x4$$

$$y2 \sim x5 + x6 + x7 + x8$$

$$y3 \sim y1 + y2$$

Esta é a abordagem adotada pelo lavaan. As equações de regressão múltipla são simplesmente um conjunto de fórmulas de regressão, usando a sintaxe típica de uma fórmula R.

Uma terceira extensão dos modelos MEE é que eles incluem variáveis latentes contínuas. No lavaan, qualquer fórmula de regressão pode conter variáveis latentes, tanto como uma variável dependente como independente. Por exemplo, na sintaxe mostrada abaixo, as variáveis que começam com "f" são variáveis latentes:

$$y1 \sim f1 + f2 + x1 + x2$$

$$f1 \sim x1 + x2$$

Esta parte da sintaxe do modelo corresponderia à "parte estrutural" de um modelo MEE. Para descrever a "parte de medição" do modelo, precisamos especificar os indicadores (observados ou latentes) para cada uma das variáveis latentes. Em lavaan, isso é feito com o operador especial " \sim ". O lado esquerdo desta fórmula contém o nome da variável latente. O lado direito contém os indicadores dessa variável latente, separados pelo operador "+". Por exemplo:

$$f1 \sim item1 + item2 + item3$$

$$f2 \sim item4 + item5 + item6 + item7$$

$$f3 \sim f1 + f2$$

Neste exemplo, as variáveis item1 até item7 são variáveis observadas. Portanto, as variáveis latentes f1 e f2 são fatores de primeira ordem. A variável latente f3 é um fator de segunda ordem, já que todos os seus indicadores são as próprias variáveis latentes.

Neste trabalho, teremos apenas um fator de primeira ordem (f1), pois queremos estimar apenas uma habilidade individual, que é a proficiência de cada respondente. Sendo assim, o modelo será a soma de todos os itens:

$$\boxed{\text{Modelo} = \sim \text{item1} + \text{item2} + \dots + \text{itemJ}.} \quad (4.3)$$

Após especificado o modelo e o banco de dados a ser trabalhado, será utilizado a função *cfa* que é uma função dedicada para a análise fatorial condicional (CFA):

$$\boxed{R > \text{fit} = \text{cfa}(\text{Modelo}, \text{data} = \text{'banco de dados'}, \text{name} = \text{'variáveis observáveis'})}$$

O primeiro argumento (*Modelo*) é o objeto que contém a sintaxe do modelo lavaan. O segundo argumento (*data*) é o conjunto de dados que contém as variáveis observadas. E o terceiro argumento (*name*) é o conjunto com o nome das variáveis observadas.

Com o modelo especificado, sempre é possível inspecionar a tabela de parâmetros usando o comando

$$\boxed{\text{parTable}(\text{fit})}$$

A partir desse comando é possível selecionar parâmetros de interesse para o estudo em questão. No caso desse trabalho serão selecionadas as cargas fatoriais (λ_j), os limiares (μ_j) e as variâncias residuais (φ_j).

Figura 4.1 Saída das estimativas de φ , utilizando o pacote Lavaan do software R.

variances:	Estimate	Std.Err	z-value	P(> z)
.Item.1	0.228			
.Item.2	0.541			
.Item.3	0.230			
.Item.4	0.378			
.Item.5	0.270			
.Item.6	0.770			
.Item.7	0.300			
.Item.8	0.228			
.Item.9	0.284			
.Item.10	0.371			
.Item.11	0.369			
.Item.12	0.727			
.Item.13	0.610			
.Item.14	0.308			
.Item.15	0.543			
.Item.16	0.413			
.Item.17	0.398			
.Item.18	0.260			
.Item.19	0.498			
.Item.20	0.309			
.Item.21	0.396			
.Item.22	0.244			
.Item.23	0.309			
.Item.24	0.626			
.Item.25	0.277			
.Item.26	0.421			
.Item.27	0.611			
.Item.28	0.373			
.Item.29	0.422			
.Item.30	0.319			
model	0.772	0.003	282.995	0.000

Figura 4.2 Saída das estimativas de μ , utilizando o pacote *Lavaan* do software *R*.

Thresholds:				
	Estimate	Std.Err	z-value	P(> z)
Item.1 t1	-0.314	0.004	-77.728	0.000
Item.2 t1	0.651	0.004	152.047	0.000
Item.3 t1	0.165	0.004	41.408	0.000
Item.4 t1	0.121	0.004	30.439	0.000
Item.5 t1	-0.670	0.004	-155.582	0.000
Item.6 t1	0.685	0.004	158.559	0.000
Item.7 t1	-0.830	0.005	-184.409	0.000
Item.8 t1	0.184	0.004	46.113	0.000
Item.9 t1	-0.743	0.004	-169.283	0.000
Item.10 t1	-0.172	0.004	-43.139	0.000
Item.11 t1	-0.215	0.004	-53.839	0.000
Item.12 t1	0.888	0.005	193.596	0.000
Item.13 t1	-0.988	0.005	-207.823	0.000
Item.14 t1	-0.192	0.004	-48.165	0.000
Item.15 t1	0.835	0.005	185.253	0.000
Item.16 t1	-0.524	0.004	-125.732	0.000
Item.17 t1	-1.034	0.005	-213.656	0.000
Item.18 t1	0.773	0.004	174.741	0.000
Item.19 t1	1.339	0.006	240.467	0.000
Item.20 t1	0.040	0.004	10.195	0.000
Item.21 t1	0.050	0.004	12.724	0.000
Item.22 t1	0.046	0.004	11.643	0.000
Item.23 t1	-1.362	0.006	-241.625	0.000
Item.24 t1	-0.103	0.004	-26.009	0.000
Item.25 t1	1.360	0.006	241.543	0.000
Item.26 t1	0.625	0.004	146.844	0.000
Item.27 t1	0.799	0.004	179.224	0.000
Item.28 t1	-0.087	0.004	-21.836	0.000
Item.29 t1	0.551	0.004	131.436	0.000
Item.30 t1	1.020	0.005	211.984	0.000

As Figuras (4.1) e (4.2) mostram, respectivamente, os valores de (φ) e (μ) , obtidos através do *software R*, utilizando o pacote *Lavaan*. O parâmetro (λ_j) foi obtido utilizando a restrição $\lambda_j^2 + \varphi_j = 1$, vista no Capítulo 3. Sendo assim $\lambda_j = \sqrt{(1 - \varphi_j)}$.

Com isso, é possível adquirir o parâmetro de dificuldade (δ_j) e de discriminação (β_j) proposto pela MEE, mostrado no Capítulo 3.

4.3 Resultados Obtidos Com Uma Amostra

Aqui será apresentado alguns resultados obtidos pelo *software R*, com o pacote *mirt* para estimar os parâmetros do método da TRI, e o pacote *Lavaan* para estimar os parâmetros do método da MEE. Com o objetivo maior de ilustrar o uso dos pacotes e as quantidades adotadas no estudo, será gerada uma base de dados de exemplificação. No próximo capítulo será feito um estudo mais robusto com a geração de um número razoável de réplicas e exploração de outras quantidades de interesse.

Inicialmente foram gerados valores (parâmetros verdadeiros) para os parâmetros de discriminação, dificuldade de $J = 30$ itens e as habilidades de $n = 100.000$ respondentes

segundo as distribuições:

$$a \sim \text{Uniforme}(0, 5; 2, 0)$$

$$b \sim N(0; 1)$$

$$\theta \sim N(0; 1).$$

Em seguida foi gerado um banco de dados com observações dicotômicas (zeros e uns, sendo 0 um item respondido incorretamente e 1 um item respondido corretamente). Vamos chamar este banco de dados de *Amostra 1*.

O próximo passo foi realizar as estimações pela TRI e MEE, com o objetivo de verificar se os métodos eles foram capazes de estimar as características da Amostra 1 com boa precisão, isto é, próximos aos parâmetros verdadeiros adotados para a geração dos dados. Esta etapa se denomina *recuperação de parâmetros* (*parameter recovery*).

As comparações das estimações realizadas pelos dois métodos com os parâmetros verdadeiros serão dadas nos resultados a seguir.

4.4 Resultados do método da TRI - Amostra 1

A Figura 4.3 mostra o Cruzamento do Parâmetro de Discriminação estimado pelo método da TRI com o Parâmetro de Discriminação verdadeiro. A Figura 4.4 mostra a diferença do parâmetro de discriminação verdadeiro com o estimado pela TRI. Observa-se que a estimacão desse Parâmetro por esse método apresentou uma boa precisão, e uma diferença mínima, em relação ao Parâmetro verdadeiro.

Uma medida resumo bastante adotada para quantificar a precisão no processo de estimacão é o Erro Quadrático Médio (EQM), cuja expressão, em relação a cada um dos parâmetros, é dada por:

$$EQM(a) = \frac{1}{J} \sum_{j=1}^J (\hat{a}_j - a_j)^2 \quad , \quad EQM(b) = \frac{1}{J} \sum_{j=1}^J (\hat{b}_j - b_j)^2$$

$$EQM(\theta) = \frac{1}{n} \sum_{i=1}^n (\hat{\theta}_i - \theta_i)^2.$$

Para os resultados da Amostra 1 relativamente ao parâmetros de discriminação, temos que $EQM(a) = 0.0001$, indicando que cada estimativa do parâmetro a varia em média apenas 0.0102 do verdadeiro valor.

Figura 4.3 *Cruzamento dos Parâmetros de Discriminação estimados pelo Mirt com os verdadeiros.*

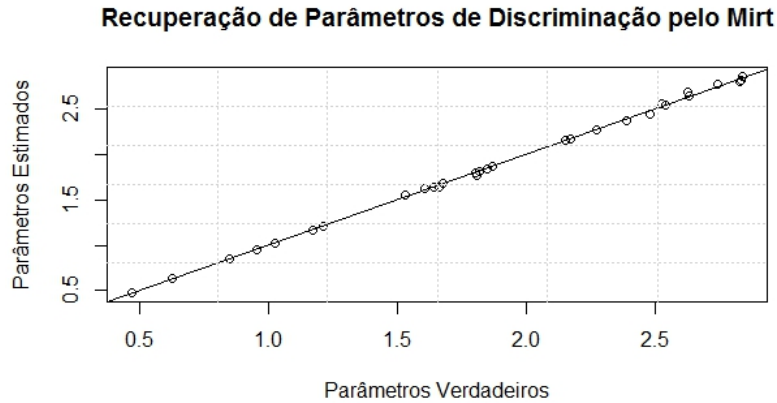
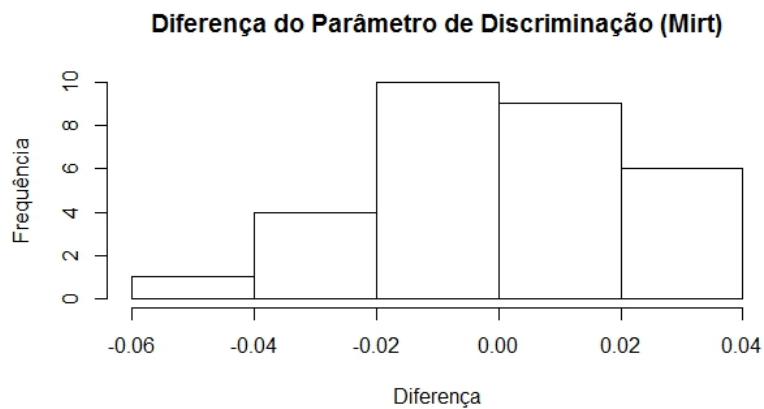


Figura 4.4 *Diferença dos Parâmetros de Discriminação Verdadeiros com os Estimados pelo Mirt.*



A Figura 4.5 mostra o Cruzamento do Parâmetro de Dificuldade estimado pelo método da TRI com o Parâmetro de Dificuldade verdadeiro. A Figura 4.6 mostra a diferença do parâmetro de dificuldade verdadeiro com o estimado pela TRI. Observa-se que a estimação desse Parâmetro por esse método, assim como no caso do parâmetro de discriminação, apresentou uma boa precisão em relação ao Parâmetro verdadeiro.

Para os resultados da Amostra 1 relativamente ao parâmetros de dificuldade, temos que $EQM(b) = 8.9805e^{-05}$, indicando que cada estimativa do parâmetro b varia em média apenas 0.0095 do verdadeiro valor.

Figura 4.5 *Cruzamento dos Parâmetros de Dificuldade estimados pelo Mirt com os verdadeiros.*

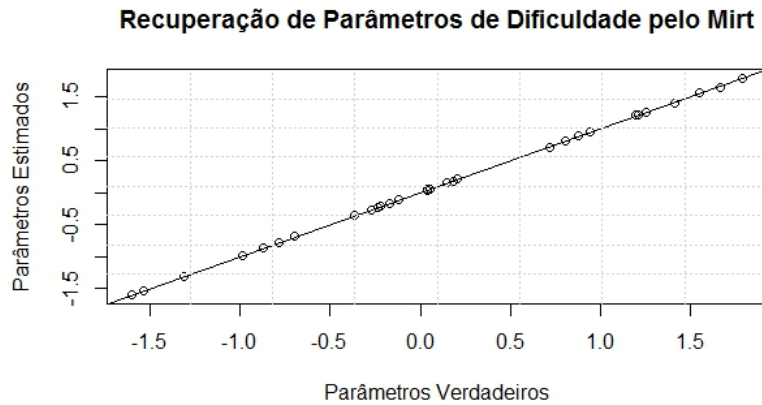
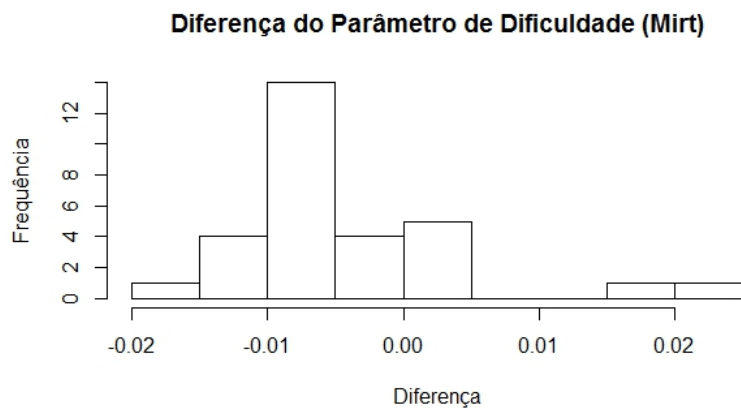
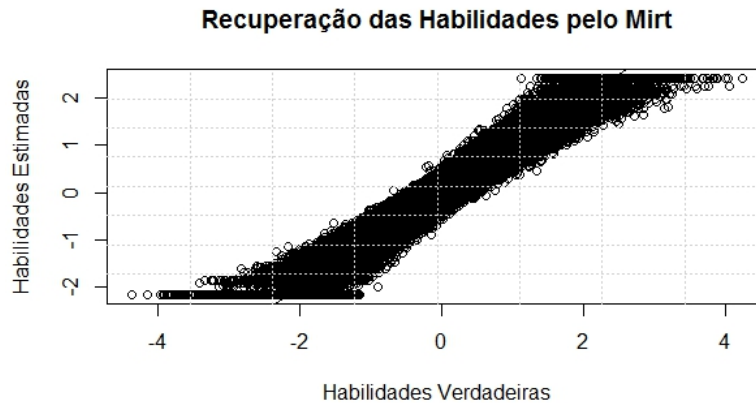
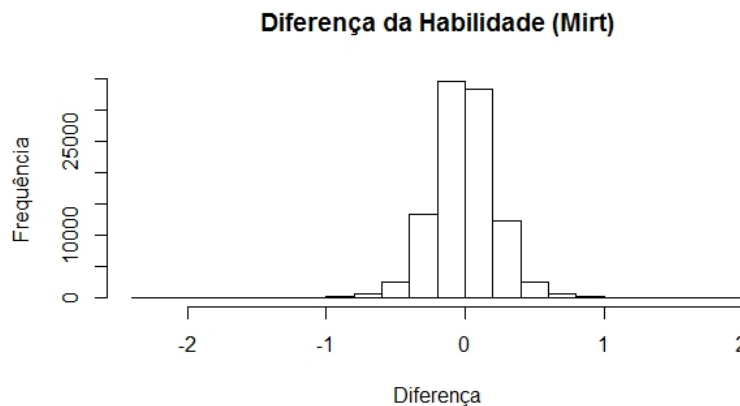


Figura 4.6 *Diferença dos Parâmetros de Dificuldade Verdadeiros com os Estimados pelo Mirt.*



E por fim, a Figura 4.7 mostra o Cruzamento das habilidades estimadas pelo método da TRI com as habilidades verdadeiras, assim como, a Figura 4.8 que mostra a diferença entre as habilidades verdadeiras e estimadas por esse método. Observa-se que a estimação das habilidades por esse método apresentou uma boa precisão em relação as habilidades verdadeiras.

Para os resultados da Amostra 1 relativamente as habilidades, temos que $EQM(\theta) = 0.0638$, indicando que cada estimativa do parâmetro θ varia em média apenas 0.2525 do verdadeiro valor.

Figura 4.7 *Cruzamento das Habilidades estimadas pelo Mirt com as verdadeiras.*Figura 4.8 *Diferença das Habilidades Verdadeiras com as Estimadas pelo Mirt.*

4.5 Resultados do método da MEE - Amostra 1

A Figura 4.9 mostra o Cruzamento do Parâmetro de Discriminação estimado pelo método da MEE com o Parâmetro de Discriminação verdadeiro. Diferente da estimação feita pelo método da TRI, o método da MEE não estimou todos os parâmetros de Discriminação de maneira precisa. Alguns parâmetros apresentaram diferença significativa, em relação aos parâmetros verdadeiros. Isso pode ser comprovado pela figura (4.10), onde é possível verificar que a diferença dos parâmetros estimados com os parâmetros verdadeiros, foi muito maior do que a diferença visto na Figura 4.4.

Para os resultados da Amostra 1 relativamente ao parâmetros de discriminação, temos que $EQM(a) = 0.0173$, indicando que cada estimativa do parâmetro a varia em média apenas 0.1314 do verdadeiro valor, uma variação maior do que a ocorrida no mirt.

Figura 4.9 *Cruzamento dos Parâmetros de Discriminação estimados pelo Lavaan com os verdadeiros.*

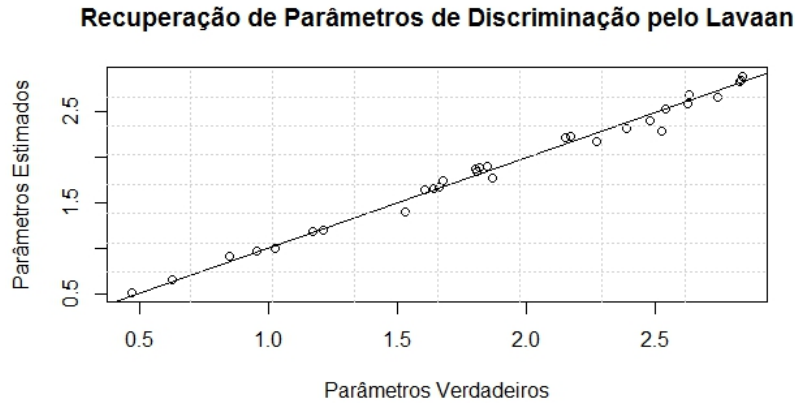
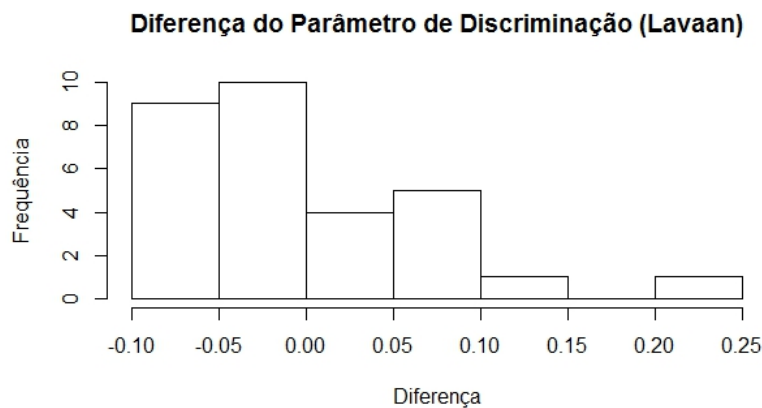


Figura 4.10 *Diferença dos Parâmetros de Discriminação Verdadeiros com os Estimados pelo Lavaan.*



Para os resultados da Amostra 1 relativamente ao parâmetro de dificuldade, temos que $EQM(b) = 0.0008$, indicando que cada estimativa do parâmetro b varia em média apenas 0.0276 do verdadeiro valor. Assim como no caso do parâmetro de discriminação, a variação foi maior maior do que a ocorrida no mirt.

A figura (4.11) mostra o Cruzamento do Parâmetro de Dificuldade estimado pelo método da MEE com o Parâmetro de Dificuldade verdadeiro. A estimação realizada pelo método da MEE apresentou algumas diferenças em relação ao parâmetro verdadeiro, e verificando a figura (4.12), nota-se uma diferença pequena dos parâmetros verdadeiros com os estimados pela MEE. Como essas diferenças não foram tão grandes então afirma-se que os parâmetros foram bem estimados, porém os parâmetros estimados pela TRI apresentou resultados mais precisos.

Figura 4.11 *Cruzamento dos Parâmetros de Dificuldade estimados pelo Lavaan com os verdadeiros.*

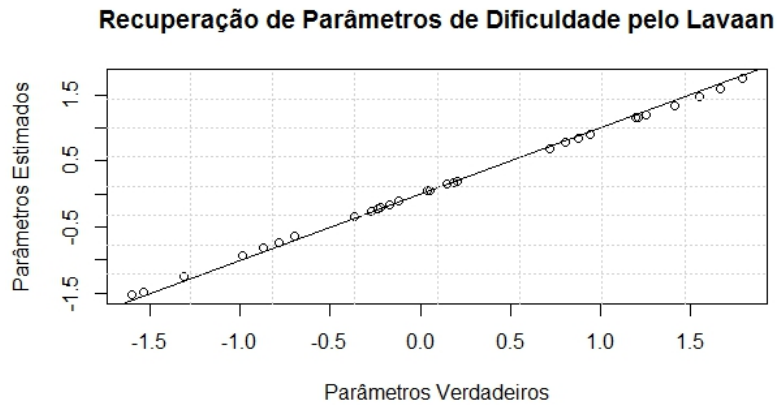
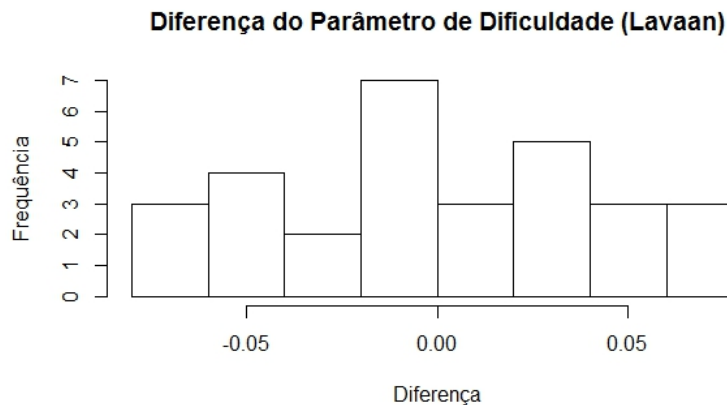
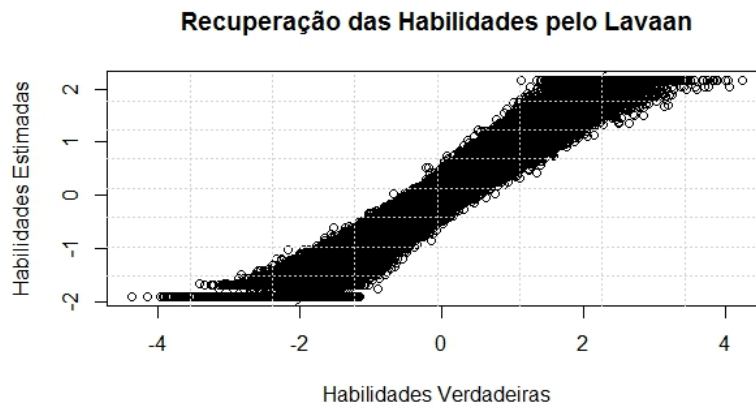
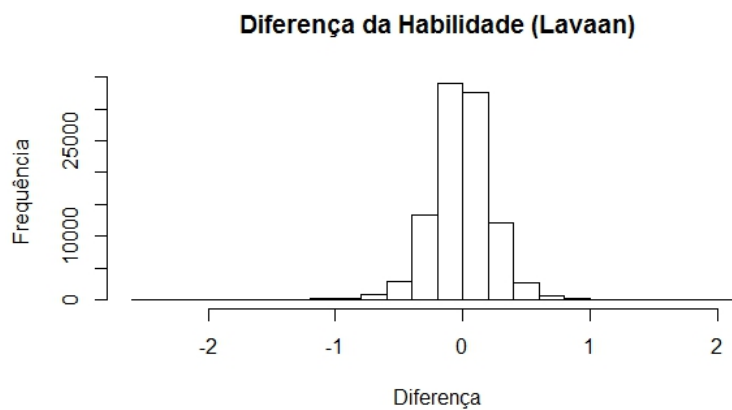


Figura 4.12 *Diferença dos Parâmetros de Dificuldade Verdadeiros com os Estimados pelo Lavaan.*



E por fim, a figura (4.13) mostra o Cruzamento das habilidades estimadas pelo método da MEE com as habilidades verdadeiras e a figura (4.14) mostra a diferença entre as habilidades verdadeiras e estimadas por esse método. Observa-se que a diferença entre as habilidades apresentou a maioria dos valores bem próximo de zero. Sendo assim pode-se afirmar que a estimação das habilidades pelo método da MEE foi precisa em relação as habilidades verdadeiras.

Para os resultados da Amostra 1 relativamente a parâmetro de dificuldade, temos que $EQM(\theta) = 0.0660$, indicando que cada estimativa do parâmetro θ varia em média apenas 0.2570 do verdadeiro valor, uma variação equivalente a variação do mirt no caso deste parâmetro.

Figura 4.13 *Cruzamento das Habilidades estimadas pelo Lavaan com as verdadeiras.*Figura 4.14 *Diferença das Habilidades Verdadeiras com as Estimadas pelo Lavaan.*

No próximo capítulo será apresentado os resultados, de uma maneira mais ampla sobre a comparação desses dois métodos, e aplicação a dados reais dos testes elaborados para a resolução deste trabalho, assim como a análise de cada resultado.

Capítulo 5

Estudo de Simulação

5.1 Introdução

Neste capítulo são apresentados os resultados de um estudo mais amplo sobre a comparação dos métodos e resultados obtidos via MEE e TRI. Foram utilizados diferentes tamanhos amostrais, $n = 100, 200, 500, 1000$ e 10000 , bem como diferentes tamanhos de testes: $J=10, 20, 50$ e 100 . Os parâmetros dos itens foram gerados aleatoriamente, de acordo com as distribuições:

$$a_j \sim U(0.5; 2.0), \quad b_j \sim N(0, 1) \quad \theta_i \sim N(0, 1).$$

Para cada par (J, n) de tamanhos amostrais e número de itens, foram realizadas $R = 100$ réplicas. Em cada réplica ($r = 1, \dots, R$) e para classe de parâmetro (a, b, θ) foi construído o *Erro Quadrático Médio* (EQM) entre as estimativas e os verdadeiros valores dos parâmetros da classe. Também foi observado o tempo de execução (T), em segundos, de forma que possamos verificar a viabilidade de aplicação a grandes bases de dados.

$$EQM(a)_r = \frac{1}{J} \sum_{j=1}^J (\hat{a}_j - a_j)^2, \quad EQM(b)_r = \frac{1}{J} \sum_{j=1}^J (\hat{b}_j - b_j)^2$$
$$EQM(\theta)_r = \frac{1}{n} \sum_{i=1}^n (\hat{\theta}_i - \theta_i)^2, \quad r = 1, \dots, R.$$

Finalmente, foram obtidas as quantidades a seguir para cada J e n por:

$$EQM(a) = \frac{1}{R} \sum_{r=1}^R EQM(a)_r, \quad EQM(b) = \frac{1}{R} \sum_{r=1}^R EQM(b)_r,$$
$$EQM(\theta) = \frac{1}{R} \sum_{r=1}^R EQM(\theta)_r$$

Tabela 5.1 *Erro Quadrático Médio*

J	n	Método	$EQM(a)$	$EQM(b)$	$EQM(\theta)$	$T(\text{segundos})$
10	100	TRI	2.2940	1.1888	0.2765	0.4924
		MEE	0.5166	0.1523	0.2765	1.0897
	200	TRI	0.7467	0.1798	0.2627	0.4372
		MEE	0.9584	0.1033	0.2623	1.5180
	500	TRI	0.1066	0.0288	0.2560	0.2924
		MEE	0.0931	0.0902	0.2587	2.9184
	1000	TRI	0.0375	0.0134	0.2502	0.3148
		MEE	0.0394	0.0825	0.2525	5.4674
	10000	TRI	0.0042	0.0011	0.2428	0.6049
		MEE	0.0165	0.0820	0.2453	34.7672
20	100	TRI	1.0197	0.2710	0.1628	0.6194
		MEE	0.3672	0.1371	0.1629	2.1835
	200	TRI	0.1878	0.0812	0.1507	0.5267
		MEE	0.1561	0.1093	0.1522	2.7976
	500	TRI	0.0639	0.0258	0.1449	0.5123
		MEE	0.0616	0.0837	0.1465	4.6955
	1000	TRI	0.0293	0.0114	0.1416	0.6871
		MEE	0.0341	0.0811	0.1432	7.8636
	10000	TRI	0.0025	0.0011	0.1401	2.0748
		MEE	0.0128	0.0728	0.1417	49.2978
50	100	TRI	0.6976	1.7463	0.0797	1.6949
		MEE	0.3374	0.1362	0.0833	14.8851
	200	TRI	0.1560	0.0816	0.0703	1.3479
		MEE	0.1293	0.1092	0.0714	16.1112
	500	TRI	0.0512	0.0233	0.0651	1.4715
		MEE	0.0524	0.0883	0.0665	20.3339
	1000	TRI	0.0233	0.0104	0.0622	2.2170
		MEE	0.0283	0.0795	0.0633	27.0357
	10000	TRI	0.0022	0.0010	0.0613	8.1453
		MEE	0.0124	0.0720	0.0622	117.2858
100	100	TRI	0.5422	0.3711	0.0453	4.0062
		MEE	0.2833	0.1382	0.0474	254.4917
	200	TRI	0.1254	0.0661	0.0386	3.5228
		MEE	0.1221	0.1057	0.0398	262.1207
	500	TRI	0.0432	0.0225	0.0355	4.4585
		MEE	0.0474	0.0851	0.0363	269.6715
	1000	TRI	0.0215	0.0105	0.0329	6.2984
		MEE	0.0291	0.0791	0.0338	300.3522
	10000	TRI	0.0020	0.0010	0.0318	21.9067
		MEE	0.0119	0.0721	0.0324	660.1961

A Tabela (5.1) mostra os resultados do EQM para cada par (J, n) do método da TRI em comparação com o método da MEE, assim como o tempo de execução de cada método. A partir dessa tabela é observável que o EQM do parâmetro de discriminação (a) é menor no método da TRI quando o tamanho amostral é alto, e é menor pelo método da MEE quando tamanho amostral é baixo. O EQM do parâmetro de dificuldade (b) é menor no método da TRI para altos tamanhos amostrais e também para alguns baixos tamanhos amostrais. O EQM das habilidades (θ) não apresenta grande variação de um método para o outro, ou seja, ambos os métodos estimaram este parâmetro de maneira equivalente. É facilmente observado que o método da TRI apresentou menor tempo de processamento (T) para todo par (J, n) .

A seguir será apresentado graficamente os EQMs de cada parâmetro, assim como o tempo de processamento (T), em duas partes: a primeira sendo os EQMs em relação ao número de itens e a segunda sendo em relação ao número de indivíduos.

5.2 Resultados dos EQMs da TRI e da MEE

Os gráficos a seguir apresentam a comparação dos EQMs obtidos pelo método da TRI (barra vermelha) com os EQMs obtidos pelo método do MEE (barra azul).

As Figuras (5.1), (5.2), (5.3), (5.4) mostram a variação do EQM do parâmetro de discriminação, do método da TRI e do método da MEE, fixando os valores de J em 10, 20, 50 e 100, respectivamente. De acordo com essas figuras foi observado que o EQM do parâmetro de discriminação pelo método da MEE foi menor do que o EQM do parâmetro de discriminação pelo método da TRI, para um tamanho de amostra pequeno. Quando o tamanho de amostra fica maior, o método com menor EQM fica sendo o método da TRI.

Figura 5.1 *EQMs do parâmetro de discriminação com tamanho de itens igual a 10.*

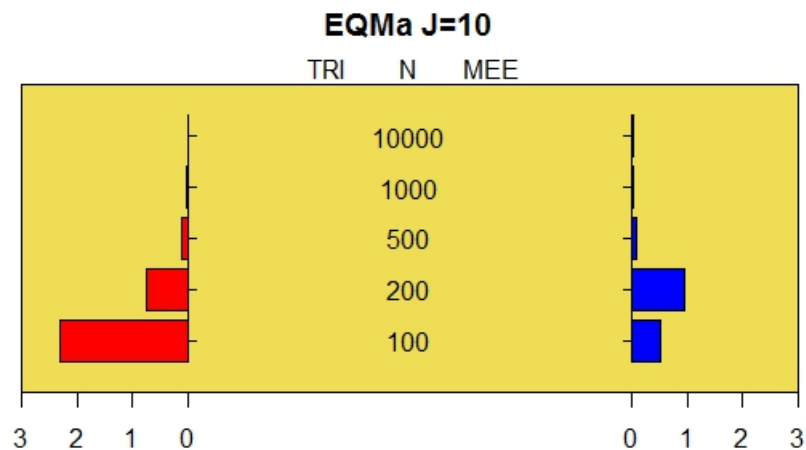


Figura 5.2 *EQMs do parâmetro de discriminação com tamanho de itens igual a 20.*

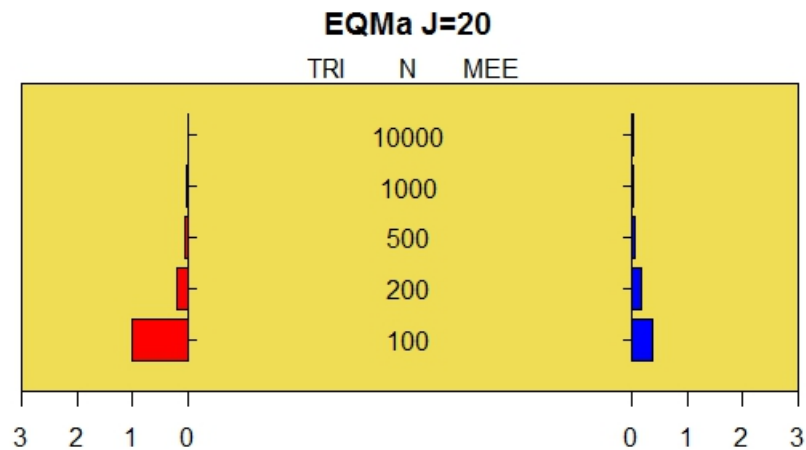
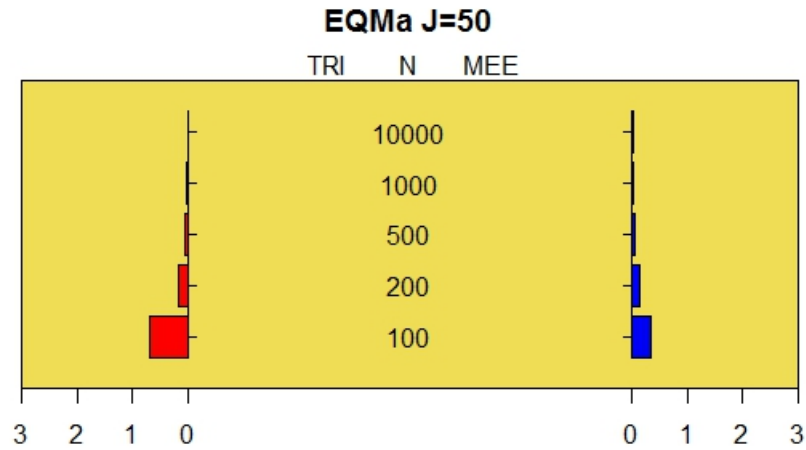
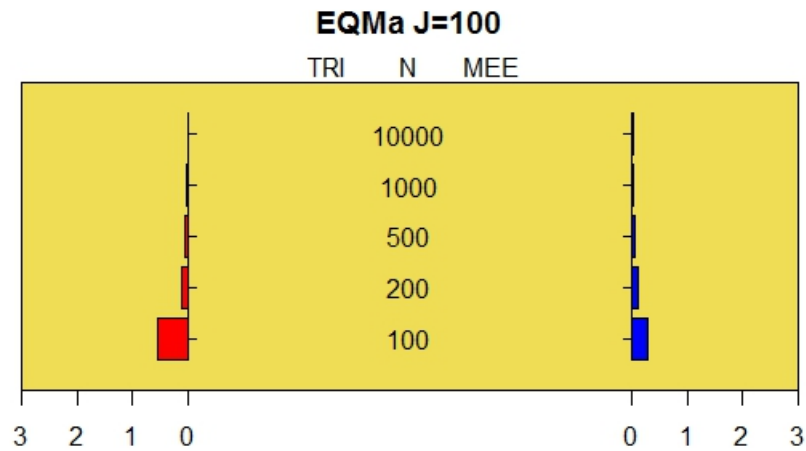


Figura 5.3 *EQMs do parâmetro de discriminação com tamanho de itens igual a 50.*Figura 5.4 *EQMs do parâmetro de discriminação com tamanho de itens igual a 100.*

As Figuras (5.5), (5.6), (5.7), (5.8) mostram a variação do EQM do parâmetro de dificuldade, do método da TRI e do método da MEE, fixando os valores de J em 10, 20, 50 e 100, respectivamente. Da mesma maneira que nas figuras (5.1), (5.2), (5.3), (5.4) o EQM do parâmetro de dificuldade pelo método da MEE foi menor do que o EQM do parâmetro de dificuldade pelo método da TRI, para um tamanho de amostra pequeno, enquanto que, quando o tamanho de amostra fica maior, o método com menor EQM fica sendo o método da TRI. Podemos notar que o EQM do parâmetro de dificuldade pelo método da MEE não sofreu variação significativa quando o tamanho do teste aumenta.

Figura 5.5 EQMs do parâmetro de dificuldade com tamanho de itens igual a 10.

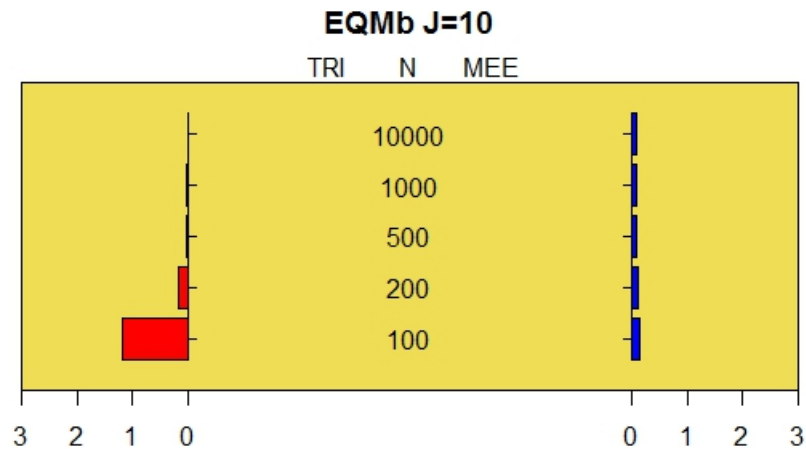


Figura 5.6 EQMs do parâmetro de dificuldade com tamanho de itens igual a 20.

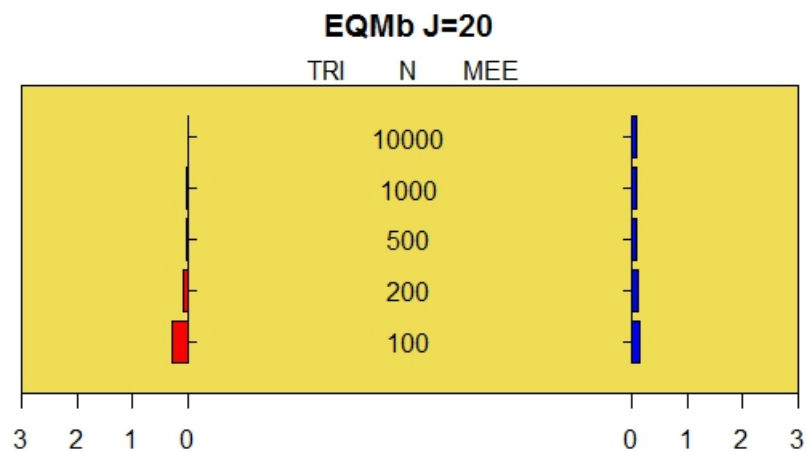


Figura 5.7 EQMs do parâmetro de dificuldade com tamanho de itens igual a 50.

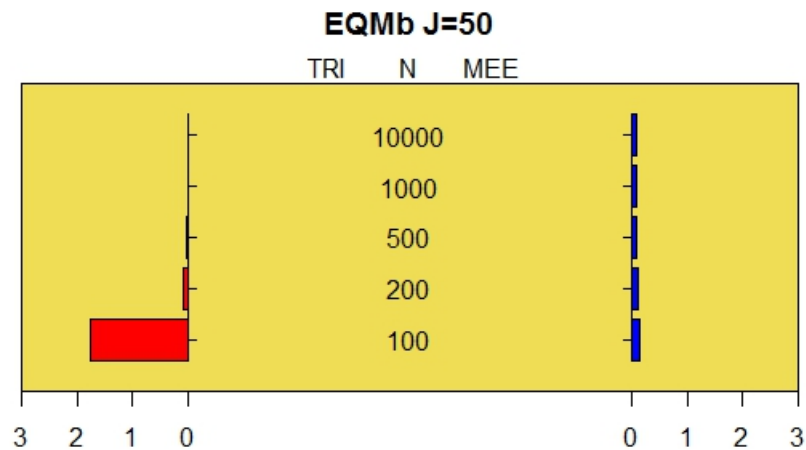
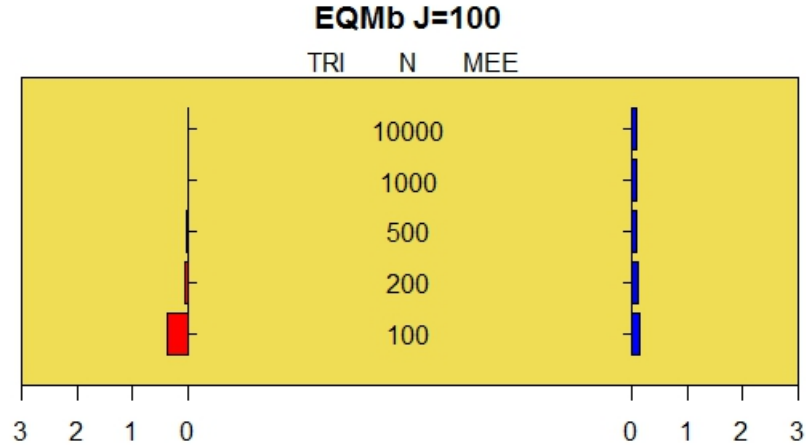


Figura 5.8 *EQMs do parâmetro de dificuldade com tamanho de itens igual a 100.*

As Figuras (5.9), (5.10), (5.11), (5.12) mostram a variação do EQM das habilidades, do método da TRI e do método da MEE, fixando os valores de J em 10, 20, 50 e 100, respectivamente. De acordo com essas figuras pode-se verificar que o EQM das habilidades obtido pelo método da MEE e o EQM das habilidades obtido pelo método da TRI foram estimados de maneira equivalente, ou seja, não houve diferença significativa no EQM das habilidades obtido por ambos os métodos.

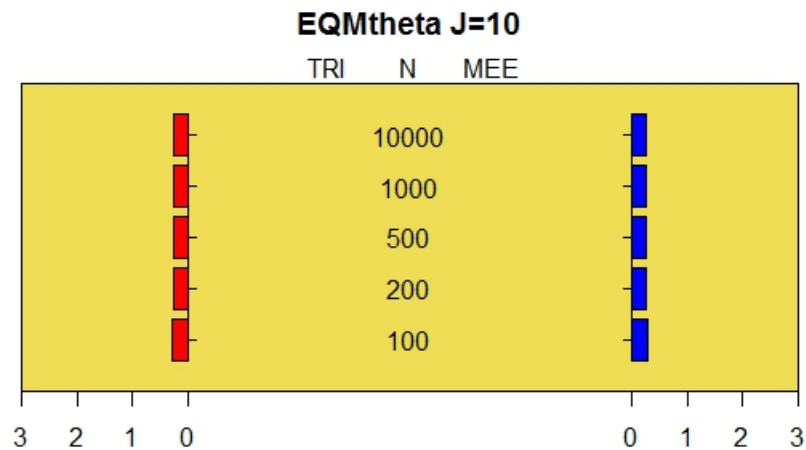
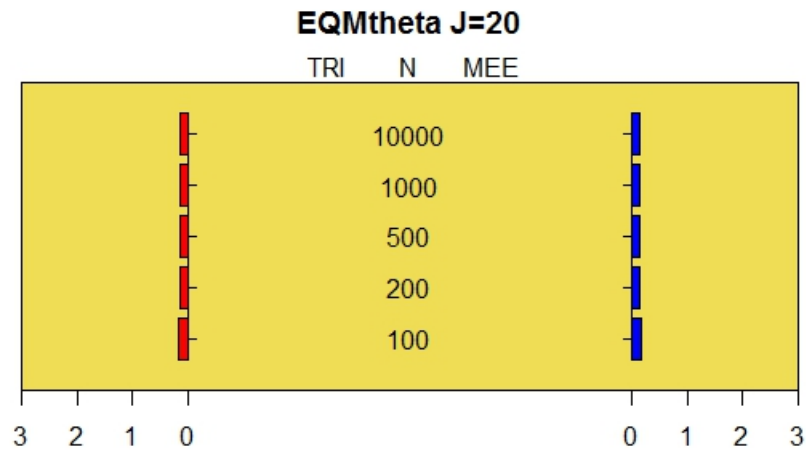
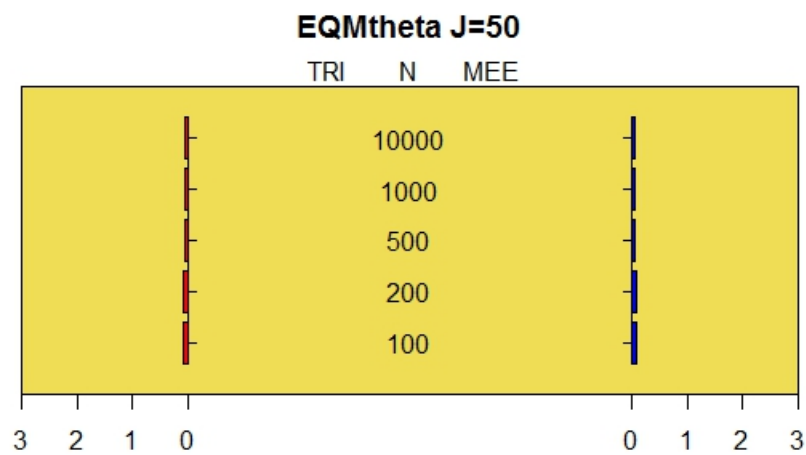
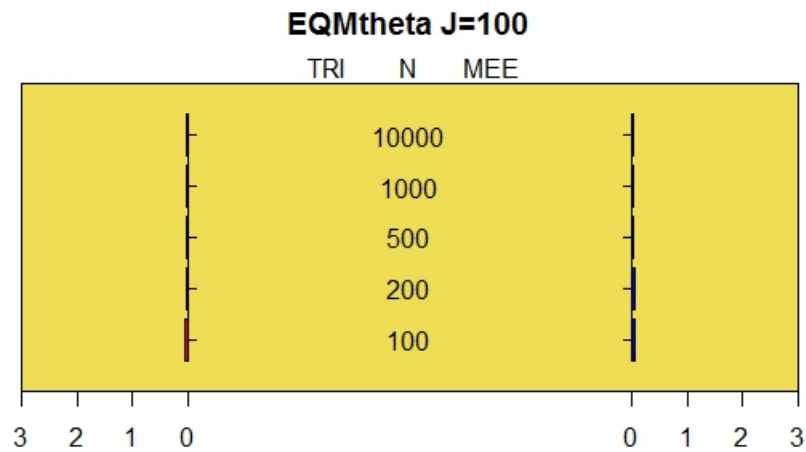
Figura 5.9 *EQMs das habilidades com tamanho de itens igual a 10.*

Figura 5.10 *EQMs das habilidades com tamanho de itens igual a 20.*Figura 5.11 *EQMs das habilidades com tamanho de itens igual a 50.*Figura 5.12 *EQMs das habilidades com tamanho de itens igual a 100.*

No próximo capítulo serão apresentadas as conclusões obtidas na elaboração deste trabalho, assim como as expectativas para trabalhos futuros.

Capítulo 6

Conclusões

6.1 Conclusões Obtidas neste Trabalho

Este trabalho teve como objetivo a comparação dos métodos Teoria da Resposta ao Item (TRI) e Modelos de Equações Estruturais (MEE). No Capítulo 1 foi apresentado um resumo em relação aos dois métodos, assim como o objetivo e a importância da realização do mesmo. No Capítulo 2 foi feita uma breve descrição da TRI, com a descrição do modelo e das funções de estimação dos parâmetros e das habilidades. No Capítulo 3 foi feita a descrição do MEE, comparando analiticamente com a TRI e mostrando a equivalência entre os métodos quando considerado o Modelo Logístico de 2 parâmetros para a TRI. No Capítulo 4 foram mostrados alguns aspectos computacionais, o programa utilizado para realizar as estimativas dos métodos (software R), tais como os pacotes utilizados no programa (Mirt e Lavaan). Foram mencionados os comandos realizados para estimar e recuperar os parâmetros, assim como uma prévia dos resultados para uma única amostra. E no Capítulo 5 foram mostrados os resultados com algumas combinações de tamanho de amostras (n) e tamanho dos testes (J) e as análises feitas a partir dos pares (J, n) .

Com todas as informações obtidas neste trabalho, agora poderá ser alcançado o seu real objetivo: qual o melhor método para estimar parâmetros de itens e variáveis latentes? De acordo com os resultados demonstrados nos Capítulos 4 e 5 foi verificado que o método da TRI foi o melhor método de estimação quando o tamanho da amostra e o tamanho do teste são altos, enquanto que o método da MEE se apresentou o melhor método quando o tamanho da amostra e o tamanho do teste são baixos.

Analisando os dois métodos por cada valor do tamanho do teste, confirmamos que, o parâmetro de discriminação foi melhor estimado pelo método da MEE quando o tamanho da amostra e o tamanho do teste são baixos, enquanto que, o método da TRI foi o melhor método de estimação quando o tamanho da amostra e o tamanho do teste são altos.

Da mesma maneira que o parâmetro de discriminação e analisando novamente os dois métodos por cada valor do tamanho do teste, é confirmado que, o parâmetro de dificuldade foi melhor estimado pelo método da MEE quando quando o tamanho da amostra e o tamanho do teste são baixos, enquanto que, o método da TRI foi o melhor método de estimação quando o tamanho da amostra e o tamanho do teste são altos.

As habilidades foram estimadas de maneira equivalente em ambos os métodos, pois não houve diferença significativa em relação aos valores do tamanho do teste e do tamanho da amostra.

O menor tempo de execução T , em cada uma das combinações (J, n) , foi obtido pelo método da TRI.

6.2 Perspectivas para Trabalhos Futuros

Para trabalhos futuros, caso se utilize esses métodos, pretende-se:

- Aplicar esses métodos utilizando grupos distintos;
- Utilizar uma habilidade multidimensional;
- Realizar uma aplicação em dados reais;
- Realizar uma análise para a estimação das habilidades extremas (acima de 2, por exemplo) dos indivíduos e verificar qual método estimou melhor.
- Verificar como otimizar a estimação das habilidades no Lavaan de forma a diminuir o tempo computacional.

Apêndice A

Desenvolvimento das questões do Capítulo 2

A.1 Expressões da seção 2.2.2

Nesta seção serão apresentadas as extensões das fórmulas da seção 2.2.2 do Capítulo 2.

Para obtermos as derivadas de (2.8) serão necessários os procedimentos a seguir:

$$\begin{aligned}\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_j} &= \frac{\partial}{\partial \zeta_j} \left\{ \sum_{i=1}^s r_i \log P(\mathbf{x}_i | \zeta, \eta) \right\} \\ &= \sum_{i=1}^s r_i \frac{1}{P(\mathbf{x}_i | \zeta, \eta)} \frac{\partial P(\mathbf{x}_i | \zeta, \eta)}{\partial \zeta_j}.\end{aligned}\tag{A.1}$$

Mas

$$\begin{aligned}\frac{\partial P(\mathbf{x}_i | \zeta, \eta)}{\partial \zeta_j} &= \frac{\partial}{\partial \zeta_j} \int_{\mathbb{R}} P(\mathbf{x}_i | \theta, \zeta) \mathbf{g}(\theta | \eta) \mathbf{d}\theta \\ &= \int_{\mathbb{R}} \left(\frac{\partial}{\partial \zeta_j} P(\mathbf{x}_i | \theta, \zeta) \right) g(\theta | \eta) d\theta \\ &= \int_{\mathbb{R}} \left(\frac{\partial}{\partial \zeta_j} \prod_{l=1}^J P(\mathbf{x}_{il} | \theta, \zeta_l) \right) g(\theta | \eta) d\theta\end{aligned}\tag{A.2}$$

$$\begin{aligned}\frac{\partial P(\mathbf{x}_i | \zeta, \eta)}{\partial \zeta_j} &= \int_{\mathbb{R}} \left(\prod_{l \neq j} P(\mathbf{x}_{il} | \theta, \zeta_l) \right) \left(\frac{\partial}{\partial \zeta_j} P(\mathbf{x}_{ij} | \theta, \zeta_j) \right) g(\theta | \eta) d\theta \\ &= \int_{\mathbb{R}} \left(\frac{\partial P(\mathbf{x}_{ij} | \theta, \zeta_j) / \partial \zeta_j}{P(\mathbf{x}_{ij} | \zeta_j)} \right) P(\mathbf{x}_i | \theta, \zeta) g(\theta | \eta) d\theta.\end{aligned}\tag{A.3}$$

Com a notação $P_j = P(\mathbf{X}_{ij} = 1 | \theta, \zeta_j)$ e $Q_j = 1 - P_j$, temos que

$$P(\mathbf{X}_{ij} = x_{ij} | \theta, \zeta_j) = P(\mathbf{X}_{ij} = 1 | \theta, \zeta_j)^{x_{ij}} P(\mathbf{X}_{ij} = 0 | \theta, \zeta_j)^{1-x_{ij}} = P_j^{x_{ij}} Q_j^{1-x_{ij}}.\tag{A.4}$$

Portanto, rescrevendo $P(\mathbf{x}_i|\theta, \zeta_j)$, teremos que

$$\begin{aligned} \frac{\partial}{\partial \zeta_j} P(\mathbf{x}_{ij}|\theta, \zeta_j) &= \left(P_j^{x_{ij}} Q_j^{1-x_{ij}} \right) \\ &= x_{ij} P_j^{x_{ij}-1} \left(\frac{\partial P_j}{\partial \zeta_j} \right) Q_j^{1-x_{ij}} + (1-x_{ij}) Q_j^{-x_{ij}} \left(-\frac{\partial P_j}{\partial \zeta_j} \right) P_j^{x_{ij}} \\ &= \left(x_{ij} P_j^{x_{ij}-1} Q_j^{1-x_{ij}} - (1-x_{ij}) Q_j^{-x_{ij}} P_j^{x_{ij}} \right) \frac{\partial P_j}{\partial \zeta_j}. \end{aligned}$$

Notemos agora que o uso entre parênteses vale 1 quando $x_{ij} = 1$ e vale -1 quando $x_{ij} = 0$, portanto podemos rescrevê-lo como $(-1)^{x_{ij}+1}$ e com isso,

$$\frac{\partial}{\partial \zeta_j} P(\mathbf{x}_{ij}|\theta, \zeta_j) = (-1)^{x_{ij}+1} \frac{\partial P_j}{\partial \zeta_j}. \quad (\text{A.5})$$

Notando agora que

$$\frac{(-1)^{x_{ij}+1} P_j Q_j}{P_j^{x_{ij}} Q_j^{1-x_{ij}}} = \begin{cases} Q_j, & x_{ij} = 1 \\ -P_j, & x_{ij} = 0, \end{cases} \quad (\text{A.6})$$

podemos escrever este termo como $x_{ij} - P_j$. Segue que (A.3) pode ser escrita como

$$\frac{\partial P(\mathbf{x}_i|\zeta, \eta)}{\partial \zeta_j} = \int_{\mathbb{R}} \left[\frac{\mathbf{x}_{ij} - P_j}{P_j Q_j} \left(\frac{\partial P_j}{\partial \zeta_j} \right) \right] P(\mathbf{x}_i|\theta, \zeta) g(\theta|\eta) d\theta. \quad (\text{A.7})$$

Usando a notação

$$g_i^*(\theta) \equiv g(\theta|x_i, \zeta, \eta) = \frac{P(\mathbf{x}_i|\theta, \zeta) g(\theta|\eta)}{P(\mathbf{x}_i|\zeta, \eta)} \quad (\text{A.8})$$

teremos que a função de verossimilhança (A.1) pode ser escrita como

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_j} = \sum_{i=1}^s r_i \int_{\mathbb{R}} \left[\frac{\mathbf{x}_{ij} - P_j}{P_j Q_j} \left(\frac{\partial P_j}{\partial \zeta_j} \right) \right] g_i^*(\theta) d\theta. \quad (\text{A.9})$$

Agora resta a obtenção das equações específicas para cada parâmetro do vetor $\zeta_j = (a_j, b_j)$. As expressões para as derivadas do modelo logístico P_i são dadas por:

$$\frac{\partial P_j}{\partial a_j} = D(\theta - b_j) P_j Q_j \quad (\text{A.10})$$

$$\frac{\partial P_j}{\partial b_j} = -D a_j P_j Q_j \quad (\text{A.11})$$

Para obter a equação de estimação para o parâmetro de discriminação, a_i , notemos que da expressão (A.9) temos que

$$\begin{aligned} \frac{\partial \log L(\zeta, \eta)}{\partial a_j} &= \sum_{i=1}^s r_i \int_{\mathbb{R}} \left[\frac{(\mathbf{x}_{ij} - P_j)}{P_j Q_j} \left(\frac{\partial P_j}{\partial a_j} \right) \right] g_i^*(\theta) d\theta \\ &= \sum_{i=1}^s r_i \int_{\mathbb{R}} \left[\frac{(\mathbf{x}_{ij} - P_j)}{P_j Q_j} D(\theta - b_j) P_j Q_j \right] g_i^*(\theta) d\theta \\ &= D \sum_{i=1}^s r_i \int_{\mathbb{R}} [(\mathbf{x}_{ij} - P_j)(\theta - b_j)] g_i^*(\theta) d\theta \end{aligned} \quad (\text{A.12})$$

Para obter a equação de estimação para o parâmetro de dificuldade, b_i , temos que

$$\begin{aligned} \frac{\partial \log L(\zeta, \eta)}{\partial b_j} &= \sum_{i=1}^s r_i \int_{\mathbb{R}} \left[\frac{(\mathbf{x}_{ij} - P_j)}{P_j Q_j} \left(\frac{\partial P_j}{\partial b_j} \right) \right] g_i^*(\theta) d\theta \\ &= \sum_{i=1}^s r_i \int_{\mathbb{R}} \left[\frac{(\mathbf{x}_{ij} - P_j)}{P_j Q_j} (-1) D a_j P_j Q_j \right] g_i^*(\theta) d\theta \\ &= -D a_j \sum_{i=1}^s r_i \int_{\mathbb{R}} [(\mathbf{x}_{ij} - P_j)] g_i^*(\theta) d\theta \end{aligned} \quad (\text{A.13})$$

Em resumo, as equações de estimação para os parâmetros a_j e b_j são respectivamente,

$$a_j : D \sum_{i=1}^s r_i \int_{\mathbb{R}} [(\mathbf{x}_{ij} - P_j)(\theta - b_j)] g_i^*(\theta) d\theta = 0 \quad (\text{A.14})$$

$$b_j : -D a_j \sum_{i=1}^s r_i \int_{\mathbb{R}} [(\mathbf{x}_{ij} - P_j)] g_i^*(\theta) d\theta = 0 \quad (\text{A.15})$$

Estas equações não possuem solução analítica e devido a isso precisaremos de algum método iterativo para a obtenção das estimativas de máxima verossimilhança marginal dos parâmetros dos itens. Neste trabalho será abordado o método de Newton-Raphson para a obtenção dessas estimativas.

A.1.1 Algoritmo Newton-Raphson

Seja $l(\zeta) = \log L(\zeta)$ a log-verossimilhança onde $\zeta = (\zeta_1, \dots, \zeta_J)$, com $\zeta_j = (a_j, b_j)$. Considerando valores iniciais $\widehat{\zeta}_j^{(0)} = (a_j^{(0)}, b_j^{(0)})$ encontrados para ζ_j , então uma estimativa atualizada será, $(\widehat{\zeta}_j^{(1)}) = (\widehat{\zeta}_j^{(0)}) + (\Delta \widehat{\zeta}_j^{(0)})$, ou seja,

$$\begin{aligned}\hat{a}_j^{(1)} &= \hat{a}_j^{(0)} + \Delta\hat{a}_j^{(0)}, \\ \hat{b}_j^{(1)} &= \hat{b}_j^{(0)} + \Delta\hat{b}_j^{(0)},\end{aligned}\tag{A.16}$$

onde $\Delta\hat{a}_j^{(0)}$ e $\Delta\hat{b}_j^{(0)}$ são erros de aproximação. Usando a expansão em séries de Taylor de $\frac{\partial l(\zeta)}{\partial \zeta_j}$ em torno de $\hat{\zeta}_j^{(0)}$, teremos

$$\begin{aligned}\frac{\partial l(\zeta)}{\partial a_j} &= \frac{\partial l(\hat{\zeta}_j^{(0)})}{\partial a_j} + \Delta\hat{a}_j^{(0)} \frac{\partial^2 l(\hat{\zeta}_j^{(0)})}{\partial a_j^2} + \Delta\hat{b}_j^{(0)} \frac{\partial^2 l(\hat{\zeta}_j^{(0)})}{\partial a_j \partial b_j} + R_{a_j}(\hat{\zeta}_j^{(0)}), \\ \frac{\partial l(\zeta)}{\partial b_j} &= \frac{\partial l(\hat{\zeta}_j^{(0)})}{\partial b_j} + \Delta\hat{a}_j^{(0)} \frac{\partial^2 l(\hat{\zeta}_j^{(0)})}{\partial b_j \partial a_j} + \Delta\hat{b}_j^{(0)} \frac{\partial^2 l(\hat{\zeta}_j^{(0)})}{\partial b_j^2} + R_{b_j}(\hat{\zeta}_j^{(0)}),\end{aligned}$$

Nessas expressões usa-se que $\frac{\partial l(\zeta)}{\partial \zeta_j}$ é função apenas de ζ_j , não dependendo de ζ_l para $l \neq j$. Por isso, poderemos representá-la de forma simplificada por $\frac{\partial l(\zeta_j)}{\partial \zeta_j}$. Fazendo

$$\frac{\partial l(\zeta_j)}{\partial a_j} = \frac{\partial l(\zeta_j)}{\partial b_j} = 0,$$

usando a notação

$$\begin{aligned}L_1 &= \frac{\partial l(\hat{\zeta}_j^{(0)})}{\partial a_j}; L_{11} = \frac{\partial^2 l(\hat{\zeta}_j^{(0)})}{\partial a_j^2}; L_{12} = \frac{\partial^2 l(\hat{\zeta}_j^{(0)})}{\partial a_j \partial b_j}; \\ L_2 &= \frac{\partial l(\hat{\zeta}_j^{(0)})}{\partial b_j}; L_{21} = \frac{\partial^2 l(\hat{\zeta}_j^{(0)})}{\partial b_j \partial a_j}; L_{22} = \frac{\partial^2 l(\hat{\zeta}_j^{(0)})}{\partial b_j^2};\end{aligned}$$

e desprezando os restos $R_{a_j}(\hat{\zeta}_j^{(0)})$ e $R_{b_j}(\hat{\zeta}_j^{(0)})$, teremos

$$\begin{aligned}0 &= L_1 + L_{11}\Delta\hat{a}_j^{(0)} + L_{12}\Delta\hat{b}_j^{(0)}, \\ 0 &= L_2 + L_{21}\Delta\hat{a}_j^{(0)} + L_{22}\Delta\hat{b}_j^{(0)},\end{aligned}$$

Colocando o resultado em forma matricial, teremos

$$-\begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} \Delta\hat{a}_j^{(0)} \\ \Delta\hat{b}_j^{(0)} \end{pmatrix}.$$

Resolvendo o sistema para $\Delta\widehat{\zeta}_j^{(0)}$, teremos

$$\begin{pmatrix} \Delta\hat{a}_j^{(0)} \\ \Delta\hat{b}_j^{(0)} \end{pmatrix} = - \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}^{-1} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix},$$

e finalmente, pela equação (A.16), teremos

$$\begin{pmatrix} \hat{a}_j^{(1)} \\ \hat{b}_j^{(1)} \end{pmatrix} = \begin{pmatrix} \hat{a}_j^{(0)} \\ \hat{b}_j^{(0)} \end{pmatrix} - \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}^{-1} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}.$$

Após obtido $\widehat{\zeta}_j^{(1)}$, este é considerado um novo ponto inicial para a obtenção de $\widehat{\zeta}_j^{(2)}$, e assim por diante. Este processo é repetido até que algum critério de parada seja alcançado. Como por exemplo, até que $\Delta\widehat{\zeta}_j^{(t)} = \widehat{\zeta}_j^{(t)} - \widehat{\zeta}_j^{(t-1)}$ seja suficientemente pequeno, se aproximando de 0, ou que um número pré-definido de t interações seja alcançado.

Para a aplicação do algoritmo Newton-Raphson, é preciso das derivadas segundas de $L(\zeta, \eta)$. As expressões para as derivadas segundas são obtidas a partir de

$$\begin{aligned} \frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta'_j} &= \frac{\partial}{\partial \zeta_l} \left[\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_j} \right]' \\ &= \frac{\partial}{\partial \zeta_l} \left[\sum_{i=1}^s r_i \frac{1}{P(\mathbf{x}_i | \zeta, \eta)} \frac{\partial P(\mathbf{x}_i | \zeta, \eta)}{\partial \zeta_j} \right]' \\ &= \sum_{i=1}^s r_i \left\{ \frac{\partial^2 P(\mathbf{x}_i | \zeta, \eta) / (\partial \zeta_l \partial \zeta'_j)}{P(\mathbf{x}_i | \zeta, \eta)} - \left(\frac{\partial P(\mathbf{x}_i | \zeta, \eta) / \partial \zeta_l}{P(\mathbf{x}_i | \zeta, \eta)} \right) \left(\frac{\partial P(\mathbf{x}_i | \zeta, \eta) / \partial \zeta'_j}{P(\mathbf{x}_i | \zeta, \eta)} \right)' \right\} \end{aligned} \quad (\text{A.17})$$

para $j, l = 1, \dots, J$.

Considerando $\widehat{\zeta}_j^{(t)}$ a estimativa de ζ na iteração t , então na iteração $t + 1$ teremos que

$$\widehat{\zeta}_j^{(t+1)} = \widehat{\zeta}_j^{(t)} - [\mathbf{H}_{\mathbf{PI}}(\widehat{\zeta}_j^{(t)})]^{-1} \mathbf{h}_{\mathbf{PI}}(\widehat{\zeta}_j^{(t)}), \quad (\text{A.18})$$

onde por (A.9) teremos,

$$\begin{aligned} \mathbf{h}(\zeta_j) &\equiv \frac{\partial \log L(\zeta, \eta)}{\partial \zeta_j} \\ &\equiv \sum_{i=1}^s r_j \int_{\mathbb{R}} \{(\mathbf{x}_{ij} - P_j)\} \mathbf{h}_j g_i^*(\theta) d\theta, \end{aligned} \quad (\text{A.19})$$

de mesmo modo por (A.17) teremos,

$$\begin{aligned}
\mathbf{H}(\zeta_j) &\equiv \frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta'_j} \\
&\equiv \sum_{i=1}^s r_j \{ \mathbf{H}_{j\mathbf{l}(i)} - \mathbf{h}_{j\mathbf{l}(i)} \mathbf{h}'_{\mathbf{l}(i)} \}.
\end{aligned} \tag{A.20}$$

A.2 Expressões da seção 2.2.3

Nesta seção serão apresentadas as extensões das fórmulas da seção 2.2.3 do Capítulo 2.

Para obtermos a derivada de (2.10) serão necessários os procedimentos a seguir: Podemos notar pela equação (2.9), que

$$\begin{aligned}
\frac{\partial \log L(\theta)}{\partial \theta_i} &= \sum_{j=1}^J \left\{ \mathbf{x}_{ij} \frac{\partial \log P_{ij}}{\partial \theta_i} + (1 - \mathbf{x}_{ij}) \frac{\partial \log Q_{ij}}{\partial \theta_i} \right\} \\
&= \sum_{j=1}^J \left\{ \mathbf{x}_{ij} \frac{1}{P_{ij}} \frac{\partial P_{ij}}{\partial \theta_i} - (1 - \mathbf{x}_{ij}) \frac{1}{Q_{ij}} \frac{\partial P_{ij}}{\partial \theta_i} \right\} \\
&= \sum_{j=1}^J \left\{ \mathbf{x}_{ij} \frac{1}{P_{ij}} - (1 - \mathbf{x}_{ij}) \frac{1}{Q_{ij}} \right\} \left\{ \frac{\partial P_{ij}}{\partial \theta_i} \right\}, \\
&= \sum_{j=1}^J \left\{ \frac{(\mathbf{x}_{ij} - P_{ij})}{P_{ij} Q_{ij}} \right\} \left\{ \frac{\partial P_{ij}}{\partial \theta_i} \right\},
\end{aligned} \tag{A.21}$$

Como

$$\frac{\partial P_{ij}}{\partial \theta_i} = D a_j P_{ij} Q_{ij}, \tag{A.22}$$

obtêm-se

$$\begin{aligned}
\frac{\partial \log L(\theta)}{\partial \theta_i} &= \sum_{j=1}^J \left\{ \frac{(\mathbf{x}_{ij} - P_{ij})}{P_{ij} Q_{ij}} \right\} \left\{ \frac{\partial P_{ij}}{\partial \theta_i} \right\} \\
&= \sum_{j=1}^J \left\{ (\mathbf{x}_{ij} - P_{ij}) D a_j P_{ij} Q_{ij} \frac{1}{P_{ij} Q_{ij}} \right\}, \\
&= D \sum_{j=1}^J a_j (\mathbf{x}_{ij} - P_{ij}),
\end{aligned} \tag{A.23}$$

Então a equação de estimação (2.10) para $\theta_i, i = 1, \dots, n$, é

$$\theta_i : D \sum_{j=1}^J a_j (\mathbf{x}_{ij} - P_{ij}) = 0. \quad (\text{A.24})$$

Assim como na seção anterior, a equação de estimação não apresenta solução explícita para θ_i , e por isso, é preciso de algum método iterativo para obter as estimativas desejadas. Mais uma vez será apresentado o método de Newton-Raphson a fim de obter as expressões desejadas.

A.2.1 Algoritmo de Newton-Raphson

De forma similar ao que foi feito na seção A.1.1, e considerando $\hat{\theta}_j^{(t)}$ a estimativa de θ_j na iteração t , então na iteração $t + 1$ do algoritmo de Newton-Raphson teremos que

$$\hat{\theta}_j^{(t+1)} = \hat{\theta}_j^{(t)} - [H(\hat{\theta}_j^{(t)})]^{-1} h(\hat{\theta}_j^{(t)}). \quad (\text{A.25})$$

Para obter as expressões de $H(\theta_i)$ e $h(\theta_i)$ usadas em (A.25), notemos que de (A.21),

$$\begin{aligned} \frac{\partial^2 \log L(\theta)}{\partial \theta_i^2} &= \sum_{j=1}^J \left\{ \left[\frac{\partial}{\partial \theta_i} \left(\frac{\mathbf{x}_{ij} - P_{ij}}{P_{ij} Q_{ij}} \right) \right] \left(\frac{\partial P_{ij}}{\partial \theta_i} \right) + \left(\frac{\mathbf{x}_{ij} - P_{ij}}{P_{ij} Q_{ij}} \right) \left(\frac{\partial P_{ij}}{\partial \theta_i} \right) \right\} \\ &= \sum_{j=1}^J \left\{ \left(\frac{\mathbf{x}_{ij} - \mathbf{P}_{ij}}{P_{ij} Q_{ij}} \right) \left(\frac{\partial^2 P_{ij}}{\partial \theta_i^2} \right) - \left(\frac{\mathbf{x}_{ij} - \mathbf{P}_{ij}}{P_{ij} Q_{ij}} \right)^2 \left(\frac{\partial P_{ij}}{\partial \theta_i} \right)^2 \right\}. \end{aligned} \quad (\text{A.26})$$

A segunda parcela em (A.26) é obtida por (A.27). Com relação a essa temos,

$$\frac{\partial^2 P_{ij}}{\partial \theta_i^2} = D^2 a_j^2 P_{ij} Q_{ij} (1 - 2P_{ij}). \quad (\text{A.27})$$

Sejam

$$h_{ij} = (P_{ij} Q_{ij})^{-1} \left(\frac{\partial P_{ij}}{\partial \theta_i} \right) = D a_j, \quad (\text{A.28})$$

$$H_{ij} = (P_{ij} Q_{ij})^{-1} \left(\frac{\partial^2 P_{ij}}{\partial \theta_i^2} \right) = D^2 a_j^2 (1 - 2P_{ij}). \quad (\text{A.29})$$

Com isso, de (A.23) temos que

$$\begin{aligned}
h(\theta_i) &\equiv \frac{\partial \log L(\theta)}{\partial \theta_i} \\
&= \sum_{j=1}^J \left\{ (\mathbf{x}_{ij} - P_{ij}) \frac{1}{P_{ij} Q_{ij}} \right\} (P_{ij} Q_{ij}) h_{ij} \\
&= \sum_{j=1}^J \{ (\mathbf{x}_{ij} - P_{ij}) \} h_{ij},
\end{aligned} \tag{A.30}$$

e retornando a (A.26)

$$\begin{aligned}
H(\theta_i) &\equiv \frac{\partial^2 \log L(\theta)}{\partial \theta_i^2} \\
&= \sum_{j=1}^J \left\{ \left(\frac{\mathbf{x}_{ij} - P_{ij}}{P_{ij} Q_{ij}} \right) (P_{ij} Q_{ij}) H_{ij} - \left(\frac{\mathbf{x}_{ij} - P_{ij}}{P_{ij} Q_{ij}} \right)^2 (P_{ij} Q_{ij})^2 h_{ij}^2 \right\} \\
&= \sum_{j=1}^J (\mathbf{x}_{ij} - P_{ij}) \{ H_{ij} - (\mathbf{x}_{ij} - P_{ij}) h_{ij}^2 \}.
\end{aligned} \tag{A.31}$$

Apêndice B

Algoritmo usado na simulação

Neste Apêndice serão apresentados os algoritmos utilizados neste trabalho. O algoritmo a seguir, consiste na geração de um banco de dados dicotômicos com os valores dos parâmetros de interesse do modelo logístico. A geração dos dados dicotômicos desse banco é aleatória, de forma que os parâmetros de interesse mudam a cada geração do banco de dados.

Os métodos da TRI e MEE utilizaram os bancos de dados gerados a partir desse algoritmo para realizar seus respectivos processos de estimação.

Os parâmetros obtidos neste algoritmo, foram utilizados com os parâmetros estimados pelos métodos da TRI e MEE, nas equações do Erro Quadrático Médio, com o intuito de verificar qual método apresentou o menor erro.

B.1 Geração do Bando de Dados

```
PL=2 # Logistic Model (1,2,3 parameters)

I= 30 # Número de Itens
n = 100000
a = runif(I,0.5, 2.0) # U(0.5 , 2.0)
b = rnorm(I,0.0, 1.0) # N(0 , 1)
c = runif(I,0.0, 0.25) # U(0 , 0.25)
d=-a*b # Intercepto

if (PL==2) c=c*0; if (PL==1) {c=c*0; a=rep(1,I)}

theta = rnorm(n,0,1) # N(0 , 1)
eta = theta%% t(a) + matrix(d,n,I,byrow=TRUE) # n x I
```

```

P = c + (1-c)/(1+exp(-1.7*eta)) # n x I
X = runif(n*I) # Uniform(0,1)
dim(X)=c(n,I); dim(X)=c(n,I) # matriz n x I de U(0,1)
U = 1*(X<P) ; U=as.data.frame(U)
colnames(U)=paste0("Item.",1:I)

rm(P,X,eta)

```

B.2 MIRT

O algoritmo a seguir utilizou o pacote `mirt` para realizar o processo de estimação pelo método da TRI.

```

#####
# CALIBRAÇÃO DOS DADOS PELO MIRT
#####

if(!require(mirt)){install.packages("mirt"); library(mirt) }

mirt.2PL = mirt(U, 1, itemtype = '2PL')
PAR=coef(mirt.2PL,simplify=TRUE)$items[,1:2] ## Parâmetros da Tabela

# Estimativa da discriminação
a_est=PAR[,1]
a_mirt=a_est/1.7
cbind(a,a_mirt)
plot(a,a_mirt,main="Recuperação de Parâmetros de Discriminação pelo Mirt",
xlab="Parâmetros Verdadeiros",
ylab="Parâmetros Estimados");
lines(c(0,3),c(0,3))
grid(nx = 6, ny = 6, col = "lightgray",
lty = "dotted", lwd = par("lwd"), equilogs = TRUE)
line(a,a_mirt)

```



```
# Estimativa da dificuldade
d_est=PAR[,2]
b_est=-d_est/a_est
#cbind(b,b_est);
plot(b,b_est,main="Recuperação de Parâmetros de Dificuldade pelo Mirt",
xlab="Parâmetros Verdadeiros",
ylab="Parâmetros Estimados");
lines(c(-3,3),c(-3,3));
grid(nx = 8, ny = 8, col = "lightgray",

lty = "dotted", lwd = par("lwd"), equilog = TRUE)
line(b,b_est)

# ..... Estimação dos thetas com o mirt .....
sco.mirt=fscores(mirt.2PL) ## Scores
hist(sco.mirt); mean(sco.mirt); sd(sco.mirt)
plot(theta,sco.mirt,main="Recuperação das Habilidades pelo Mirt",
xlab="Habilidades Verdadeiras",ylab="Habilidades Estimadas");
lines(c(-3,3),c(-3,3));
grid(nx = 8, ny = 8, col = "lightgray",
lty = "dotted", lwd = par("lwd"), equilog = TRUE)

# Cálculo dos EQMs pelo mirt
EQMa_mirt=mean((a_mirt-a)^2);
EQMb_mirt=mean((b_est-b)^2);
EQMt_mirt=mean((sco.mirt-theta)^2);
```

B.3 LAVAAN

O próximo algoritmo utilizou o pacote lavaan para realizar o processo de estimação pelo método do MEE.

```
#####
# ESTIMAÇÃO PELO LAVAAN
#####
if(!require(lavaan)){install.packages("lavaan"); library(lavaan)}

# single factor model
model= ' model =~ Item.1 ' ; for (i in 2:I) model=paste0(model,' + Item.',i)

# binary CFA
fit = cfa(model, data=U, ordered=names(U)); summary(fit, fit.measures=TRUE)

# extract partable
partable <- parTable(fit)

Latent=partable[1:I,14]; Latent
Thresholds=partable[(I+1):(2*I),14]; Thresholds
Variances=partable[(2*I+1):(3*I),14]; Variances
Lambda=sqrt(1-Variances)

# Estimativa da dificuldade
Delta=Thresholds/Lambda; Delta
cbind(b,Delta)

# Estimativa da discriminação
Beta=Lambda/sqrt(Variances); Beta
cbind(a,Beta)

plot(a,Beta,main="Recuperação de Parâmetros de Discriminação pelo Lavaan",
xlab="Parâmetros Verdadeiros",
```

```

ylab="Parâmetros Estimados");
lines(c(0,3),c(0,3));
grid(nx = 8, ny = 8, col = "lightgray",
lty = "dotted", lwd = par("lwd"), equilogs = TRUE)
line(a,Beta)

```

```

plot(b,Delta,main="Recuperação de Parâmetros de Dificuldade pelo Lavaan",
xlab="Parâmetros Verdadeiros",
ylab="Parâmetros Estimados");
lines(c(-3,3),c(-3,3));
grid(nx = 8, ny = 8, col = "lightgray",
lty = "dotted", lwd = par("lwd"), equilogs = TRUE)
line(b,Delta)

```

```

## ..... Estimação dos thetas com o mirt .....

```

```

sco.lavaan <- predict(fit) # Estimação dos thetas
#hist(sco.lavaan); mean(sco.lavaan); sd(sco.lavaan)
#plot(theta,sco.lavaan, main="Recuperação das Habilidades pelo Lavaan",
xlab="Habilidades Verdadeiras",
ylab="Habilidades Estimadas");
#lines(c(-3,3),c(-3,3));
grid(nx = 8, ny = 8, col = "lightgray",
lty = "dotted", lwd = par("lwd"), equilogs = TRUE)
#.....

```

```

##### OBTENDO A RELAÇÃO DA REPARAMETRIZAÇÃO #####

```

```

# PADRONIZANDO OS SCO DO LAVAAN

```

```

m= mean(sco.lavaan); s=sd(sco.lavaan);

```

```

sco.lavaanp=(sco.lavaan-m)/s; #score padronizado

```

```

hist(sco.lavaanp); mean(sco.lavaanp); sd(sco.lavaanp);

```

```
plot(theta,sco.lavaanp, main="Recuperação das Habilidades pelo Lavaan",
      xlab="Habilidades Verdadeiras",
      ylab="Habilidades Estimadas");
lines(c(-3,3),c(-3,3));
grid(nx = 8, ny = 8, col = "lightgray",
      lty = "dotted", lwd = par("lwd"), equilogs = TRUE)
```

```
EQMa_lavaan=mean((Beta-a)^2);
EQMb_lavaan=mean((Delta-b)^2);
EQMt_lavaan=mean((sco.lavaanp-theta)^2);
```

Referências Bibliográficas

- Andrade, D., Tavares, H., and Valle, R. (2000). Introdução a teoria da resposta ao item: Conceitos e aplicações. 14o sinape: Caxambu.
- Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual review of psychology*, 53(1):605–634.
- Brei, V. A. and Liberali Neto, G. (2006). O uso da técnica de modelagem em equações estruturais na área de marketing: um estudo comparativo entre publicações no brasil e no exterior. *Revista de Administração Contemporânea*, 10(4):131–151.
- Campana, A., Tavares, M., and Silva, D. (2009). Modelagem de equações estruturais: Apresentação de uma abordagem estatística multivariada para pesquisas em educação física. *Motricidade*, 5(4):59–80.
- Chalmers, R. P. et al. (2012). mirt: A multidimensional item response theory package for the r environment. *Journal of Statistical Software*, 48(6):1–29.
- de Andrade, D. F., Tavares, H. R., and da Cunha Valle, R. (2000). Teoria da resposta ao item: conceitos e aplicações. *ABE, Sao Paulo*.
- Glockner-Rist, A. and Hoijtink, H. (2003). The best of both worlds: Factor analysis of dichotomous data using item response theory and structural equation modeling. *Structural Equation Modeling*, 10(4):544–565.
- Isaac, E. (1966). On and hb keller. *Analysis of Numerical Methods*.
- Lord, F. (1952). A theory of test scores. *Psychometric monographs*.
- Marôco, J. (2010). *Análise de equações estruturais: Fundamentos teóricos, software & aplicações*. ReportNumber, Lda.

-
- Mueller, R. O. (1999). *Basic principles of structural equation modeling: An introduction to LISREL and EQS*. Springer Science & Business Media.
- Pasquali, L. and Primi, R. (2003). Fundamentos da teoria da resposta ao item: Tri. *Avaliação Psicológica*, 2(2):99–110.
- Pilati, R. and Laros, J. A. (2007). Modelos de equações estruturais em psicologia: conceitos e aplicações. *Psicologia: teoria e pesquisa*, 23(2):205–216.
- Rosseel, Y. (2012). Lavaan: An r package for structural equation modeling and more. version 0.5–12 (beta). *Ghent, Belgium: Ghent University*.
- Soares, T. M. (2005). Utilização da teoria da resposta ao item na produção de indicadores sócio-econômicos. *Pesquisa Operacional*, 25(1):83–112.
- Takane, Y. and De Leeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika*, 52(3):393–408.
- Tavares, H. R., de Carvalho, A. M., and Leite, W. L. (2016). Curva de crescimento em matemática via teoria da resposta ao item. *Estudos em Avaliação Educacional*, 27(64):242–266.
- Van Montfort, K., Oud, J., and Satorra, A. (2004). *Recent developments on structural equation models: Theory and applications*, volume 19. Springer Science & Business Media.