

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Estudo comparativo de métodos de estimação do modelo de resposta gradual para dados de burnout em enfermeiras

Juliana Marambaia Maia

Dissertação de Mestrado do Programa Interinstitucional de Pós-Graduação em Estatística (PIPGES)

Juliana Marambaia Maia

Estudo comparativo de métodos de estimação do modelo de resposta gradual para dados de burnout em enfermeiras

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Mestra em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística. VERSÃO REVISADA

Área de Concentração: Estatística

Orientadora: Profa. Dra. Cibele Maria Russo Novelli

**USP – São Carlos
Maio de 2020**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

M217e Maia, Juliana Marambaia
 Estudo comparativo de métodos de estimação do
 modelo de resposta gradual para dados de burnout em
 enfermeiras / Juliana Marambaia Maia; orientadora
 Cibele Maria Russo Novelli. -- São Carlos, 2020.
 104 p.

 Dissertação (Mestrado - Programa
 Interinstitucional de Pós-graduação em Estatística) --
 Instituto de Ciências Matemáticas e de Computação,
 Universidade de São Paulo, 2020.

 1. Teoria de resposta ao item. 2. Modelo de
 resposta gradual. 3. Inferência Bayesiana. 4.
 Inferência clássica. 5. Síndrome de burnout. I.
 Novelli, Cibele Maria Russo, orient. II. Título.

Juliana Marambaia Maia

**Comparative study of methods for estimating the gradual
response model for burnout data in nurses**

Master dissertation submitted to the Institute of
Mathematics and Computer Sciences – ICMC- USP and
to the Department of Statistics – DEs- UFSCar, in partial
fulfillment of the requirements for the degree of the
Master Interagency Program Graduate in Statistics.
FINAL VERSION

Concentration Area: Statistics

Advisor: Profa. Dra. Cibele Maria Russo Novelli

**USP – São Carlos
May 2020**



Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado da candidata Juliana Marambaia Maia, realizada em 27/04/2020:

Profa. Dra. Cibele Maria Russo Novelli
USP

Prof. Dr. Marcelo Andrade da Silva

Profa. Dra. Rosineide Fernando da Paz
UFC

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Cibele Maria Russo Novelli Marcelo Andrade da Silva, Rosineide Fernando da Paz e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ão) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.

Profa. Dra. Cibele Maria Russo Novelli

Este trabalho é dedicado a toda população que torna possível o acesso a educação e ciência nas universidades públicas do Brasil.

AGRADECIMENTOS

Agradeço a toda minha família por estar ao meu lado em todos os momentos e me apoiar em todas as decisões. Em especial as mulheres que me inspiram: minha mãe Zorá, minha irmã Joana, minhas sobrinhas Nana e Marina, minhas tias Rosa e Regi e minha avó Regina que está no meu coração e na minha memória em todos os momentos.

A minha orientadora Cibele Russo pelos ensinamentos, apoio, paciência, sugestões e críticas construtivas.

Ao meu namorado Kaique pelo amor, companheirismo, suporte e pelas dicas de programação.

As amizades especiais cultivadas ao longo da vida e as que pude fazer em São Carlos pelos momentos de estudo e descontração, que são essenciais para nossa saúde mental.

Aos professores que contribuíram para a minha formação pelos ensinamentos, dedicação e conselhos.

Ao HPC - High Performance Computing, pois esta pesquisa foi desenvolvida com o auxílio dos recursos de HPC disponibilizados pela Superintendência de Tecnologia da Informação da Universidade de São Paulo. Estendo o agradecimento à equipe pelo apoio técnico dado via e-mail e telefone.

Aos funcionários das secretarias da pós-graduação do ICMC-USP e do PIPGES na UFSCar por todo suporte dado quando necessário.

A CAPES pelo auxílio financeiro concedido para o desenvolvimento deste trabalho.

Por fim, agradeço a todos que direta ou indiretamente contribuem e lutam pelo ensino público gratuito de qualidade e pela pesquisa brasileira.

"Educaão nunca foi despesa. Sempre foi investimento com retorno garantido."

(Sir Arthur Lewis)

RESUMO

MAIA, J. M.. **Estudo comparativo de métodos de estimação do modelo de resposta gradual para dados de burnout em enfermeiras**. 2020. 104 f. Dissertação (Mestrado em em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

O modelo de resposta gradual é um modelo da Teoria de Resposta ao Item (TRI) para itens com resposta politômica amplamente conhecido na literatura. A motivação para este trabalho parte de um conjunto de dados pertencente ao projeto RN4CAST relacionado à síndrome de *burnout*, que é uma síndrome psicológica de configuração multidimensional de exaustão emocional, despersonalização e realização pessoal reduzida. As medidas de *burnout* são baseadas no Inventário de *burnout* de Maslach, um questionário de 22 itens para ser respondido em uma escala Likert de 7 pontos. O objetivo deste trabalho é o estudo do modelo de resposta gradual unidimensional e multiunidimensional, sob as abordagens clássica e Bayesiana, motivado pelos dados da síndrome de *burnout*, e para isso realize-se um estudo de simulação para verificar o comportamento de forma empírica do modelo. Os modelos são ajustados considerando o método da máxima verossimilhança marginal e o Monte Carlo cadeia de Markov via amostrador de Gibbs. O estudo de simulação mostra que, em geral, o método Monte Carlo via cadeia de Markov amostrador de Gibbs produz bons resultados na estimação dos parâmetros dos itens e do traço latente do modelo de resposta gradual unidimensional. Os resultados do estudo do modelo de resposta gradual multiunidimensional apresenta bons resultados à medida que o tamanho amostral e o do teste aumentam.

Palavras-chave: Teoria de resposta ao item, modelo de resposta gradual, inferência Bayesiana, inferência clássica, síndrome de *burnout*.

ABSTRACT

MAIA, J. M.. **Estudo comparativo de métodos de estimação do modelo de resposta gradual para dados de burnout em enfermeiras**. 2020. 104 f. Dissertação (Mestrado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

The graded response model is an Item Response Theory (IRT) model where items admit a polytomous response widely known in the literature. The motivation for this study comes from a data set which belongs to the RN4CAST project related to the burnout syndrome. This psychological syndrome has a multidimensional configuration of emotional exhaustion, depersonalization and personal accomplishment. The burnout measurements are obtained from the Maslach Burnout Inventory, a 22-item questionnaire to be answered on a 7-point Likert scale. Our proposal in this work is a study of the unidimensional and multiunidimensional graded response model, under the frequentist and Bayesian approaches, motivated by the data from the *burnout* syndrome, and for that, a simulation study is carried out to verify the behavior of empirical form of the model. Models are fitted with the marginal maximum likelihood and the Monte Carlo Markov chain via Gibbs sampler methods. The simulation study shows that, in general, the Monte Carlo Markov chain via Gibbs sampler method produces good results in the estimation of the item parameters and latent trait of the unidimensional gradual response model. The results of the multiunidimensional gradual response model study yield good results as the sample size and test size increase.

Key-words: Item response theory, graded response model, Bayesian inference, classical inference, burnout syndrome .

LISTA DE ILUSTRAÇÕES

Figura 1 –	Frequências em termos absolutos das respostas das enfermeiras nos itens 1, 2, 4, 5, 12 e 15.	13
Figura 2 –	Quatro curvas características do item (CCI) considerando diferentes valores do parâmetro de dificuldade do modelo logístico de 1 parâmetro.	17
Figura 3 –	Quatro curvas características do item (CCI) considerando diferentes valores do parâmetro de discriminação e mesmo valor do parâmetro de dificuldade ($b_i = 1,0$) do modelo logístico de 2 parâmetros.	18
Figura 4 –	Três curvas características do item (CCI) considerando diferentes valores do parâmetro de adivinhação, e com parâmetros de discriminação e de dificuldade fixos ($a_i = 1,0$ e $b_i = 1,0$) do modelo logístico de 3 parâmetros.	19
Figura 5 –	Três curvas características do item (CCI) dos itens 3, 11 e 38 apresentados em Reckase (2009).	21
Figura 6 –	Curvas características das categorias acumuladas para um item ajustado pelo modelo de resposta gradual com $a_i = 1,0$, $b_{i1} = -2,0$, $b_{i2} = -1,0$, $b_{i3} = 0,5$, $b_{i4} = 1,0$ e $b_{i5} = 1,8$	24
Figura 7 –	Curvas características das categorias para um item ajustado pelo modelo de resposta gradual com $a_i = 1,0$, $b_{i1} = -2,0$, $b_{i2} = -1,0$, $b_{i3} = 0,5$, $b_{i4} = 1,0$ e $b_{i5} = 1,8$	25
Figura 8 –	Curvas características das categorias do item 2 - No fim do dia de trabalho, sinto-me exausto - relacionado a exaustão emocional da enfermeira. Neste item, $\hat{a} = 2,86$, $\hat{b}_1 = -1,61$, $\hat{b}_2 = -0,87$, $\hat{b}_3 = -0,49$, $\hat{b}_4 = 0,20$, $\hat{b}_5 = 0,48$ e $\hat{b}_6 = 1,68$	49
Figura 9 –	Curvas características das categorias do item 6 - Trabalhar com pessoas o dia todo é, de fato, um esforço para mim - relacionado a exaustão emocional da enfermeira. Neste item, $\hat{a} = 0,84$, $\hat{b}_1 = 0,11$, $\hat{b}_2 = 1,61$, $\hat{b}_3 = 2,33$, $\hat{b}_4 = 3,73$, $\hat{b}_5 = 5,09$ e $\hat{b}_6 = 6,20$	50
Figura 10 –	Curvas características das categorias do item 9 - Sinto que estou a influenciar positivamente a vida de outras pessoas com o meu trabalho - relacionado a realização pessoal reduzida da enfermeira. Neste item, $\hat{a} = 1,18$, $\hat{b}_1 = -1,29$, $\hat{b}_2 = 0,17$, $\hat{b}_3 = 0,75$, $\hat{b}_4 = 1,57$, $\hat{b}_5 = 2,11$ e $\hat{b}_6 = 2,71$	51

Figura 11 – Curvas características das categorias do item 22 - Sinto que alguns pacientes me culpam por alguns dos meus problemas - relacionado a despersonalização da enfermeira. Neste item, $\hat{a} = 0,78$, $\hat{b}_1 = -0,56$, $\hat{b}_2 = 1,31$, $\hat{b}_3 = 1,85$, $\hat{b}_4 = 2,97$, $\hat{b}_5 = 3,92$ e $\hat{b}_6 = 5,42$	52
Figura 12 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a exaustão emocional do modelo de resposta gradual unidimensional.	53
Figura 13 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a despersonalização do modelo de resposta gradual unidimensional.	53
Figura 14 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a realização pessoal reduzida do modelo de resposta gradual unidimensional.	54
Figura 15 – Curvas características das categorias dos itens 2 e 6, curvas de informação do item e função de informação do teste dos itens relacionados à exaustão emocional das enfermeiras de acordo com a estimação dos parâmetros do modelo de resposta gradual via método da máxima verossimilhança marginal.	56
Figura 16 – Curvas características das categorias dos itens 5 e 10, curvas de informação do item e função de informação do teste dos itens relacionados à despersonalização das enfermeiras de acordo com a estimação dos parâmetros do modelo de resposta gradual via método da máxima verossimilhança marginal.	57
Figura 17 – Curvas características das categorias dos itens 4 e 19, curvas de informação do item e função de informação do teste dos itens relacionados à realização pessoal reduzida das enfermeiras de acordo com a estimação dos parâmetros do modelo de resposta gradual via método da máxima verossimilhança marginal.	58
Figura 18 – Curvas características das categorias do item 2 - No fim do dia do trabalho, sinto-me exausto - relacionado à exaustão emocional da enfermeira. Neste item, $\hat{a} = 2,83$, $\hat{b}_1 = -1,67$, $\hat{b}_2 = -0,84$, $\hat{b}_3 = -0,40$, $\hat{b}_4 = 0,24$, $\hat{b}_5 = 0,50$ e $\hat{b}_6 = 1,53$	66
Figura 19 – Curvas características das categorias do item 5 - Sinto que trato alguns pacientes como se fossem objetos impessoais - relacionado à despersonalização da enfermeira. Neste item, $\hat{a} = 2,60$, $\hat{b}_1 = -0,51$, $\hat{b}_2 = 0,26$, $\hat{b}_3 = 0,60$, $\hat{b}_4 = 1,07$, $\hat{b}_5 = 1,42$ e $\hat{b}_6 = 2,20$	67
Figura 20 – Curvas características das categorias do item 4 - Consigo compreender facilmente como os meus pacientes se sentem acerca das coisas - relacionado à realização pessoal reduzida da enfermeira. Neste item, $\hat{a} = 0,87$, $\hat{b}_1 = 0,04$, $\hat{b}_2 = 1,56$, $\hat{b}_3 = 2,45$, $\hat{b}_4 = 3,41$, $\hat{b}_5 = 4,59$ e $\hat{b}_6 = 5,62$	68
Figura 21 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a exaustão emocional do modelo de resposta gradual multiunidimensional.	69

Figura 22 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a despersonalização do modelo de resposta gradual multiunidimensional.	69
Figura 23 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a realização pessoal reduzida do modelo de resposta gradual multiunidimensional.	70

LISTA DE ALGORITMOS

Algoritmo 1 – Metropolis-Hastings	84
Algoritmo 2 – Amostrador de Gibbs	86

LISTA DE CÓDIGOS-FONTE

Código-fonte 1 – Código JAGS para o modelo de resposta gradual unidimensional . . .	87
Código-fonte 2 – Código JAGS para o modelo de resposta gradual multiunidimensional	88

LISTA DE TABELAS

Tabela 1 – Versão em português dos itens do inventário de <i>burnout</i> de Maslach.	9
Tabela 2 – Estatísticas descritivas de variáveis relacionadas às enfermeiras.	10
Tabela 3 – Frequências (em percentagem) das categorias de respostas por item.	12
Tabela 4 – Parâmetros de itens para 3 itens (3, 11 e 38) de teste com pontuação dicotômica selecionados dentre os 50 itens apresentados em Reckase (2009).	20
Tabela 5 – Parâmetros dos itens usados na simulação dos dados na abordagem unidimensional.	40
Tabela 6 – Frequências (em percentagem) de resposta da população de enfermeiras da Bélgica para as categorias de respostas por item.	45
Tabela 7 – Frequências (em percentagem) de resposta da amostra de 498 enfermeiras da Bélgica para as categorias de respostas por item.	46
Tabela 8 – Estimativas dos parâmetros dos itens do modelo de resposta gradual unidimensional nos dados do MBI via amostrador Gibbs.	48
Tabela 9 – Estimativas dos parâmetros dos itens do modelo de resposta gradual unidimensional nos dados do MBI via método da máxima verossimilhança marginal.	55
Tabela 10 – Parâmetros dos itens usados na simulação dos dados na abordagem multiunidimensional.	60
Tabela 11 – Estimativas dos parâmetros dos itens do modelo de resposta gradual multiunidimensional nos dados do MBI via amostrador Gibbs.	65
Tabela 12 – REQM médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) sob diferentes métodos de estimação, tamanhos amostrais e tamanhos do teste	92
Tabela 13 – Viés médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) sob diferentes métodos de estimação, tamanhos amostrais e tamanhos do teste	93
Tabela 14 – Correlação média das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) sob diferentes métodos de estimação, tamanhos amostrais e tamanhos do teste	94

Tabela 15 – REQM médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subtteste (S1, S2 e S3) do modelo de resposta gradual multiundimensional estimado via amostrador de Gibbs, considerando associação fraca entre os traços latentes (ρ_{fr}) e diferentes tamanhos amostrais e tamanhos do teste	95
Tabela 16 – Viés médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subtteste (S1, S2 e S3) do modelo de resposta gradual multiundimensional estimado via amostrador de Gibbs, considerando associação fraca entre os traços latentes (ρ_{fr}) e diferentes tamanhos amostrais e tamanhos do teste	96
Tabela 17 – Correlação média das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subtteste (S1, S2 e S3) do modelo de resposta gradual multiundimensional estimado via amostrador de Gibbs, considerando associação fraca entre os traços latentes (ρ_{fr}) e diferentes tamanhos amostrais e tamanhos do teste	97
Tabela 18 – REQM médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subtteste (S1, S2 e S3) do modelo de resposta gradual multiundimensional estimado via amostrador de Gibbs, considerando associação média entre os traços latentes (ρ_{me}) e diferentes tamanhos amostrais e tamanhos do teste	98
Tabela 19 – Viés médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subtteste (S1, S2 e S3) do modelo de resposta gradual multiundimensional estimado via amostrador de Gibbs, considerando associação média entre os traços latentes (ρ_{me}) e diferentes tamanhos amostrais e tamanhos do teste	99
Tabela 20 – Correlação média das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subtteste (S1, S2 e S3) do modelo de resposta gradual multiundimensional estimado via amostrador de Gibbs, considerando associação média entre os traços latentes (ρ_{me}) e diferentes tamanhos amostrais e tamanhos do teste	100
Tabela 21 – REQM médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subtteste (S1, S2 e S3) do modelo de resposta gradual multiundimensional estimado via amostrador de Gibbs, considerando associação forte entre os traços latentes (ρ_{fo}) e diferentes tamanhos amostrais e tamanhos do teste	101

Tabela 22 – Viés médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação forte entre os traços latentes (ρ_{f_0}) e diferentes tamanhos amostrais e tamanhos do teste	102
Tabela 23 – Correlação média das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação forte entre os traços latentes (ρ_{f_0}) e diferentes tamanhos amostrais e tamanhos do teste	103
Tabela 24 – Correlação média entre os traços latentes estimados do modelo de resposta gradual multiunidimensional nos diferentes tamanhos amostrais (75, 150, 300, 500 e 750), tamanhos de teste (8, 12, 16 e 22) e níveis de associação (fraca, média e forte).	104

LISTA DE ABREVIATURAS E SIGLAS

BUGS *Bayesian inference using Gibbs sampling*

CCI curva característica do item

DP despersonalização

EE exaustão emocional

JAGS Just another Gibbs sampler

MBI inventário de *burnout* de Maslach

MCMC Monte Carlo via cadeias de Markov

MVM máxima verossimilhança marginal

PA realização pessoal reduzida

REQM . . . raiz quadrada do erro quadrático médio

RN4CAST *Registered Nurse Forecasting*

TRI Teoria de Resposta ao Item

SUMÁRIO

1	INTRODUÇÃO	1
2	MOTIVAÇÃO	7
2.1	Descrição do projeto RN4CAST	7
2.2	A Síndrome de <i>Burnout</i>	8
2.3	Análises descritivas	9
3	TEORIA DE RESPOSTA AO ITEM	15
3.1	Modelos para itens dicotômicos	16
3.1.1	<i>Modelo logístico de 1 parâmetro</i>	16
3.1.2	<i>Modelo logístico de 2 parâmetros</i>	18
3.1.3	<i>Modelo logístico de 3 parâmetros</i>	19
3.2	Modelos para itens politômicos	21
3.2.1	<i>Modelo de resposta gradual</i>	21
3.2.2	<i>Modelos Multidimensionais</i>	25
3.2.2.1	<i>Modelo multiunidimensional para dados dicotômicos</i>	26
3.2.2.2	<i>Modelo multiunidimensional para dados politômicos - o modelo de resposta gradual</i>	27
4	MÉTODOS DE ESTIMAÇÃO	29
4.1	Estimação Clássica	29
4.1.1	<i>Estimação dos parâmetros dos itens</i>	30
4.1.2	<i>Estimação dos traços latentes</i>	33
4.2	Estimação Bayesiana	34
4.2.1	<i>Implementação do MCMC</i>	36
5	RESULTADOS	39
5.1	Abordagem unidimensional	39
5.1.1	<i>Um estudo de simulação</i>	39
5.1.1.1	<i>Simulação dos dados</i>	39
5.1.1.2	<i>Estimação dos parâmetros</i>	40
5.1.2	<i>Uma aplicação aos dados de enfermeiras</i>	43
5.1.2.1	<i>Amostragem dos dados</i>	43
5.1.2.2	<i>Modelagem dos dados</i>	44

5.2	Abordagem multiunidimensional	59
5.2.1	<i>Um estudo de simulação</i>	59
5.2.1.1	<i>Simulação dos dados</i>	59
5.2.1.2	<i>Estimação dos parâmetros</i>	60
5.2.2	<i>Uma aplicação aos dados de enfermeiras</i>	63
5.2.2.1	<i>Amostragem dos dados</i>	63
5.2.2.2	<i>Modelagem dos dados</i>	63
6	DISCUSSÃO	71
	REFERÊNCIAS	73
APÊNDICE A	MÉTODO DE MONTE CARLO VIA CADEIAS DE MARKOV	83
A.1	O algoritmo Metropolis-Hastings	84
A.2	O amostrador de Gibbs	84
APÊNDICE B	CÓDIGOS	87
APÊNDICE C	RESULTADOS DO ESTUDO DE SIMULAÇÃO	91

INTRODUÇÃO

A Teoria de Resposta ao Item (TRI) tem como objetivo propor modelos para traços latentes. Traços latentes são características não-observáveis de um indivíduo, ou seja, que não podem ser observadas diretamente. A ideia é estudar a relação entre variáveis observáveis, como por exemplo as características de itens de um teste ou questionário, e o traço latente dos indivíduos. A partir de uma resposta dada pelo indivíduo, pode-se estimar o traço latente através de relações hipotéticas entre a resposta e o nível do traço latente. A TRI utiliza algumas equações matemáticas para expressar essas relações.

O desenvolvimento de técnicas, pesquisas e aplicações de modelos da TRI esteve relacionado principalmente com a necessidade de resolver problemas estatísticos e práticos na área educacional, ou seja, na mensuração de aptidão e habilidade. Apenas recentemente, a aplicação de modelos da TRI relacionados ao campo da personalidade, psicopatologia, resultados relatados pelo paciente e qualidade de vida relacionado à saúde começou a se proliferar entre os pesquisadores. Assim como no campo educacional, estes tipos de aplicações não educacionais ganharam popularidade com a finalidade de resolver problemas práticos e técnicos de mensuramento do traço latente (REISE; REVICKI, 2014).

Os primeiros modelos da TRI foram desenvolvidos na década de 50. Estes modelos tratavam apenas de respostas dicotômicas. Inicialmente considerou-se o modelo na forma da função ogiva normal, e depois utilizou-se a forma matemática da função logística, que é mais conveniente e amplamente utilizada atualmente (ANDRADE; TAVARES; VALLE, 2000). Em 1952, Frederic Lord foi o primeiro a elaborar, além de um modelo teórico, métodos para estimação dos parâmetros dos itens (LORD, 1952b; LORD, 1952a; LORD, 1953). Este modelo é unidimensional com 2 parâmetros do item, que são o parâmetro de discriminação e o parâmetro de dificuldade. Após aplicações, Lord sentiu a necessidade de incorporar um parâmetro de acerto ao acaso, e então surgiu o modelo de 3 parâmetros. Ambos modelos criados por Lord utilizavam a função ogiva normal. Rasch (1960) propôs o modelo de 1 parâmetro, que também utilizava a

função ogiva normal. Posteriormente, para os modelos de 2 e 3 parâmetros, e para o modelo de 1 parâmetro, [Birnbaum \(1968\)](#) e [Wright \(1968\)](#), respectivamente, propuseram substituir a função ogiva normal pela função logística, que é uma função explícita dos parâmetros do item e do traço latente e não envolve integração.

A partir do final dos anos 60, foram desenvolvidos novos modelos para dados de itens politômicos, em que avalia-se qual a resposta dada pelo indivíduo, e não somente se respondeu corretamente ou não. O modelo de resposta gradual foi desenvolvido por [Samejima \(1969\)](#) com a proposta de analisar itens com categorias ordenadas, como dados em uma escala *Likert* ([LIKERT, 1932](#)). [Bock \(1972\)](#) propôs o modelo de resposta nominal, com o objetivo de analisar itens com categorias não ordenadas, como itens de múltipla escolha. O modelo de escala gradual foi proposto por [Andrich \(1978\)](#), e é um caso particular do modelo de resposta gradual de [Samejima \(1969\)](#), ou seja, é um modelo adequado para itens de categorias de resposta ordenadas, porém supõe-se que os escores das categorias são igualmente espaçados. O modelo de crédito parcial ([MASTERS, 1982](#)) é uma extensão do modelo de [Rasch \(1960\)](#) para itens politômicos com categorias ordenadas. Já o modelo de crédito parcial generalizado proposto por [Muraki \(1992\)](#) baseou-se no modelo de crédito parcial de [Masters \(1982\)](#) desconsiderando a hipótese de poder de discriminação uniforme para todos os itens. O modelo de resposta gradual de [Samejima \(1969\)](#) é um dos modelos para itens politômicos mais utilizados em diversas áreas, como educação, psicologia e *marketing*. Existe uma gama de trabalhos encontrados na literatura que utilizam o modelo de resposta gradual, como, por exemplo, [Koch \(1983\)](#), [Reise e Yu \(1990\)](#), [Ankenmann e Stone \(1992\)](#), [Baker \(1992\)](#), [Cohen, Kim e Baker \(1993\)](#), [Reise, Widaman e Pugh \(1993\)](#), [DeMars \(2002\)](#), [Ferrando \(2009\)](#), [LaHuis, Clark e O'Brien \(2011\)](#), [Kieftenbeld e Natesan \(2012\)](#), [Kuo e Sheng \(2016\)](#) e [Silva, Bazán e Huggins-Manley \(2017\)](#).

Nos modelos da teoria de resposta ao item convencionais existem duas suposições principais: unidimensionalidade e independência local. Na primeira assume-se que o conjunto de itens do teste mensura um único traço latente. Visto que a performance de um indivíduo no teste pode ser afetada por diversos fatores cognitivos e/ou de personalidade, considera-se que existe um fator “dominante” que afeta a performance do indivíduo. A segunda suposição, de independência local, considera que dado o traço latente do indivíduo, as respostas do indivíduo aos itens são estatisticamente independentes. Se a suposição de unidimensionalidade não é válida, isto é, o conjunto de itens do teste mensura múltiplos traços latentes, o modelo é denominado multidimensional. Modelos multidimensionais envolvem grande complexidade comparados aos modelos unidimensionais.

Utilizar modelos unidimensionais para ajustar os dados de respostas de testes que mensuram múltiplos traços latentes pode ser problemático. Segundo [Reckase \(1997\)](#), a metodologia dos modelos da teoria de resposta ao item multidimensionais surgiu para ajudar a entender as habilidades necessárias para responder corretamente a itens de um teste, as características de um indivíduo que afetam a probabilidade de responder um item e a complexidade por trás das

equações dos testes. Na literatura podem ser encontrados, por exemplo, os seguintes trabalhos relacionados aos modelos da teoria de resposta ao item multidimensionais: [Dragow e Parsons \(1983\)](#), [Reckase \(1985\)](#), [Kelderman e Rijkes \(1994\)](#), [Ackerman, Gierl e Walker \(2003\)](#), [Walker e Beretvas \(2003\)](#), [Kacmar *et al.* \(2006\)](#), [Zhang e Stone \(2008\)](#), [Li, Jiao e Lissitz \(2012\)](#), [Kose e Demirtasli \(2012\)](#), [Kose e Demirtasli \(2012\)](#), [Wetzel e Hell \(2014\)](#), [Svetina *et al.* \(2017\)](#).

Alguns testes ou questionários consistem em vários subtestes ou subquestionários, em que cada um concentra-se em uma habilidade específica, e os itens em cada subteste são projetados para medir uma habilidade em comum ([SHENG; WIKLE, 2007](#)). Este é um caso especial dos modelos multidimensionais, e na literatura são chamados de modelos da teoria de resposta ao item multidimensionais com estrutura simples ou entre itens, ou ainda, utiliza-se o termo modelos multiunidimensionais, sendo este mais intuitivo. Trabalhos sobre os modelos da teoria de resposta ao item multiunidimensionais podem ser encontrados em [Oshima, Raju e Flowers \(1997\)](#), [Torre e Patz \(2005\)](#), [Martelli \(2014\)](#) e [Kuo e Sheng \(2016\)](#).

Outras estruturas multidimensionais podem ser consideradas no modelo. A estrutura bi-fatorial, introduzida por [Holzinger e Swineford \(1937\)](#), pode ser utilizada quando um único item mensura mais do que um traço latente. Neste caso, assume-se que existe uma habilidade geral e habilidades específicas que afetam as respostas dos indivíduos a cada item, e considera-se que os fatores são ortogonais. Estudos que levam em consideração a estrutura bi-fatorial nos modelos da teoria de resposta ao item multidimensionais podem ser vistos em [Reise, Morizot e Hays \(2007\)](#), [Reise \(2012\)](#) e [Toland *et al.* \(2017\)](#). A estrutura aditiva é similar a estrutura bi-fatorial, porém leva-se em consideração que os fatores podem ser correlacionados. Trabalhos com este tipo de estrutura podem ser vistos em [Sheng e Wikle \(2008\)](#) e [Martelli, Matteucci e Mignani \(2016\)](#). Já na estrutura hierárquica, assume-se que itens em um mesmo subteste mensuram uma habilidade específica, que é influenciada por uma habilidade geral, em que diferentes estruturas hierárquicas podem ser utilizadas a depender da relação entre as habilidades específicas e geral. Estudos de modelos da teoria de resposta ao item multidimensionais com estrutura hierárquica pode ser visto em [Sheng e Wikle \(2008\)](#) e [Zhou \(2019\)](#).

Uma das etapas mais importantes da TRI é a estimação dos parâmetros. Isso foi um grande desafio no passado, mas nas últimas décadas numerosas técnicas eficientes de estimação foram desenvolvidas. Os parâmetros são relacionados aos itens e aos indivíduos, e existem três diferentes cenários possíveis: quando se conhece as habilidades dos indivíduos e tem-se o interesse em estimar os parâmetros dos itens, ou quando se conhece os parâmetros dos itens e tem-se o interesse em estimar as habilidades dos indivíduos, ou ainda, se os parâmetros dos itens e as habilidades dos indivíduos são desconhecidas e tem-se o interesse em estimar ambos, sendo esta a situação mais comum. Na TRI, o processo de estimação dos parâmetros dos itens é chamado de calibração ([ANDRADE; TAVARES; VALLE, 2000](#)). A estimação, nas três situações citadas anteriormente, pode ser feita por duas abordagens: a clássica e a Bayesiana. Na inferência clássica, geralmente os parâmetros são estimados pelo método da máxima verossimilhança.

Na inferência Bayesiana pode ser utilizado uma extensão Bayesiana do método da máxima verossimilhança, porém, avanços computacionais permitiram a utilização de métodos iterativos que utilizam Monte Carlo via Cadeia de Markov (MCMC) para a estimação dos parâmetros.

Na inferência clássica os métodos da máxima verossimilhança conjunta e da máxima verossimilhança condicional foram os primeiros métodos propostos para estimação dos parâmetros de modelos da teoria de resposta ao item. Na situação em que deseja-se estimar tantos os parâmetros do item quanto os parâmetros dos indivíduos, a estimação pelo método da máxima verossimilhança conjunta considera a estimação simultânea de todos os parâmetros. Isso leva a uma enorme exigência computacional, uma vez que é grande o número de parâmetros a serem estimados ao mesmo tempo, envolvendo inversão de matrizes desta ordem. Este método pode produzir estimadores inconsistentes para os parâmetros dos itens (ANDERSEN, 1973). A estimação pelo método da máxima verossimilhança condicional envolve a estimação dos parâmetros dos itens considerando inicialmente as habilidades conhecidas, e após a obtenção das estimativas dos parâmetros do item, a estimação dos parâmetros das habilidades é realizada considerando conhecidos os parâmetros dos itens. Andersen (1973) notou para o modelo Rasch que na presença de um número muito grande de indivíduos (ou itens), existe um problema de falta de consistência dos estimadores dos parâmetros dos itens (ou habilidades).

Por conta dos problemas dos métodos citados acima, Bock e Lieberman (1970) desenvolveram um método de estimação em duas etapas, considerando o método da máxima verossimilhança marginal: baseia-se na existência de uma distribuição associada à habilidade dos indivíduos, possibilitando a estimação dos parâmetros dos itens considerando a verossimilhança marginal, e posteriormente, a estimação das habilidades é realizada individualmente por máxima verossimilhança, por exemplo. Ainda assim, este método pode envolver inversão de matrizes de ordem grande, o que implica muita exigência computacional. Para contornar este problema, Bock e Aitkin (1981) modificaram o método proposto anteriormente em 1970, adicionando a suposição de independência entre os itens, e propondo que as estimativas de máxima verossimilhança sejam obtidas com a aplicação do algoritmo EM, proposto por Dempster, Laird e Rubin (1977). Dragow (1989), Ogasawara (2001) e Casabianca e Lewis (2015) são exemplos de trabalhos na literatura que utilizam e estudam o método da máxima verossimilhança marginal na TRI.

Na literatura da TRI, a estimação dos parâmetros utilizando inferência Bayesiana inclui uma extensão de métodos de máxima verossimilhança. A ideia de utilizar métodos Bayesianos na Teoria de Resposta ao Item é tentar aprimorar as acurácias das estimativas dos parâmetros, além de contornar os problemas existentes na estimação clássica, comentados anteriormente. Inicialmente, Swaminathan e Gifford (1982), Swaminathan e Gifford (1985), Swaminathan e Gifford (1986) propuseram uma estimação Bayesiana conjunta dos parâmetros dos modelos Rasch, logístico de 2 parâmetros e logístico de 3 parâmetros, respectivamente, como uma extensão do método da máxima verossimilhança clássico, em que a estimação dos parâmetros dos itens e dos indivíduos é realizada simultaneamente. Mislevy (1986), Tsutakawa e Lin (1986)

propuseram, em contrapartida, métodos de estimação Bayesiana que são extensões do método da máxima verossimilhança marginal de [Bock e Aitkin \(1981\)](#).

Nas últimas duas décadas, com o avanço computacional, os métodos Bayesianos utilizando MCMC vem ganhando um crescente interesse na literatura da TRI. Os métodos MCMC são uma classe de algoritmos utilizados para obter amostras de uma distribuição de probabilidade (por exemplo, a distribuição a posteriori) com base na construção de uma cadeia de Markov. O algoritmo MCMC que é comumente aplicado em modelos da TRI é o amostrador de Gibbs ([GEMAN; GEMAN, 1984](#)). [Fox \(2010\)](#) discute métodos Bayesianos para analisar dados de resposta de itens e, além disso, trabalhos utilizando algoritmos MCMC na estimação de parâmetros de modelos da teoria de resposta ao item são encontrados na literatura, como, por exemplo, [Patz e Junker \(1999\)](#), [Kim \(2001\)](#), [Sahu \(2002\)](#), [Kim e Bolt \(2007\)](#), [Kieftenbeld e Natesan \(2012\)](#), [Natesan *et al.* \(2016\)](#), [Chang \(2017\)](#) e [Marcoulides \(2018\)](#).

A motivação deste trabalho partiu de um conjunto de dados relacionado com a síndrome de *burnout* em enfermeiras. *Burnout* é uma síndrome psicológica que leva a um cansaço físico e emocional, perda de motivação para o trabalho e aparecimento de sentimentos de inadequação e de fracasso. A síndrome tem uma configuração multidimensional de exaustão emocional, despersonalização e realização pessoal reduzida. Os dados motivacionais são respostas de enfermeiras a um questionário de 22 itens, que são estados de sentimentos relatados no trabalho, para serem respondidos por meio de uma escala Likert de sete pontos. Neste questionário, cada dimensão de *burnout* é tratada em conjunto de itens, que são os subquestionários.

A configuração do questionário relacionado com a síndrome de *burnout* em enfermeiras motivou a ideia de comparar, sob diferentes cenários, a performance do modelo de resposta gradual com a configuração unidimensional, em que considera-se cada dimensão da síndrome de *burnout* separadamente e o modelo de resposta gradual com a configuração multiunidimensional, sendo as dimensões da síndrome de *burnout* as dimensões do modelo multiunidimensional.

Este trabalho compara a performance de estimação dos parâmetros do modelo de resposta gradual unidimensional e multiunidimensional em diferentes cenários propostos no estudo de simulação. Além disso, é discutida uma aplicação desse modelo ao conjunto de dados referente à síndrome de *burnout* em enfermeiras. Neste trabalho, considera-se as abordagens clássica e Bayesiana para a estimação dos parâmetros utilizando os métodos da máxima verossimilhança marginal e MCMC via amostrador de Gibbs.

O Capítulo 2 apresenta a motivação do trabalho, originada por um conjunto de dados relacionado à síndrome de *burnout* em enfermeiras provenientes do projeto Registered Nurse Forecasting. O Capítulo 3 apresenta modelos da Teoria de Resposta ao Item dicotômicos, bem como o modelo de resposta gradual. Os métodos de estimação dos parâmetros dos modelos de resposta gradual são descritos no Capítulo 4, em que considera-se as abordagens clássica e Bayesiana. Os resultados do estudo de simulação e da aplicação aos dados de enfermeiras são observados no Capítulo 5. Por fim, o Capítulo 6 apresenta uma discussão e propostas futuras.

MOTIVAÇÃO

2.1 Descrição do projeto RN4CAST

O projeto *Registered Nurse Forecasting* (RN4CAST), descrito em [Sermeus et al. \(2011\)](#), é um dos maiores estudos sobre a força de trabalho de enfermagem que foi realizado na Europa. A ideia do projeto é contribuir para a precisão dos modelos de previsão e gerar novas abordagens para uma gestão mais eficaz dos recursos de enfermagem.

Segundo [Sermeus et al. \(2011\)](#), a profissão na área de saúde com maior número de profissionais é a enfermagem. Além disso, o trabalho realizado pelos enfermeiros tem grande impacto nos resultados e na segurança dos pacientes e nos custos envolvidos. Por isso, o estudo e o planejamento relacionados a enfermeiros podem dar suporte a ações em nível político e público. Análises recentes dos recursos humanos para saúde mostram que todos os países podem estimular benefícios na área da saúde através de mais investimentos estratégicos e administração do trabalho dos enfermeiros, ou seja, estratégias efetivas no trabalho da enfermagem podem melhorar o desempenho do sistema de saúde.

O projeto RN4CAST foi realizado entre 2009 e 2011 e envolveu 33.731 enfermeiros registrados em 2.169 unidades de enfermagem em 486 hospitais em 12 países da Europa (Bélgica, Inglaterra, Finlândia, Alemanha, Grécia, Irlanda, Noruega, Polônia, Espanha, Suécia, Suíça e Holanda). A amostragem multinível foi realizada de tal forma que, em cada um dos 12 países, um mínimo de 30 hospitais gerais (não especializados) foram selecionados aleatoriamente, com exceção da Irlanda e da Noruega onde os hospitais selecionados representavam todas as instituições relevantes. Pelo menos duas unidades gerais de enfermagem médica e cirúrgica para cada hospital foram selecionadas aleatoriamente, porque a ligação entre a carga de trabalho e o ambiente de trabalho dos enfermeiros com a segurança do paciente e os desfechos clínicos é melhor documentada nestes tipos de unidades de enfermagem. Todas as unidades de enfermagem que foram selecionadas aleatoriamente nos hospitais participantes concordaram em participar do

estudo e todos os enfermeiros envolvidos em atividades diretas de assistência ao paciente foram convidados para o estudo (SERMEUS *et al.*, 2011).

As pesquisas foram traduzidas para 11 línguas a partir da sua versão original em inglês americano, assegurando simultaneamente a sua relevância para a prática de enfermagem e contextos de cuidados de saúde de 12 países. Não foram permitidas alterações ao modelo do questionário original, de modo que as perguntas foram redigidas de forma semelhante em todos os países. A qualidade das traduções foi avaliada por um júri de 7 a 11 especialistas bilíngues (incluindo enfermeiros) para obter índices de validade de conteúdo, de modo a permitir a comparabilidade dos conceitos entre países.

Fatores como características dos ambientes de trabalho e as qualificações da força de trabalho do enfermeiro podem afetar o bem-estar dos enfermeiros, causando um esgotamento, denominado síndrome de *burnout*. A síndrome de *burnout* tem efeitos importantes na satisfação do trabalho dos enfermeiros, na alta rotatividade de funcionários (um empregado é admitido e outro desligado de maneira sucessiva) e na satisfação do paciente (SERMEUS *et al.*, 2011). Assim, a sua investigação, prevenção e tratamento podem ser estratégias para melhorar o desempenho do sistema de saúde.

2.2 A Síndrome de *Burnout*

O termo *burnout* foi inicialmente introduzido na literatura científica por [Freudenberger \(1974\)](#). Ele utilizou esse termo para descrever uma síndrome que parecia ser frequente nos profissionais de saúde como consequência da tensão emocional e do esforço elevado causados pelo contato com pessoas doentes e vulneráveis ([RIBAS, 2011](#)).

Segundo [Ribas \(2011\)](#), o conceito de *burnout* foi evoluindo devido a contribuição de diversos autores, porém a definição mais consolidada atualmente é “um cansaço físico e emocional que leva a uma perda de motivação para o trabalho, que pode evoluir até ao aparecimento de sentimentos de inadequação e de fracasso” ([MASLACH; JACKSON; LEITER, 1996](#)).

Segundo [Maslach, Jackson e Leiter \(1996\)](#), *burnout* é uma síndrome psicológica que tem uma configuração multidimensional de **exaustão emocional**, **despersonalização** e **realização pessoal reduzida**. A exaustão emocional ocorre quando a pessoa percebe nela mesma a impressão de que não dispõe de recursos suficientes para dar aos outros, levando a sentimentos de excessivo stress emocional. A despersonalização corresponde ao desenvolvimento por parte do profissional de saúde de atitudes negativas e insensíveis em relação as pessoas com as quais trabalha, tratando-as como objetos. Já a realização pessoal reduzida, normalmente conduz a uma avaliação negativa e baixa de si mesmo provocando uma diminuição das expectativas sobre si próprio ([RIBAS, 2011](#)). As consequências de *burnout* são potencialmente sérias para os trabalhadores, seus clientes e as instituições nas quais interagem.

Um instrumento muito utilizado em pesquisas da síndrome de *burnout* relacionado ao trabalho é o inventário de *burnout* de Maslach (MBI) (MASLACH; JACKSON; LEITER, 1996) de 22 itens. Como a exaustão emocional, a despersonalização e a realização pessoal reduzida são variáveis latentes, isto é, não podem ser medidas diretamente, uma possibilidade é medi-las por meio de questões que possam ser respondidas utilizando, por exemplo, a escala Likert (LIKERT, 1932), um tipo de escala de resposta psicométrica que mede atitudes e comportamentos utilizando opções de resposta que variam de um extremo a outro. Especificamente neste questionário, os itens são estados de sentimentos relatados no trabalho, para ser respondido por meio de uma escala Likert de sete pontos (0 – nunca; 1 – algumas vezes por ano ou menos 2 – uma vez por mês; 3 – algumas vezes por mês; 4 – uma vez por semana; 5 – algumas vezes por semana; 6 – todos os dias). Os itens apresentam-se divididos em três sub escalas: exaustão emocional (EE), despersonalização (DP) e realização pessoal reduzida (PA). A EE consiste em nove itens (1; 2; 3; 6; 8; 13; 14; 16; 20), enquanto a DP corresponde a cinco itens (5; 10; 11; 15; 22) e a PA a oito (4; 7; 9; 12; 17; 18; 19; 21), que estão apresentados na Tabela 1. A síndrome de *burnout* é indicada por maiores pontuações de EE e DP e menores pontuações em PA.

Tabela 1 – Versão em português dos itens do inventário de *burnout* de Maslach.

Descrição dos itens	Frequência
1. Sinto-me vazio emocionalmente, por causa do meu trabalho.	0 1 2 3 4 5 6
2. No fim do dia de trabalho, sinto-me exausto.	0 1 2 3 4 5 6
3. Sinto-me fatigado quando acordo de manhã e tenho que enfrentar mais um dia de trabalho.	0 1 2 3 4 5 6
4. Consigo compreender facilmente como os meus pacientes se sentem acerca das coisas.	6 5 4 3 2 1 0
5. Sinto que trato alguns pacientes como se fossem objetos impessoais.	0 1 2 3 4 5 6
6. Trabalhar com pessoas o dia todo é, de fato, um esforço para mim.	0 1 2 3 4 5 6
7. Lido muito eficazmente com os problemas dos meus pacientes.	6 5 4 3 2 1 0
8. Sinto-me esgotado devido ao meu trabalho.	0 1 2 3 4 5 6
9. Sinto que estou a influenciar positivamente a vida de outras pessoas com o meu trabalho.	6 5 4 3 2 1 0
10. Tornei-me mais insensível em relação às pessoas, desde que comecei este trabalho.	0 1 2 3 4 5 6
11. Preocupo-me que este trabalho me esteja a “endurecer” emocionalmente.	0 1 2 3 4 5 6
12. Sinto-me muito enérgico.	6 5 4 3 2 1 0
13. Sinto-me muito frustrado com o meu trabalho.	0 1 2 3 4 5 6
14. Sinto que estou a trabalhar demasiado no meu emprego.	0 1 2 3 4 5 6
15. De fato, não me interessa o que acontece a alguns pacientes.	0 1 2 3 4 5 6
16. Trabalhar diretamente com pessoas coloca-me sob demasiada tensão.	0 1 2 3 4 5 6
17. Consigo facilmente criar uma atmosfera relaxada com os meus pacientes.	6 5 4 3 2 1 0
18. Conseguí realizar muitas coisas importantes nesta profissão.	6 5 4 3 2 1 0
19. Sinto-me entusiasmado depois de trabalhar de perto com os meus pacientes.	6 5 4 3 2 1 0
20. Sinto que estou no meu limite (“fim de linha”).	0 1 2 3 4 5 6
21. No meu trabalho, lido com os problemas emocionais com muita calma.	6 5 4 3 2 1 0
22. Sinto que os pacientes me culpam por alguns dos seus problemas	0 1 2 3 4 5 6

2.3 Análises descritivas

O projeto RN4CAST envolveu 33.731 enfermeiras, porém, aqui são apresentadas análises de 28.683 enfermeiras. As 5.048 enfermeiras que não estão sendo consideradas nesta análise

inicial são as que não responderam pelo menos um item do MBI (dado faltante). A seguir estão apresentadas estatísticas descritivas de algumas variáveis relacionadas as enfermeiras presentes no estudo RN4CAST, como por exemplo, o tempo de experiência de trabalho. Essas variáveis podem ser medidas nos níveis das enfermeiras, das unidades de enfermagem, dos hospitais e dos países, como está apresentado na Tabela 2.

A variável *experiência de trabalho* expressa os anos que as enfermeiras trabalham de forma registrada. Observa-se na Tabela 2 que, no nível das enfermeiras, o intervalo do tempo de experiência de trabalho varia de 0 até 50 anos de experiência. No nível das unidades de enfermagem, a menor média de anos de experiência também é de 0 ano e a maior é de 41 anos de experiência, enquanto no nível dos hospitais, a variação das médias diminui, sendo que a menor média de anos de experiência é de aproximadamente 4 anos e a maior é de aproximadamente 25 anos. A média geral de experiência de trabalho é de 13 anos.

A variável *ambiente de trabalho* representa um resumo médio geral do ambiente de trabalho, com base na escala de ambiente de prática do índice de trabalho de enfermagem (LAKE, 2002). Uma pontuação em escala Likert de 4 pontos é fornecida para cada item (por exemplo, “Elogio e reconhecimento por um trabalho bem feito”). A escala é: “Totalmente concordo”=4, “Concordo”=3, “Não concordo”=2 e “Totalmente não concordo”=1, de tal forma que valores elevados refletem um ambiente positivo. A média do ambiente de trabalho é cerca de 2,4, o que representa um sentimento global neutro sobre o ambiente de trabalho. Esta covariável é bastante estável entre os países, mas varia mais entre os hospitais e as unidades de enfermagem.

As unidades de enfermagem foram classificadas como cirúrgicas (cerca de 46%) ou médicas. As unidades de enfermagem registradas como cirúrgicas e médicas foram classificadas como médicas.

Tabela 2 – Estatísticas descritivas de variáveis relacionadas às enfermeiras.

	Experiência de trabalho (anos) ¹	Ambiente de trabalho ¹	Unid. de enfer. cirúrgica
País	—	2,40 (1,98;2,76)	—
Hospital	13,81 (3,76;25,19)	2,40 (1,37;3,50)	—
Unid. de enfer.	13,57 (0,00;41,00)	2,42 (1,00;4,25)	45,45%
Enfermeira	12,70 (0,00;50,00)	—	—

¹Média (e intervalo)

A Tabela 3 expõe a frequência, em porcentagem, que as enfermeiras responderam as categorias de cada item. Entre os itens de exaustão emocional, a categoria 3 (algumas vezes por mês) dos itens 1 (“Sinto-me vazio emocionalmente, por causa do meu trabalho”), 3 (“Sinto-me fatigado quando acordo de manhã e tenho que enfrentar mais um dia de trabalho”) e 14 (“Sinto

que estou a trabalhar demasiado no meu trabalho”), respectivamente foi a de maior frequência de respostas das enfermeiras, sendo 24,1% no item 1, 19,7% no item 3 e 21,4% no item 14. Além disso, nos itens 3 e 14, a categoria 5 (algumas vezes por semana) também possui uma grande frequência de resposta das enfermeiras, sendo de 19,0% e 20,7%, respectivamente, e no item 1, a categoria 1 (algumas vezes por ano ou menos) também possui grande frequência (21,4% das enfermeiras). Já nos itens 6 (“Trabalhar com pessoas o dia todo é, de fato, um esforço para mim”), 8 (“Sinto-me esgotado devido ao meu trabalho”), 16 (“Trabalhar diretamente com pessoas coloca-me sob demasiada tensão”) e 20 (“Sinto que estou no meu limite (“fim de linha”)”), as categorias 0 (nunca) e 1 (algumas vezes por ano ou menos) foram as respostas de maior frequência das enfermeiras, sendo que 66,2% das enfermeiras responderam as categoria 0 ou 1 no item 6, 53,6% no item 8, 77,1% no item 16 e 65,0% no item 20. As categorias 5 (algumas vezes por semana) e 3 (algumas vezes por mês) foram de maior frequência no item 2 (“No fim do dia de trabalho, sinto-me exausto”), sendo de 24,6% e 22,5%, respectivamente. No item 13 (“Sinto-me muito frustrado com o meu trabalho”), 22,0% das enfermeiras responderam a categoria 1 (algumas vezes por ano ou menos), sendo a categoria de maior frequência de resposta neste item.

As frequências de resposta nas categorias dos itens de despersonalização também estão apresentadas na Tabela 3. Em todos os itens relacionados a despersonalização, 5 (“Sinto que trato alguns pacientes, como se fossem objetos impessoais”), 10 (“Tornei-me mais insensível em relação às pessoas, desde que comecei este trabalho”), 11 (“Preocupo-me que este trabalho me esteja a “endurecer” emocionalmente”), 15 (“De fato, não me interessa o que acontece a alguns pacientes”) e 22 (“Sinto que os pacientes me culpam por alguns dos seus problemas”), a maior frequência de resposta das enfermeiras são das categorias 0 (nunca) ou 1 (algumas vezes por ano ou menos), sendo de 69,5% no item 5, 77,8% no item 10, 71,2% no item 11, 88,8% no item 15 e 66,2% no item 22. Porém, ainda assim, existe um número grande de enfermeiras que afirmam se sentir ou agir da forma descrita nos itens com uma frequência de pelo menos uma vez por mês.

A Tabela 3 também apresenta a frequência de resposta das categorias dos itens relacionados a realização pessoal reduzida, em que para esta dimensão, a síndrome de *burnout* é indicada por menores pontuações da escala Likert. Nos itens 4 (“Consigo compreender facilmente como os meus pacientes se sentem acerca das coisas”), 7 (“Lido muito eficazmente com os problemas dos meus pacientes”), 9 (“Sinto que estou a influenciar positivamente a vida de outras pessoas com o meu trabalho”), 17 (“Consigo facilmente criar uma atmosfera relaxada com os meus pacientes”), 18 (“Consegui realizar muitas coisas importantes nesta profissão”), 19 (“Sinto-me entusiasmado depois de trabalhar de perto com os meus pacientes”) e 21 (“No meu trabalho, lido com os problemas emocionais com muita calma”), as categorias que tiveram maior frequência de resposta das enfermeiras foram 6 (todos os dias) ou 5 (algumas vezes por semana), sendo de 81,9% no item 4, 84,7% no item 7, 68,5% no item 9, 78,4% no item 17, 69,4% no item 18, 61,6% no item 19 e 66,2% no item 21. As categorias de maior frequência de resposta das enfermeiras no item 12 (“Sinto-me muito enérgico”) foram 5 (algumas vezes por semana) e 3 (algumas vezes

por mês), em que 33,6% das enfermeiras responderam a categoria 5 e 18,8% a categoria 3.

A Figura 1 apresenta as frequências em termos absolutos das respostas das enfermeiras nos itens 1, 2, 4, 5, 12 e 15.

Tabela 3 – Frequências (em porcentagem) das categorias de respostas por item.

	Categorias						
	0	1	2	3	4	5	6
Item 1	5,6	21,4	17,9	24,1	10,4	15,4	5,2
Item 2	3,7	11,4	12,9	22,5	12,6	24,6	12,3
Item 3	9,3	16,8	14,9	19,7	12,1	19,0	8,2
Item 4	0,7	1,3	2,1	6,6	7,4	27,4	54,5
Item 5	50,4	19,1	9,5	9,2	4,9	5,3	1,6
Item 6	38,8	27,4	12,9	10,3	4,4	4,3	1,9
Item 7	1,2	1,4	1,8	5,0	5,9	30,5	54,2
Item 8	25,6	28,0	12,9	13,9	6,1	8,8	4,7
Item 9	2,0	3,0	4,4	11,7	10,4	31,8	36,7
Item 10	59,0	18,8	7,5	6,7	3,0	3,0	2,0
Item 11	46,9	24,3	9,3	8,2	3,5	4,0	3,8
Item 12	4,5	5,5	9,4	18,8	13,3	33,6	14,9
Item 13	15,1	22,0	14,3	19,0	10,6	12,8	6,2
Item 14	5,8	14,6	13,1	21,4	11,6	20,7	12,8
Item 15	73,5	15,3	4,7	3,0	1,5	1,1	0,9
Item 16	47,6	29,5	9,5	6,9	2,5	2,8	1,2
Item 17	1,8	2,3	3,2	6,8	7,5	33,7	44,7
Item 18	1,2	3,1	4,7	11,2	10,4	34,0	35,4
Item 19	2,2	3,3	5,9	13,8	13,2	34,9	26,7
Item 20	35,7	29,3	10,6	10,6	4,9	6,1	2,8
Item 21	1,1	3,7	6,0	12,9	10,1	32,9	33,3
Item 22	35,8	30,4	10,9	10,2	5,1	5,7	1,9

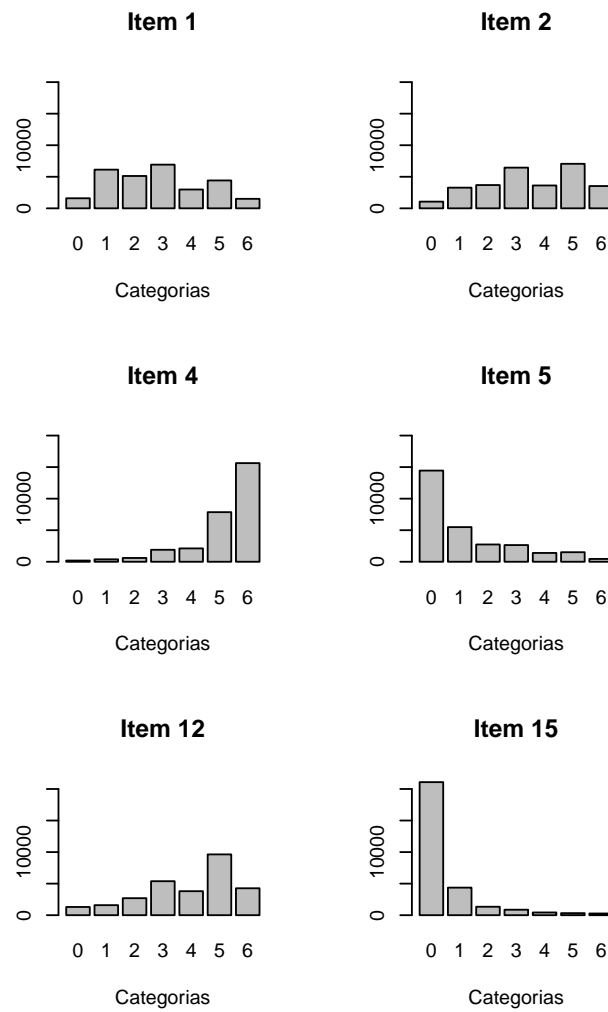


Figura 1 – Frequências em termos absolutos das respostas das enfermeiras nos itens 1, 2, 4, 5, 12 e 15.

No próximo capítulo, será revisitada a teoria de resposta ao item e seus modelos. Serão apresentados modelos para respostas dicotômicas, além do modelo de resposta gradual, para dados com resposta politômica.

TEORIA DE RESPOSTA AO ITEM

Testes educacionais ou psicológicos são instrumentos utilizados para obter amostras de comportamentos ou habilidades, que usualmente são quantificados de forma a obter escores numéricos (LORD, 1980). Na Teoria Clássica das Medidas, procedimentos e modelos populares para construir estes tipos de teste e interpretar seus escores foram amplamente utilizados por especialistas há muito tempo. Porém, a forma como os testes educacionais e psicológicos eram construídos, avaliados e usados possui muitas falhas, como, por exemplo, a estimativa da habilidade do indivíduo depender da escolha específica dos itens selecionados para um teste. A partir de 1980, especialistas progrediram no estudo de um novo sistema de medição, a Teoria de Resposta ao Item (TRI), de forma a tratar essas falhas comuns nas práticas de medição (HAMBLETON; SWAMINATHAN; ROGERS, 1991).

Segundo Hambleton, Swaminathan e Rogers (1991), a TRI baseia-se em dois princípios básicos: a performance do indivíduo em um item do teste pode ser predita ou explicada por um conjunto de fatores chamado traços latentes ou habilidades; e a relação entre a performance do indivíduo em um item e o conjunto de traços subjacentes a performance do item pode ser descrita por uma função monótona crescente chamada função característica do item ou curva característica do item (CCI). Ou seja, para inferir sobre o nível de habilidade do indivíduo é necessário saber como sua habilidade determina a resposta em um item, e essa relação é dada pela curva característica do item. Assim, a metodologia da TRI sugere uma forma de representar a relação entre a probabilidade de um indivíduo dar uma certa resposta a um item e seu traço latente (ANDRADE; TAVARES; VALLE, 2000).

Todos os modelos da TRI contêm um ou mais parâmetros que descrevem o item e um ou mais parâmetros que descrevem o indivíduo, e um dos primeiros passos é a estimação destes parâmetros. Porém, os modelos TRI incluem suposições sobre os dados, como comentado em Andrade, Tavares e Valle (2000), em que é fundamental que estejam satisfeitas:

Unidimensionalidade apenas uma habilidade é medida por um conjunto de itens em um teste.

É suficiente admitir que haja uma habilidade “dominante” que influencia o desempenho de um indivíduo no teste, visto que muitos fatores cognitivos e de personalidade podem influenciar a sua performance. Os modelos em que se assume que mais do que uma habilidade do indivíduo é levada em consideração no teste são chamados multidimensionais, e são mais complexos.

Independência local assume-se que, para uma dada habilidade, as respostas aos diferentes itens do teste são independentes. Ou seja, as habilidades especificadas no modelo são os únicos fatores que influenciam as respostas dos indivíduos aos itens do teste. Assim, para um dado indivíduo, a probabilidade de uma resposta padrão em um conjunto de itens é igual ao produto das probabilidades associadas com as respostas do indivíduo aos itens individuais.

A seguir serão apresentados modelos de resposta ao item populares para dados com resposta dicotômica.

3.1 Modelos para itens dicotômicos

Os modelos que serão apresentados nesta seção são utilizados para itens que possuem resposta dicotômica ou que foram dicotomizadas, como por exemplo itens de múltipla escolha corrigidos como certo ou errado. Os três modelos mais utilizados para este tipo de dado são os modelos logísticos de 1, 2 e 3 parâmetros, e são nomeados dessa forma por causa do número de parâmetros relacionados ao item que cada modelo possui. Os modelos para itens dicotômicos apresentados a seguir têm o objetivo de uma construção didática dos modelos de repostas politômicas, que são o foco deste trabalho.

3.1.1 Modelo logístico de 1 parâmetro

As curvas características do item para o modelo logístico de 1 parâmetro, que também é chamado de Modelo Rasch, são dadas pela equação

$$P(Y_{ij} = 1 | \theta_j) = \frac{1}{1 + e^{-D(\theta_j - b_i)}}, \quad (3.1)$$

com $i = 1, \dots, I$, $j = 1, \dots, n$, $\theta_j \in \mathbb{R}$ e $b_i \in \mathbb{R}$, em que

Y_{ij} é uma variável dicotômica que pode ser representada da seguinte maneira

$$Y_{ij} = \begin{cases} 1, & \text{se o } j\text{-ésimo indivíduo pontuar no } i\text{-ésimo item;} \\ 0, & \text{se o } j\text{-ésimo indivíduo não pontuar no } i\text{-ésimo item.} \end{cases}$$

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

$P(Y_{ij} = 1|\theta_j)$ é a probabilidade do indivíduo j com habilidade θ_j pontuar no i -ésimo item.

b_i é o parâmetro de dificuldade do item i .

D é um fator de escala constante e igual a 1 se deseja-se obter resultado da função logística. O valor 1,7 é utilizado para obter a aproximação para a distribuição normal.

I é o número total de itens no teste.

n é o número total de indivíduos.

Neste modelo é assumido que a **dificuldade do item**, b_i , é a única característica do item que influencia a performance do indivíduo. O parâmetro b_i é um parâmetro de posição, e indica a posição da CCI na escala da habilidade, e seu valor é dado pelo ponto na escala da habilidade em que a probabilidade de pontuar é igual a 0,5. Assim, quanto maior a dificuldade do item, maior é a habilidade necessária para que o indivíduo tenha 50% de probabilidade de pontuar no item. Desta forma, é desejável que um instrumento de medição contenha itens com diferentes valores para este parâmetro, contemplando toda a escala do traço latente. A Figura 2 apresenta quatro CCI's do modelo logístico de 1 parâmetro, em que se considera os seguintes valores para o parâmetro de dificuldade: $b_i = -1,0; 0,0; 1,0$ e $2,0$.

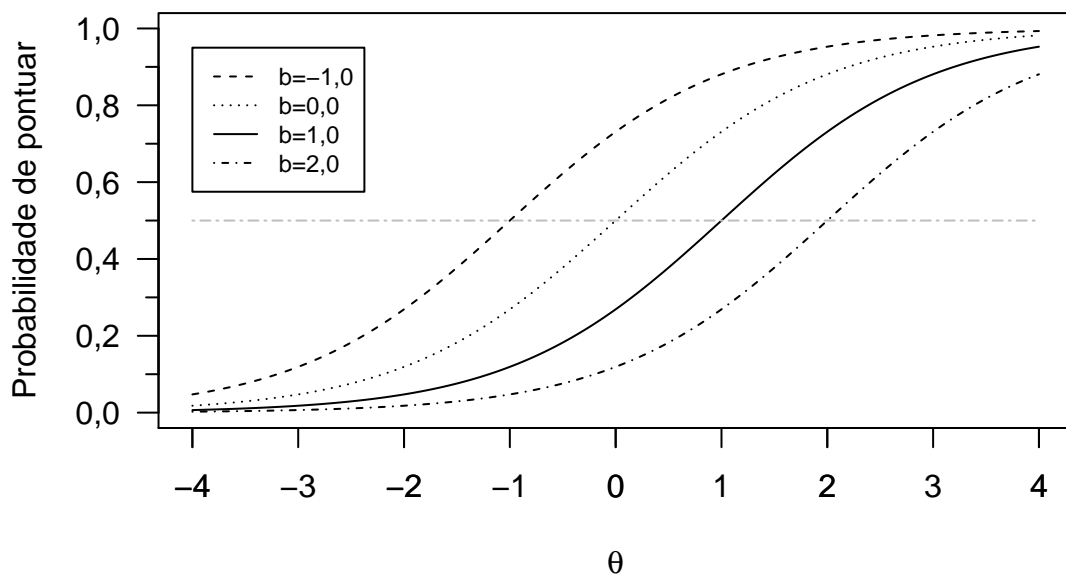


Figura 2 – Quatro curvas características do item (CCI) considerando diferentes valores do parâmetro de dificuldade do modelo logístico de 1 parâmetro.

Fonte: Elaborada pelo autor.

3.1.2 Modelo logístico de 2 parâmetros

A curva característica do item do modelo logístico de 2 parâmetros é dada pela seguinte equação

$$P(Y_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \quad (3.2)$$

com $i = 1, \dots, I$, $j = 1, \dots, n$, $\theta_j \in \mathbb{R}$, $b_i \in \mathbb{R}$ e $a_i \in \mathbb{R}_+$, em que $P(Y_{ij} = 1|\theta_j)$, θ_j , b_i e D são definidos como na equação (3.1), e a_i é o **parâmetro de discriminação do item i** . Este é um parâmetro de inclinação da CCI na escala da habilidade. Assim, itens com valores do parâmetro de discriminação maiores são mais úteis para discriminar indivíduos com traço latente em torno do parâmetro de dificuldade. Além disso, o parâmetro de discriminação deve ser positivo. A Figura 3 apresenta quatro CCI's do modelo logístico de 2 parâmetros com o parâmetro de dificuldade fixo em $b_i = 1,0$, mas com diferentes valores do parâmetro de discriminação: $a_i = 0,5; 1,0; 1,5$ e $2,0$.

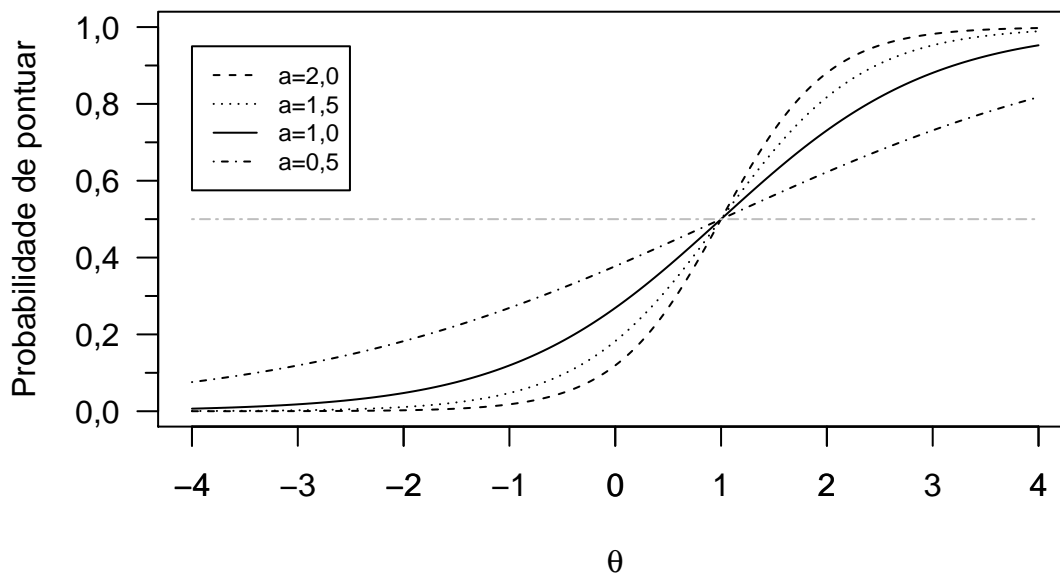


Figura 3 – Quatro curvas características do item (CCI) considerando diferentes valores do parâmetro de discriminação e mesmo valor do parâmetro de dificuldade ($b_i = 1,0$) do modelo logístico de 2 parâmetros.

Fonte: Elaborada pelo autor.

3.1.3 Modelo logístico de 3 parâmetros

A expressão matemática para o modelo logístico de 3 parâmetros é dada por:

$$P(Y_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \quad (3.3)$$

com $i = 1, \dots, I$, $j = 1, \dots, n$, $\theta_j \in \mathbb{R}$, $b_i \in \mathbb{R}$, $a_i \in \mathbb{R}_+^*$ e $c_i \in [0, 1]$, em que $P(Y_{ij} = 1|\theta_j)$, θ_j , b_i , D e a_i são definidos como no modelo logístico de 2 parâmetros dado pela expressão (3.2). O parâmetro c_i no modelo logístico de 3 parâmetros representa **a probabilidade de indivíduos com baixa habilidade pontuar no item i** , ou seja, representa a probabilidade de um acerto casual, e pode ser chamado de "parâmetro de adivinhação". A Figura 4 apresenta três CCI's do modelo logístico de 3 parâmetros, em que os parâmetros de dificuldade e de discriminação são fixos em $a_i = 1,0$ e $b_i = 1,0$, e diferentes valores do parâmetro de adivinhação, que são $c_i = 0,0; 0,1$ e $0,3$.

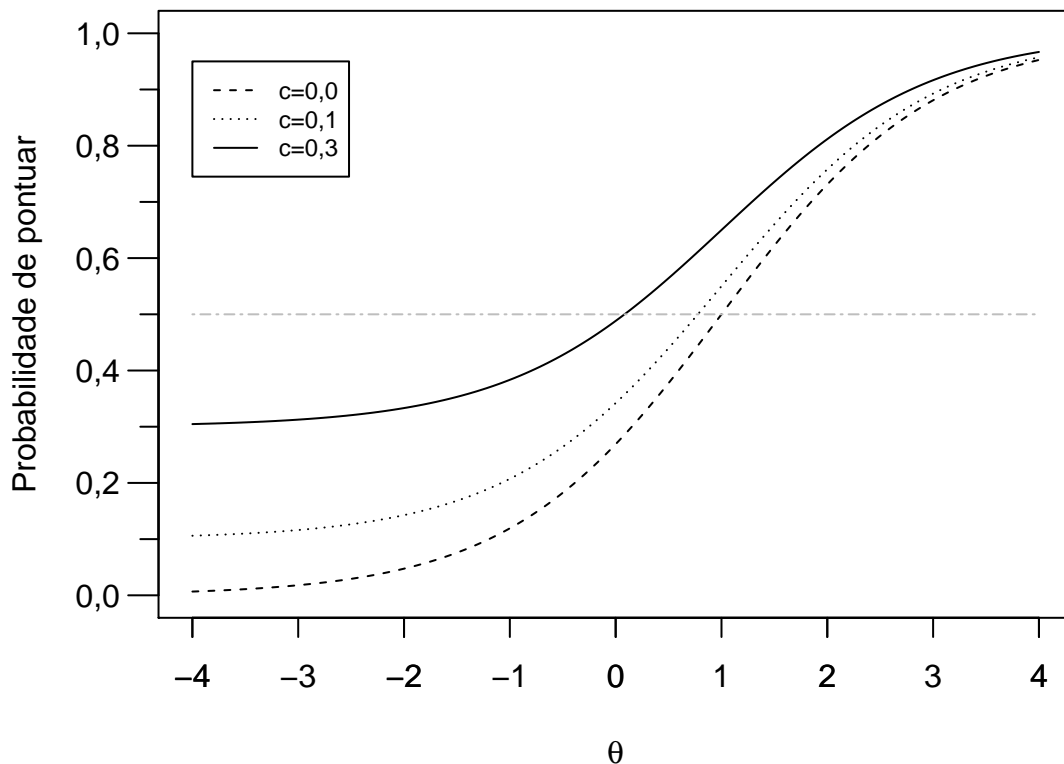


Figura 4 – Três curvas características do item (CCI) considerando diferentes valores do parâmetro de adivinhação, e com parâmetros de discriminação e de dificuldade fixos ($a_i = 1,0$ e $b_i = 1,0$) do modelo logístico de 3 parâmetros.

Fonte: Elaborada pelo autor.

Um exemplo de aplicação

No capítulo 2 do livro de [Reckase \(2009\)](#) é apresentado um exemplo de parâmetros de itens para um conjunto de 50 itens. Os itens podem ser de múltipla escolha ou abertos e são classificados como corretos ou incorretos. Os parâmetros de adivinhação para os itens de resposta aberta foram considerados como 0, pois não havia nenhuma chance de adivinhar a resposta para esses itens.

Com o objetivo de apresentar um exemplo simples e didático, dentre os 50 itens do exemplo apresentado no livro, 3 itens foram selecionados para as curvas características dos itens serem apresentadas neste trabalho. A Tabela 4 apresenta os parâmetros de discriminação, a , dificuldade, b , e adivinhação, c , dos itens 3, 11 e 38 e a Figura 5 apresenta as curvas características de cada item.

Tabela 4 – Parâmetros de itens para 3 itens (3, 11 e 38) de teste com pontuação dicotômica selecionados dentre os 50 itens apresentados em [Reckase \(2009\)](#).

Item	a	b	c
3	1,47	0,06	0,00
11	0,70	1,49	0,00
38	2,01	1,19	0,28

Os itens 3 e 38 possuem valores altos do parâmetro de discriminação, a , sendo, respectivamente 1,47 e 2,01 e, é possível notar na Figura 5 que as curvas desses itens são mais íngrimes, indicando uma discriminação maior entre os indivíduos com traço latente em torno do parâmetro de dificuldade do que no item 11, que o parâmetro de discriminação é 0,70. Já os itens 11 e 38 possuem os maiores valores do parâmetro de dificuldade, b , sendo iguais a 1,49 e 1,19, respectivamente. Porém o item 38 leva em consideração o acerto casual, sendo o parâmetro de adivinhação, c , igual a 0,28. Assim, neste item, indivíduos com baixo valor do traço latente possui uma maior probabilidade de pontuar do que nos outros itens, que o parâmetro de adivinhação é igual a 0,00. Já no item 11, nota-se que existe uma maior dificuldade de indivíduos com baixo valor do traço latente pontuar do que no item 3, que o parâmetro de dificuldade é 0,06.

É importante salientar que é desejável que os itens do instrumento de medição, que neste caso é um teste, possua diferentes valores para os parâmetros de dificuldade, de forma que indivíduos com diferentes níveis de traço latente, inclusive baixos, tenham probabilidade de acerto maior do que 0,5 em alguns itens do teste.

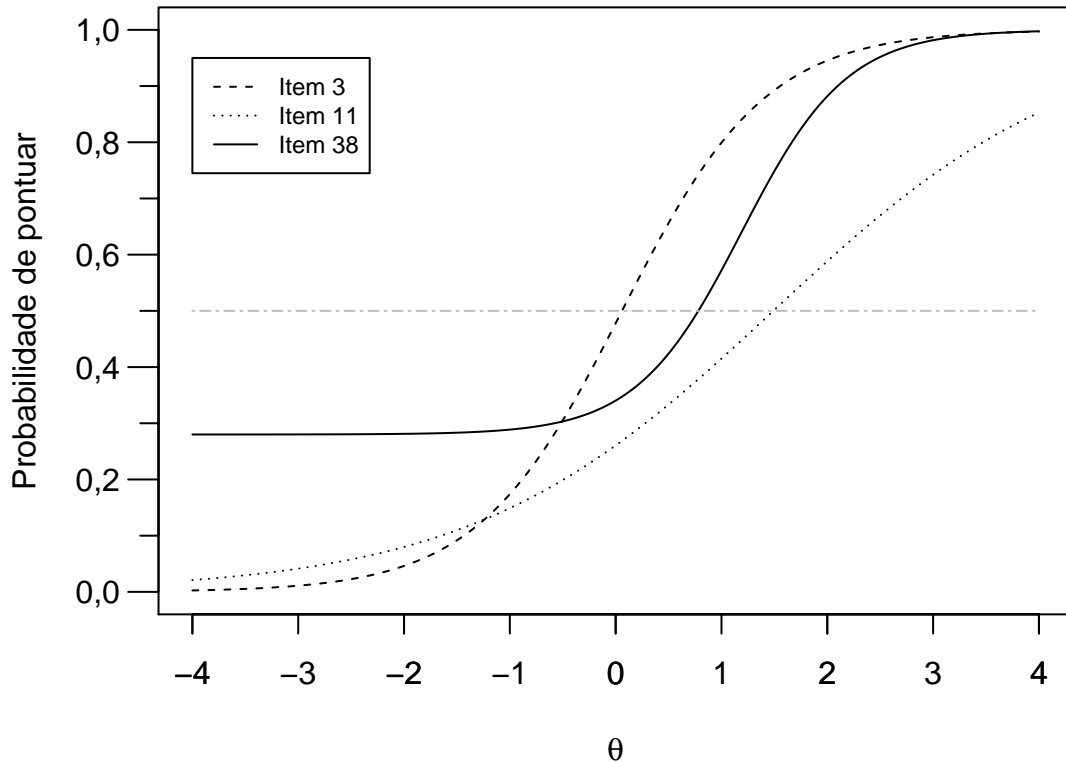


Figura 5 – Três curvas características do item (CCI) dos itens 3, 11 e 38 apresentados em [Reckase \(2009\)](#).

Fonte: Elaborada pelo autor.

3.2 Modelos para itens politômicos

Os itens com resposta politômica são aqueles que possuem mais do que duas categorias de resposta, como, por exemplo, itens de múltipla escolha em que se leva em consideração não somente se o indivíduo respondeu o item corretamente ou incorretamente, mas também qual a resposta dada pelo indivíduo. O modelo de resposta gradual ([SAMEJIMA, 1969](#)), o modelo de resposta nominal ([BOCK, 1972](#)), o modelo de escala gradual ([ANDRICH, 1978](#)), o modelo de crédito parcial ([MASTERS, 1982](#)) e o modelo de crédito parcial generalizado ([MURAKI, 1992](#)) são exemplos de modelos para itens politômicos. A seguir será apresentado o modelo de resposta gradual proposto por [Samejima \(1969\)](#), que será o foco deste trabalho.

3.2.1 Modelo de resposta gradual

O modelo de resposta gradual assume que as categorias de resposta de um item podem ser ordenadas, como por exemplo dados em uma escala Likert. Supõe-se que as categorias dos itens são ordenadas da menor para a maior, isto é, $k = 0, 1, \dots, m_i$, em que $(m_i + 1)$ é o número

de categorias do i -ésimo item.

Pode-se definir a probabilidade do j -ésimo indivíduo responder uma categoria específica a partir de probabilidades acumuladas $P_{i,k}^+(\theta_j)$, que são probabilidades do j -ésimo indivíduo selecionar determinada categoria k ou outra superior no i -ésimo item. A expressão é dada por um extensão do modelo logístico de 2 parâmetros:

$$P_{i,k}^+(\theta_j) = P(Z_{ij} \geq k | \theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{ik})}}, \quad (3.4)$$

com $i = 1, \dots, I, j = 1, \dots, n, k = 0, \dots, m_i, \theta_j \in \mathbb{R}, b_{ik} \in \mathbb{R}$ e $a_i \in \mathbb{R}_+^*$. Z_{ij} representa a categoria respondida pelo j -ésimo indivíduo no i -ésimo item e b_{ik} representa o parâmetro de dificuldade da k -ésima categoria do item i . Esse parâmetro é conhecido como parâmetro "limiar" e refere-se ao nível do traço latente em que a probabilidade de um indivíduo selecionar uma categoria ou outra de ordem superior é de 0,5 (VINCENZI *et al.*, 2018). Os demais parâmetros do modelo possuem as mesmas definições apresentadas no modelo logístico de 2 parâmetros.

Pode-se definir que:

$$Y_{ijk} = \begin{cases} 1, & \text{se o } j\text{-ésimo indivíduo escolhe a categoria } k \text{ no } i\text{-ésimo item;} \\ 0, & \text{caso contrário.} \end{cases}$$

Considere também as seguintes notações:

- $\mathbf{Y}_{ij.} = (Y_{ij1}, \dots, Y_{ijm_i})^\top$ é o vetor aleatório de respostas do j -ésimo indivíduo ao i -ésimo item;
- $\mathbf{Y}_{.j.} = (\mathbf{Y}_{1j.}^\top, \dots, \mathbf{Y}_{Ij.}^\top)^\top$ é o conjunto de respostas do j -ésimo indivíduo aos I itens;
- $\mathbf{Y}_{...} = (\mathbf{Y}_{.1.}^\top, \dots, \mathbf{Y}_{.n.}^\top)^\top$ é o conjunto de respostas dos n indivíduos aos I itens;
- $\boldsymbol{\xi}_i = (a_i, b_{i1}, \dots, b_{im_i})^\top$ é o vetor de parâmetros do i -ésimo item;
- $\boldsymbol{\xi} = (\boldsymbol{\xi}_1^\top, \dots, \boldsymbol{\xi}_I^\top)^\top$ é o conjunto de parâmetros dos I itens;
- $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)^\top$ é o vetor de traços latentes dos n indivíduos.

Em um item com $(m_i + 1)$ categorias, é necessário estimar m_i parâmetros de dificuldades, além do parâmetro de inclinação (discriminação) do item. Logo, em um teste com I itens, e cada item contendo $(m_i + 1)$ categorias, tem-se que o número de parâmetros a serem estimados é $\sum_{i=1}^I m_i + I$.

Pela definição, deve haver necessariamente uma ordenação entre os parâmetros de dificuldade das categorias de um item de acordo com a classificação de seus escores, assim,

$$b_{i1} \leq b_{i2} \leq \dots \leq b_{im_i}.$$

Além disso, a partir das informações anteriores, tem-se que

$$Z_{ij} | \theta_j, \xi_i \sim \text{Categorica}(P_{ij0}, \dots, P_{ijm_i}),$$

em que θ_j é o traço latente do j -ésimo indivíduo e ξ_i é o vetor de parâmetros relacionados ao i -ésimo item.

No modelo de resposta gradual, a probabilidade de um indivíduo pontuar uma categoria específica é modelada a partir de probabilidades acumuladas, P_{ijk}^+ , que são as probabilidades do indivíduo j selecionar uma categoria específica k ou outra categoria maior no item i . Assim,

$$P_{ijk}^+ = P(X_{ij} \geq k | \theta_j, \xi_i) = L(\eta_{ijk}),$$

em que $L(\cdot)$ representa a função de ligação e geralmente utiliza-se a distribuição acumulada logística como função de ligação, como é apresentado na equação (3.4), por exemplo.

Portanto, a probabilidade que o indivíduo j selecione a categoria k no item i é obtida subtraindo as probabilidades acumuladas adjacentes, isto é,

$$P_{ijk} = P(X_{ij} = k | \theta_j, \xi_i) = \begin{cases} 1 - P_{ij(k+1)}^+ & \text{se } k = 0 \\ P_{ijk}^+ - P_{ij(k+1)}^+ & \text{se } 1 \leq k < m_i \\ P_{ijk}^+ & \text{se } k = m_i \end{cases}.$$

E, em termos da função de ligação,

$$P_{ijk} = P(X_{ij} = k | \theta_j, \xi_i) = \begin{cases} 1 - L(a_i(\theta_j - b_{i(k+1)})) & \text{se } k = 0 \\ L(a_i(\theta_j - b_{ik})) - L(a_i(\theta_j - b_{i(k+1)})) & \text{se } 1 \leq k < m_i \\ L(a_i(\theta_j - b_{ik})) & \text{se } k = m_i \end{cases}.$$

Nos modelos para itens politômicos, as curvas características dos itens são denominadas por alguns autores de curvas características das categorias de um item (CCCI). Na Figura 6 são apresentadas as curvas características das categorias acumuladas de um item, em que os parâmetros de dificuldade são $b_{i1} = -2,0$, $b_{i2} = -1,0$, $b_{i3} = 0,5$, $b_{i4} = 1,0$ e $b_{i5} = 1,8$ e o parâmetro de discriminação é $a_i = 1,0$. Assim como no modelo logístico de 2 parâmetros, o parâmetro de dificuldade indica a posição na escala do traço latente em que a probabilidade de selecionar a categoria ou outra maior é igual a 0,5, e as inclinações são mais íngrimes nesse ponto. A inclinação depende do parâmetro de discriminação.

Na Figura 7 são apresentadas as curvas características das categorias deste item. A curva correspondente à categoria 0 é uma função decrescente de θ , e nesta categoria o parâmetro de dificuldade $b_{i1} = -2,0$ indica a posição na escala do traço latente em que a probabilidade selecionar a categoria 0 é 0,5. Na categoria de pontuação mais alta, neste caso a categoria 5, a curva característica desta categoria indica que o parâmetro de dificuldade $b_{i5} = 1,8$ é igual a posição na escala do traço latente do indivíduo quando a probabilidade de selecionar a categoria 5 é 0,5. As outras curvas não têm relação óbvia com os parâmetros do item (RECKASE, 2009).

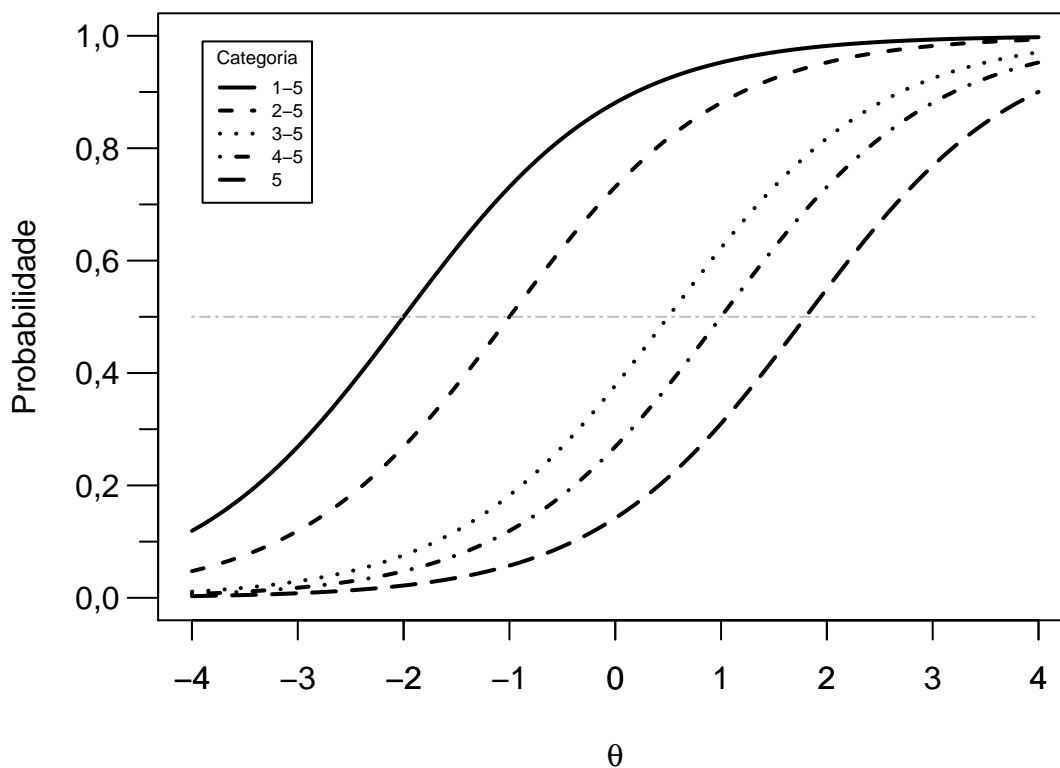


Figura 6 – Curvas características das categorias acumuladas para um item ajustado pelo modelo de resposta gradual com $a_i = 1,0$, $b_{i1} = -2,0$, $b_{i2} = -1,0$, $b_{i3} = 0,5$, $b_{i4} = 1,0$ e $b_{i5} = 1,8$.

Fonte: Elaborada pelo autor.

Nesta seção apresentamos modelos da TRI supondo a unidimensionalidade, em que apenas uma habilidade é medida pelo conjunto de itens de um teste. Entretanto, é possível que indivíduos usem mais do que uma única habilidade para responder a determinado item e/ou teste. Além disso, um teste pode requerer que o indivíduo utilize múltiplas habilidades para respondê-lo. Modelos que levam em consideração essas características são denominados multidimensionais, e possui a vantagem de refletir a interação entre os indivíduos e os itens do teste com mais precisão do que os modelos unidimensionais da TRI (ZHANG; STONE, 2008). A multidimensionalidade será tratada mais detalhadamente a seguir.

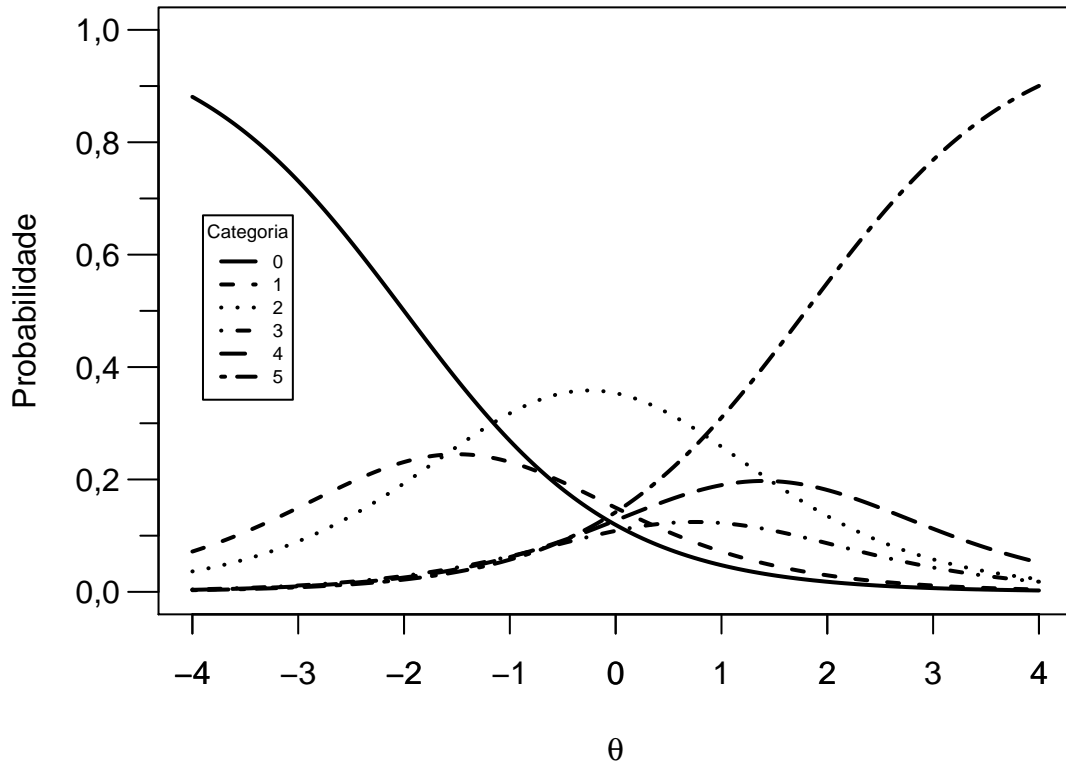


Figura 7 – Curvas características das categorias para um item ajustado pelo modelo de resposta gradual com $a_i = 1,0$, $b_{i1} = -2,0$, $b_{i2} = -1,0$, $b_{i3} = 0,5$, $b_{i4} = 1,0$ e $b_{i5} = 1,8$.

Fonte: Elaborada pelo autor.

3.2.2 Modelos Multidimensionais

Diferentes estruturas multidimensionais serão apresentadas a seguir e podem ser vistas mais detalhadamente em [Martelli \(2014\)](#).

Modelo unidimensional consecutivo: São ajustados sequencialmente modelos da TRI unidimensionais em cada subteste. Assim, é possível mensurar cada traço latente especificamente, porém, não é possível mensurar a relação entre eles.

Modelo multiunidimensional: Leva-se em consideração que os traços latentes podem ser correlacionados e a intensidade dessas associações podem ser obtidas diretamente.

Modelo bi-fatorial: Assume-se que uma habilidade geral (ou comum) e habilidades específicas afetam a resposta de cada item. Este é um caso de que um único item mensura mais do que um traço latente. Além disso, é ignorado associações entre os traços latentes.

Modelo hierárquico: A estrutura hierárquica e as construções latentes específicas são modeladas explicitamente, assim, itens em um mesmo subteste mensuram uma habilidade específica, e por sua vez, cada habilidade específica é influenciada por uma habilidade geral. A depender da relação entre a habilidade específica e a habilidade geral, diferentes estruturas hierárquicas podem ser utilizadas. Por exemplo, pode-se considerar cada habilidade específica como uma função linear da habilidade geral ou considerar que cada habilidade específica se combina linearmente para formar a habilidade geral.

Modelo aditivo: A estrutura latente é de tal modo que a resposta a um item do teste é afetada tanto pela habilidade específica quanto pela geral, tal que, as habilidades latentes formam uma estrutura aditiva. Este modelo tem uma estrutura parecida com o modelo bi-fatorial, porém aqui todas as habilidades podem ser correlacionadas.

O modelo multiunidimensional para dados dicotômicos e para dados politômicos será apresentado com mais detalhes a seguir.

3.2.2.1 Modelo multiunidimensional para dados dicotômicos

Levando-se em consideração agora a multidimensionalidade de um teste, temos que o conjunto de itens mensura mais do que um único traço latente. Cada indivíduo j é caracterizado por um vetor de habilidades latentes, isto é, $\theta_j = (\theta_{1j}, \theta_{2j}, \dots, \theta_{Hj})^\top$, em que $j = 1, \dots, n$ e H representa o número de traços latentes que o teste mensura. O parâmetro de discriminação também é representado como um vetor, α_i , em que cada elemento deste vetor é representado por a_{hi} , com $h = 1, \dots, H$, assim, $\alpha_i = (a_{1i}, a_{2i}, \dots, a_{Hi})^\top$, em que $i = 1, \dots, I$, tal que I é o número de itens. No caso da estrutura multiunidimensional, o vetor de parâmetros de discriminação pode ser simplificado para $\alpha_i = (0, \dots, 0, a_{hi}, 0, \dots, 0)^\top$, em que cada traço latente é relacionado com um único conjunto de itens.

Considerando um teste com I itens, estruturado com H subtestes, cada um contendo I_h itens que mensuram um único traço latente, temos que a probabilidade de um indivíduo j responda corretamente ao item i que pertence ao h -ésimo subteste é dado pela generalização multivariada do modelo logístico de 2 parâmetros, isto é,

$$P(Y_{hij} = 1 | \theta_j, \alpha_i, b_i) = \frac{1}{1 + e^{-D \sum_{h=1}^H a_{hi} \theta_{hj} - b_i}}, \quad (3.5)$$

com $i = 1, \dots, I$, $j = 1, \dots, n$, $h = 1, \dots, H$ e Y_{hij} é uma variável dicotômica que pode ser representada da seguinte maneira

$$Y_{hij} = \begin{cases} 1, & \text{se o } j\text{-ésimo indivíduo pontuar no } i\text{-ésimo item do } h\text{-ésimo subteste;} \\ 0, & \text{se o } j\text{-ésimo indivíduo não pontuar no } i\text{-ésimo item do } h\text{-ésimo subteste.} \end{cases}$$

Agora, a_{hi} representa o parâmetro de discriminação na h -ésima dimensão do i -ésimo item, θ_{hj} representa o traço latente do j -ésimo indivíduo na h -ésima dimensão, D é um fator de escala

constante e igual a 1 se deseja-se obter resultado da função logística e igual a 1,7 se deseja-se obter a aproximação para a distribuição normal e b_i é o parâmetro de dificuldade do i -ésimo item.

3.2.2.2 Modelo multiunidimensional para dados politômicos - o modelo de resposta gradual

No caso dos modelos multidimensionais para dados ordinais, muitos autores desenvolveram suas formalizações a partir de extensões dos modelos unidimensionais (MARTELLI, 2014). Nesta seção será apresentado o modelo de resposta gradual multidimensional, que é o foco deste trabalho.

Assim como no modelo de resposta gradual unidimensional, vamos considerar n indivíduos e um conjunto de I itens ordinais, em que a resposta Z_{hij} do j -ésimo indivíduo, $j = 1, \dots, n$, ao i -ésimo item, $i = 1, \dots, I_h$ do h -ésimo subteste assume um valor da categoria k , $k = 0, 1, \dots, m_i$. As categorias dos itens são ordenadas da menor para a maior e o número de categorias do i -ésimo item é $(m_i + 1)$. Agora vamos considerar a existência de múltiplos traços latentes, $\theta_j = (\theta_{1j}, \theta_{2j}, \dots, \theta_{Hj})^\top$, em que H representa o número de traços latentes que o teste mensura.

A probabilidade que o indivíduo j selecione a categoria k no item i é representada no modelo de resposta gradual multidimensional de forma similar ao modelo unidimensional apresentado anteriormente, porém com a notação vetorial do parâmetro que representa a habilidade, isto é,

$$P_{hijk}^+ = P(Z_{hij} \geq k | \theta_j, \xi_{hi}) = L(\eta_{hijk}),$$

em que ξ_{hi} é o vetor de parâmetros relacionados ao i -ésimo item do h -ésimo subteste e $L(\cdot)$ representa a função de ligação, geralmente utiliza-se a distribuição acumulada logística como função de ligação. Assim,

$$P_{hijk} = P(Z_{hij} = k | \theta_j, \xi_{hi}) = \begin{cases} 1 - P_{hij(k+1)}^+ & \text{se } k = 0 \\ P_{hijk}^+ - P_{hij(k+1)}^+ & \text{se } 1 \leq k < m_i \\ P_{hijk}^+ & \text{se } k = m_i \end{cases} .$$

No caso da estrutura multiunidimensional, cada dimensão do vetor de traços latentes, $\theta_j = (\theta_{1j}, \theta_{2j}, \dots, \theta_{Hj})^\top$, do indivíduo j é mensurado por um conjunto específico de itens do teste. Assim, vamos considerar um teste com I itens, estruturado com H subtestes, cada um contendo I_h itens que mensuram um único traço latente. Temos que a probabilidade do j -ésimo indivíduo responder uma categoria específica a partir de probabilidades acumuladas $P_{i,k}^+(\theta_{hj})$, que são probabilidades do j -ésimo indivíduo selecionar determinada categoria k ou outra maior

no i -ésimo item pertencente ao h -ésimo subteste é dada por

$$P_{i,k}^+(\boldsymbol{\theta}_{hj}) = P(Z_{hij} \geq k | \boldsymbol{\theta}_j, \mathbf{a}_i, b_i) = \frac{1}{1 + e^{-D \sum_{\ell=1}^H a_{\ell i} \theta_{\ell j} - b_{ik}}}, \quad (3.6)$$

com $i = 1, \dots, I_h$, $j = 1, \dots, n$, $k = 0, \dots, m_i$, $h = 1, \dots, H$ é o número de dimensões do teste, $\theta_{hj} \in \mathbb{R}$, $b_{ik} \in \mathbb{R}$ e $a_{hi} \in \mathbb{R}_+^*$. Agora a_{hi} representa o parâmetro de discriminação na h -ésima dimensão do i -ésimo item, θ_{hj} representa o traço latente do j -ésimo indivíduo na h -ésima dimensão, D é um fator de escala constante e igual a 1 se deseja-se obter resultado da função logística e igual a 1,7 se deseja-se obter a aproximação para a distribuição normal e b_{ik} é o parâmetro de dificuldade do i -ésimo item e k -ésima categoria.

No próximo capítulo, métodos de estimação para os parâmetros do modelos de resposta gradual serão descritos. Os métodos considerados são o método da máxima verossimilhança marginal, que utiliza a inferência clássica, e o método Bayesiano via Monte Carlo cadeia de Markov amostrador de Gibbs.

MÉTODOS DE ESTIMAÇÃO

Em geral, tanto os parâmetros que caracterizam os itens quanto os parâmetros da habilidade dos indivíduos são desconhecidos. Como a probabilidade de responder determinada alternativa em um item depende dos parâmetros do item e da habilidade do indivíduo, então uma das principais etapas na TRI é a estimação desses parâmetros. No caso do **modelo de resposta gradual**, a estimação dos parâmetros pode ser feita por uma abordagem clássica, através do método da máxima verossimilhança, ou por uma abordagem Bayesiana. A seguir, nas seções 4.1 e 4.2, serão apresentados os métodos de estimação clássico e Bayesiano, respectivamente, para esse modelo e podem ser vistos mais detalhadamente em [Azevedo \(2003\)](#) e [Andrade, Tavares e Valle \(2000\)](#). Aqui será apresentada a situação mais comum, em que ambos os **parâmetros do item e do indivíduo são desconhecidos**.

4.1 Estimação Clássica

Para o caso em que as habilidades são desconhecidas, o método da máxima verossimilhança marginal ([BOCK; LIEBERMAN, 1970](#); [BOCK; AITKIN, 1981](#)) é o processo de estimação dos parâmetros dos itens mais utilizado e com propriedades importantes. A proposta do método é fazer a estimação em duas etapas: primeiro os parâmetros dos itens, e posteriormente, as habilidades. As habilidades não são conhecidas, e por isso, é necessário marginalizar a verossimilhança integrando-a com relação à distribuição da habilidade, de forma que a verossimilhança não seja uma função das habilidades.

Neste caso, a ideia básica é considerar que o traço latente (habilidade) possui uma distribuição (latente) de probabilidade, e que os n indivíduos são uma amostra aleatória simples de uma população com essa distribuição ([AZEVEDO, 2003](#) apud [ANDERSEN, 1980](#)). Considera-se uma densidade $g(\theta|\eta)$ duplamente diferenciável, em que η é um vetor de parâmetros associados a $g(\theta)$ com as componentes conhecidas e finitas. Geralmente considera-se que θ é normalmente distribuído com média μ e variância σ^2 , ou seja, $\eta = (\mu, \sigma^2)$.

Para o processo de estimação, as seguintes suposições são necessárias:

- (S1) as respostas oriundas de indivíduos diferentes são independentes;
- (S2) os itens são respondidos de forma independente por cada indivíduo (independência local), dado o seu traço latente;
- (S3) a probabilidade de escolha da categoria k no item i pelo indivíduo j pode ser representada pelo modelo multivariado de Bernoulli

$$P(\mathbf{Y}_{ij} = \mathbf{y}_{ij} | \boldsymbol{\theta}, \boldsymbol{\xi}_i) \equiv P(\mathbf{Y}_{ij} = \mathbf{y}_{ij}) = \prod_{k=1}^{m_i} P_{ijk}^{y_{ijk}},$$

em que \mathbf{Y}_{ij} é o vetor aleatório de respostas do j -ésimo indivíduo ao i -ésimo item, \mathbf{y}_{ij} é o vetor de resposta do j -ésimo indivíduo ao i -ésimo item e P_{ijk} é a probabilidade de resposta na k -ésima categoria no i -ésimo item do j -ésimo indivíduo.

4.1.1 Estimação dos parâmetros dos itens

Inicialmente no processo de estimação deve-se construir a probabilidade marginal de respostas para um determinado indivíduo de tal forma a eliminar as habilidades. Assim, tem-se que a probabilidade marginal é dada por

$$P(\mathbf{Y}_{.j} | \boldsymbol{\xi}, \boldsymbol{\eta}) = \int_{\mathbb{R}} P(\mathbf{Y}_{.j} | \boldsymbol{\xi}, \boldsymbol{\theta}) g(\boldsymbol{\theta} | \boldsymbol{\eta}) d\boldsymbol{\theta}.$$

Pela independência entre as respostas de diferentes indivíduos, tem-se que a verossimilhança é dada por

$$L(\boldsymbol{\xi}, \boldsymbol{\eta}) = P(\mathbf{Y}_{\dots} | \boldsymbol{\xi}, \boldsymbol{\eta}) = \prod_{j=1}^n P(\mathbf{Y}_{.j} | \boldsymbol{\xi}, \boldsymbol{\eta})$$

e a log-verossimilhança é dada por

$$l(\boldsymbol{\xi}, \boldsymbol{\eta}) = \sum_{j=1}^n \ln P(\mathbf{Y}_{.j} | \boldsymbol{\xi}, \boldsymbol{\eta}) \quad (4.1)$$

Derivando (4.1) com relação aos parâmetros dos itens, obtêm-se as componentes do vetor score, que não possuem forma fechada, e por isso são necessários métodos numéricos de aproximação de integrais. O método de *Gauss-Hermite*, também chamado de método da quadratura gaussiana, é um dos métodos de aproximação de integrais mais utilizado na TRI. Se $g(\boldsymbol{\theta} | \boldsymbol{\eta})$ é uma função contínua com integral finita, ela pode ser aproximada, com qualquer grau de precisão, por uma distribuição discreta, que assume um número finito de pontos. Assim, a ideia é substituir o cálculo de uma integral por uma soma finita de áreas de q retângulos. Detalhes sobre este método podem ser vistos em [Hildebrand \(1987\)](#). Os pontos médios dos retângulos, $\bar{\boldsymbol{\theta}}_l$,

$l = 1, \dots, q$, são chamados de nós (ou pontos de quadratura). Associa-se um peso a cada ponto de quadratura, $A_l \equiv A(\bar{\theta}_l)$, que é composto pelo valor da função $g(\bar{\theta}_l|\boldsymbol{\eta})$ (altura) e o respectivo comprimento do seu intervalo Δ_l (largura), assim, $A_l = g(\bar{\theta}_l|\boldsymbol{\eta}) \times \Delta_l$.

Assim, as componentes do vetor escore em forma de pontos de quadratura são dadas por

$$S(a_i) = D \sum_{l=1}^q \sum_{k=0}^{m_i} \left\{ \frac{\bar{r}_{ilk}}{P_{ilk}} \left[(\bar{\theta}_l - b_{ik}) W_{ilk}^+ - (\bar{\theta}_l - b_{i(k+1)}) W_{i(k+1)}^+ \right] \right\} \quad (4.2)$$

e

$$S(b_{ih}) = Da_i \sum_{l=1}^q \left\{ W_{ilh}^+ \left[\frac{\bar{r}_{il(h-1)}}{P_{il(h-1)}} - \frac{\bar{r}_{ilh}}{P_{ilh}} \right] \right\}. \quad (4.3)$$

Nas equações (4.2) e (4.3), $i = 1, \dots, I$, $h = 1, \dots, m_i - 1$, $l = 1, \dots, q$, $k = 0, \dots, m_i$, D e b_{ik} são definidos no Capítulo 3 e $\bar{r}_{ilk} = \sum_{j=1}^n y_{ijk} g_j^*(\bar{\theta}_l)$, com

$$g_j^*(\bar{\theta}_l) = \frac{P(\mathbf{Y}_{.j}|\boldsymbol{\xi}, \bar{\theta}_l) g(\bar{\theta}_l|\boldsymbol{\eta})}{\sum_{l=1}^q P(\mathbf{Y}_{.j}|\boldsymbol{\xi}, \bar{\theta}_l) g(\bar{\theta}_l|\boldsymbol{\eta})},$$

desde que $\Delta_l \cong 1$, $P_{ilk} \equiv P_{ijk}$ com θ_j substituído por $\bar{\theta}_l$, $W_{ilk}^+ = P_{ilk}^+ Q_{ilk}^+$, com $P_{ilk}^+ = P(Z_{il} \geq k)$ e $Q_{ilk}^+ = 1 - P_{ilk}^+$.

Para resolução das equações de estimação (4.2) e (4.3), utiliza-se uma adaptação do **algoritmo EM**, que será descrita mais adiante. O algoritmo EM (DEMPSTER; LAIRD; RUBIN, 1977) é utilizado para o cálculo do estimador de máxima verossimilhança de forma iterativa, e geralmente é utilizado em situações envolvendo variáveis não observadas. A cada iteração do algoritmo EM, ocorre dois passos: o Passo E, que calcula o valor esperado do logaritmo da verossimilhança, e o Passo M, que encontra seu máximo. A ideia é repetir estes passos até obter a convergência, de acordo com o critério de parada, definido previamente.

No caso da TRI, a ideia é obter estimativas de $\boldsymbol{\xi}$ considerando variáveis não observadas $\boldsymbol{\theta}$. Assim, de um modo geral, \mathbf{Y}_{\dots} , definido na seção 3.2.1, é o vetor de dados incompletos e $(\mathbf{Y}_{\dots}, \boldsymbol{\theta})$ é o vetor de dados completos. A densidade conjunta dos dados completos é dada por $f(\mathbf{Y}_{\dots}, \boldsymbol{\theta}|\boldsymbol{\xi})$, e $\hat{\boldsymbol{\xi}}^{(t)}$ é uma estimativa de $\boldsymbol{\xi}$ na iteração t , então os passos EM para obtenção de $\hat{\boldsymbol{\xi}}^{(t+1)}$ são:

Passo E: Calcular $E[\ln f(\mathbf{Y}_{\dots}, \boldsymbol{\theta}|\boldsymbol{\xi})|\mathbf{Y}_{\dots}, \hat{\boldsymbol{\xi}}^{(t)}]$.

Passo M: Obter $\hat{\boldsymbol{\xi}}^{(t+1)}$ que maximiza a função no Passo E.

Para utilizar a adaptação do algoritmo EM na resolução das equações de estimação, supõe-se que os traços latentes são restritos a um conjunto de q valores, $\bar{\theta}_l$, $l = 1, \dots, q$, com probabilidades π_l . Seja f_{il} e r_{ilk} o número de indivíduos com traço latente em torno de $\bar{\theta}_l$ que respondem ao item i , e dentre esses indivíduos, o número daqueles que escolhem a categoria k , respectivamente. Define-se $\mathbf{f}_i = (f_{i1}, \dots, f_{iq})^\top$, $\mathbf{f} = (\mathbf{f}_1^\top, \dots, \mathbf{f}_I^\top)^\top$, $\mathbf{r}_{il} = (r_{il1}, \dots, r_{ilm_i})^\top$, $\mathbf{r}_l =$

$(\mathbf{r}_{1l}^\top, \dots, \mathbf{r}_{ll}^\top)^\top$, $\mathbf{r} = (\mathbf{r}_1^\top, \dots, \mathbf{r}_q^\top)^\top$, $\bar{r}_{ilk} = E(r_{ilk} | \mathbf{Y}_{\dots}, \boldsymbol{\xi})$ e $\bar{f}_{il} = E(f_{il} | \mathbf{Y}_{\dots}, \boldsymbol{\xi})$. Como as quantidades f_{il} e r_{ilk} são desconhecidas, estas são tratadas como variáveis não observadas no algoritmo EM. Assim, os passos da adaptação do algoritmo EM são dados por:

Passo E: Usar os pontos de quadratura $\bar{\theta}_l$, os pesos $A_l^{(t)}$, $l = 1, \dots, q$ e estimativas provisórias dos parâmetros dos itens, $\hat{\xi}_i^{(t)}$, $i = 1, \dots, I$, para gerar $g_j^*(\bar{\theta}_l)^{(t)}$ e, posteriormente, $\bar{r}_{ilk}^{(t)}$ e $\bar{f}_{il}^{(t)}$, $i = \dots, I$, $l = 1, \dots, q$ e $k = 0, \dots, m_i$.

Passo M: Com $\bar{\mathbf{r}}^{(t)}$ e $\bar{\mathbf{f}}^{(t)}$ obtidos no passo E, resolver as equações de estimação para ξ_i , $i = 1, \dots, I$, usando o algoritmo Newton-Raphson ou escore de Fisher.

A ideia dos métodos Newton-Raphson e escore de Fisher é determinar o ponto $\hat{\beta}$ que maximiza a função de verossimilhança $L(\beta)$ a partir da repetição de uma sequência de aproximações, em que o processo converge para a localização do máximo se a função é bem comportada e a aproximação inicial é boa. Aqui, β representa um vetor de P parâmetros a serem estimados. Seja $l(\beta) = \ln L(\beta)$ e $\hat{\beta}^{(t)}$ uma estimativa de β na iteração t , o método **Newton-Raphson** é descrito da seguinte maneira:

$$\hat{\beta}^{(t+1)} = \hat{\beta}^{(t)} - [\mathbf{J}(\hat{\beta}^{(t)})]^{-1} \mathbf{S}(\hat{\beta}^{(t)}), \quad (4.4)$$

em que:

$\mathbf{S}(\beta) = \frac{\partial l(\beta)}{\partial \beta}$ é o **vetor escore**, e $\mathbf{S}(\hat{\beta}^{(t)})$ é $\mathbf{S}(\beta)$ avaliado em $\hat{\beta}^{(t)}$;

$\mathbf{J}(\beta) = \frac{\partial^2 l(\beta)}{\partial \beta \partial \beta^\top}$ é a **matriz Hessiana**, dada pelas derivadas parciais de segunda ordem do logaritmo da função de verossimilhança, e $\mathbf{J}(\hat{\beta}^{(t)})$ é $\mathbf{J}(\beta)$ avaliado em $\hat{\beta}^{(t)}$.

O método **escore de Fisher** é descrito da seguinte maneira:

$$\hat{\beta}^{(t+1)} = \hat{\beta}^{(t)} - [\mathbf{I}(\hat{\beta}^{(t)})]^{-1} \mathbf{S}(\hat{\beta}^{(t)}), \quad (4.5)$$

em que:

$\mathbf{I}(\beta) = -E\left(\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta^\top}\right)$ é a **matriz de informação de Fisher**, que utiliza o valor esperado da matriz Hessiana, e $\mathbf{I}(\hat{\beta}^{(t)})$ é $\mathbf{I}(\beta)$ avaliado em $\hat{\beta}^{(t)}$.

Os métodos Newton-Raphson e escore de Fisher podem ser vistos mais detalhadamente em [Isaacson e Keller \(1966\)](#) e [Rao \(1973\)](#), respectivamente.

O procedimento do algoritmo EM adaptado gera estimativas consistentes para parâmetros dos itens, e a informação de Fisher, desde que positiva, fornecerá os erros-padrão assintóticos das estimativas. A superfície de verossimilhança é convexa em quase toda parte, o vetor de pontos estacionários será único, e portanto é o máximo global da função de verossimilhança ([SAMEJIMA, 1997](#); [BOCK; ZIMOWSKI, 1997](#)).

4.1.2 Estimação dos traços latentes

O processo de estimação dos traços latentes (habilidades) pelo método da máxima verossimilhança é realizado considerando que os parâmetros dos itens são conhecidos, ou seja, as estimativas dos parâmetros dos itens obtidos na primeira etapa, $\hat{\xi}$, são introduzidos na verossimilhança original $L(\xi, \theta)$, obtendo-se, então, uma verossimilhança perfilada, $L(\hat{\xi}, \theta)$, a qual possui propriedades interessantes que a fazem parecer com uma verossimilhança verdadeira. Assim, considera-se que $L(\hat{\xi}, \theta) \equiv L(\theta)$.

Para um determinado indivíduo, e considerando (S1) e (S2), a função de verossimilhança é dada por:

$$L(\theta_j) = \prod_{i=1}^I \prod_{k=0}^{m_i} P_{ijk}^{y_{ijk}},$$

e a log-verossimilhança dada por

$$l(\theta_j) = \sum_{i=1}^I \sum_{k=0}^{m_i} y_{ijk} \ln P_{ijk}. \quad (4.6)$$

A partir de (4.6) obtém-se o seguinte vetor escore

$$\begin{aligned} S(\theta_j) &= \sum_{i=1}^I \sum_{k=0}^{m_i} \frac{y_{ijk}}{P_{ijk}} \frac{\partial P_{ijk}}{\partial \theta_j} \\ &= D \sum_{i=1}^I \sum_{k=0}^{m_i} \frac{a_i y_{ijk} (W_{ijk}^+ - W_{ij(k+1)}^+)}{P_{ijk}}. \end{aligned} \quad (4.7)$$

É possível obter a matriz Hessiana e a informação de Fisher a partir da equação (4.7), dadas respectivamente por

$$J(\theta_j) = D^2 \sum_{i=1}^I \sum_{k=0}^{m_i} a_i^2 y_{ijk} \left[\frac{\left((1 - 2P_{ijk}^+) W_{ijk}^+ - (1 - 2P_{ij(k+1)}^+) W_{ij(k+1)}^+ \right)}{P_{ijk}} - \frac{\left(W_{ijk}^+ - W_{ij(k+1)}^+ \right)^2}{P_{ijk}^2} \right] \quad (4.8)$$

e

$$I(\theta_j) = D^2 a_i^2 \sum_{j=1}^n \sum_{k=0}^{m_i} \frac{\left(W_{ijk}^+ - W_{ij(k+1)}^+ \right)^2}{P_{ijk}}. \quad (4.9)$$

Portanto os métodos iterativos Newton-Raphson ou escore de Fisher descritos em (4.4) e (4.5) podem ser empregados.

É possível obter estimativas dos parâmetros dos itens e dos traços latentes utilizando o pacote ltm (RIZOPOULOS, 2006) no *software* R (R Core Team, 2018). Nesse pacote as estimativas dos parâmetros são obtidas sob o método da máxima verossimilhança marginal, utilizando a regra de quadratura *Gauss-Hermite*, como apresentado nesta seção. No pacote ltm também é possível controlar, por exemplo, o número de pontos de quadratura *Gauss-Hermite* e o método de otimização a ser utilizado. Além disso, o pacote fornece análises utilizando diversos modelos da teoria de resposta ao item, além do modelo de resposta gradual, como o modelo Rasch, o modelo logístico de 2 e 3 parâmetros, entre outros.

4.2 Estimação Bayesiana

Na abordagem Bayesiana especifica-se uma distribuição a priori a respeito de uma quantidade desconhecida. Cada parâmetro possui uma distribuição a priori, que expressa o conhecimento que se tem sobre os parâmetros desconhecidos antes de observar a amostra. A informação obtida pelos dados é representada pela função de verossimilhança. É através da função de verossimilhança que a informação dos dados modifica o conhecimento prévio dos parâmetros. Assim, na inferência Bayesiana, é necessário especificar uma distribuição a priori para os parâmetros, especificar um modelo para os dados, e após observar os dados, a informação a priori é combinada com a informação da amostra obtendo-se a distribuição a posteriori (EHLERS, 2011).

A abordagem Bayesiana baseia-se no teorema de Bayes, dado por

$$P(\alpha|\mathbf{x}) = \frac{P(\mathbf{x}, \alpha)}{P(\mathbf{x})} = \frac{P(\mathbf{x}|\alpha)P(\alpha)}{P(\mathbf{x})} = \frac{P(\mathbf{x}|\alpha)P(\alpha)}{\int_{\alpha} P(\mathbf{x}|\alpha)P(\alpha)d\alpha}, \quad (4.10)$$

em que $\mathbf{x} = (x_1, \dots, x_n)^\top$ representa a amostra observada, α representa o parâmetro de interesse no espaço paramétrico A , e $P(\alpha|\mathbf{x})$, $P(\mathbf{x}|\alpha)$, $P(\alpha)$ e $P(\mathbf{x})$ representam, respectivamente, a distribuição a posteriori, a função de verossimilhança, a distribuição a priori e uma constante, denominada constante normalizadora de $P(\alpha|\mathbf{x})$, que não depende do parâmetro α . $P(\mathbf{x})$ representa a probabilidade marginal dos dados ou ainda a distribuição preditiva a priori, e em alguns casos é possível determinar a distribuição a posteriori sem a necessidade de calcular diretamente o valor da constante $P(\mathbf{x})$. Assim, a forma usual para a densidade a posteriori para os parâmetros é

$$P(\alpha|\mathbf{x}) \propto P(\mathbf{x}|\alpha)P(\alpha), \quad (4.11)$$

que é denominado núcleo da densidade a posteriori.

Em palavras, a distribuição a posteriori é proporcional a verossimilhança multiplicada pela informação a priori. Além disso, inferências sobre parâmetros dos modelos podem ser

obtidas através do cálculo da média, mediana e moda a posteriori. Neste trabalho consideramos a média a posteriori, em que a esperança a posteriori é dada por

$$E(g(\alpha)|\mathbf{x}) = \int_A g(\alpha)P(\alpha|\mathbf{x})d\alpha. \quad (4.12)$$

Introduções mais detalhadas e outros conceitos da inferência Bayesiana podem ser vistos em [Bernardo e Smith \(2000\)](#), [Gamerman e Migon \(1993\)](#), [Berger \(1985\)](#), [Box e Tiao \(1973\)](#) e [Cornfield \(1969\)](#), por exemplo.

Em muitos casos são necessários cálculos complexos de integrais, ou seja, que não existe uma forma analítica fechada. Assim, é preciso utilizar métodos de aproximação com base em simulação para obtenção dessas integrais. Alguns métodos de aproximação são Monte Carlo simples, Monte Carlo com função de importância, métodos de reamostragem, método de Laplace e Monte Carlo via cadeias de Markov (MCMC), e podem ser vistos em [Evans e Swartz \(1995\)](#), [Carlin e Louis \(2001\)](#), [Gelman *et al.* \(2004\)](#) e [O'Hagan e Forster \(2004\)](#), por exemplo.

No caso de modelos da TRI, a estimação Bayesiana pode ser feita pelo método de estimação Bayesiana marginal utilizando a moda a posteriori, proposto por [Mislevy \(1986\)](#), que é uma extensão do método da máxima verossimilhança marginal, descrito na seção 4.1. A estimação Bayesiana marginal baseia-se no mesmo processo de construção da verossimilhança marginal, acoplando-a distribuições a priori para cada parâmetro do item no modelo, de forma que utilizar a informação dos dados (verossimilhança) e a informação a priori possibilita obter estimativas a posteriori dos parâmetros dos itens. Uma outra abordagem para estimação Bayesiana dos parâmetros é a utilização dos métodos MCMC, em que as vezes é possível estimar os parâmetros em modelos complexos, em situações com amostra pequena, e em outras aplicações que os algoritmos de máxima verossimilhança marginal não são aplicáveis ([WOLLACK *et al.*, 2002](#)). Nessa abordagem, obtém-se amostras a partir da distribuição a posteriori, em seguida pode-se usar a média a posteriori para estimar os parâmetros do item e os parâmetros da habilidade.

Um conjunto de algoritmos MCMC foram desenvolvidos. Um deles foi proposto por [Metropolis *et al.* \(1953\)](#) e generalizado por [Hastings \(1970\)](#), e por décadas não foi possível utilizá-lo em modelos com um grande número de parâmetros, como os modelos da TRI, por não ser computacionalmente viável ([CHANG, 2017](#)). O avanço computacional e o desenvolvimento de algoritmos MCMC tornou possível a aplicação em diferentes modelos da TRI.

Uma breve introdução dos métodos de Monte Carlo via Cadeias de Markov é apresentada no Apêndice A. Nesse Apêndice são descritos também os algoritmos MCMC Metropolis-Hastings e o amostrador de Gibbs.

Para o modelo de resposta gradual, a função de verossimilhança pode ser escrita como

$$L(\mathbf{Y}_{\dots}|\boldsymbol{\theta}, \boldsymbol{\xi}) = \prod_{j=1}^n \prod_{i=1}^I \prod_{k=1}^{m_i} [P_{ijk}]^{w_{ijk}},$$

em que \mathbf{Y}_{\dots} , $\boldsymbol{\theta}$, $\boldsymbol{\xi}$ e P_{ijk} são definidos na seção 3.2.1 e $w_{ijk} = 1$ se a k -ésima categoria do item i é selecionada pelo j -ésimo indivíduo, e $w_{ijk} = 0$ caso contrário.

Para especificar a distribuição a priori conjunta dos parâmetros do item e do indivíduo, assume-se uma estrutura de prioris independentes, isto é,

$$P(\boldsymbol{\theta}, \boldsymbol{\xi}) = \prod_{j=1}^n P(\theta_j) \prod_{i=1}^I \left[P(a_i) \prod_{k=1}^{m_i} P(b_{ik}) \right].$$

Em muitos casos, os pesquisadores utilizam a distribuição normal padrão como distribuição a priori da variável latente θ_j , $j = 1, \dots, n$, mas para os parâmetros do item a_i e b_{ik} , $i = 1, \dots, I$ e $k = 1, \dots, m_i$, diferentes prioris vêm sendo utilizadas (SILVA; BAZÁN; HUGGINS-MANLEY, 2017). Silva, Bazán e Huggins-Manley (2017) fazem uma análise de sensibilidade de diferentes prioris propostas na literatura para os parâmetros do item do modelo de resposta gradual, que são as prioris propostas por Zhu e Stone (2012), Kieftenbeld e Natesan (2012) e Curtis (2010). Com base nessa análise de sensibilidade da priori, no método Bayesiano, considerou-se a distribuição Normal truncada $[0, +\infty)$ com média 1 e desvio padrão 2,5 como distribuição a priori do parâmetro de discriminação, a_i , $i = 1, \dots, I$ e a distribuição Normal com média 0 e desvio padrão 2,5 como a priori para os parâmetros de dificuldade, b_{ik} , $k = 1, \dots, m_i$, sujeito a restrição $b_{i1} \leq b_{i2} \leq \dots \leq b_{im_i}$, que foram propostas por Curtis (2010). Para os traços latentes dos indivíduos considerou-se uma distribuição a priori normal padrão.

Dessa forma, a distribuição a posteriori é dada por

$$\begin{aligned} P(\boldsymbol{\theta}, \boldsymbol{\xi}|\mathbf{Y}_{\dots}) &\propto L(\mathbf{Y}_{\dots}|\boldsymbol{\theta}, \boldsymbol{\xi})P(\boldsymbol{\theta}, \boldsymbol{\xi}) \\ &\propto \prod_{j=1}^n P(\theta_j) \left[\prod_{i=1}^I \prod_{k=1}^{m_i} [P_{ijk}]^{w_{ijk}} \right] \prod_{i=1}^I \left[P(a_i) \prod_{k=1}^{m_i} P(b_{ik}) \right]. \end{aligned} \quad (4.13)$$

A função de verossimilhança e as distribuições a priori e a posteriori do modelo de resposta gradual multiunidimensional podem ser obtidas de forma similar a abordagem unidimensional, levando-se em conta a multidimensionalidade do traço latente.

4.2.1 Implementação do MCMC

A implementação de métodos MCMC apresentados no Apêndice A pode ser feita através de *softwares* que já estão disponíveis e são acessíveis facilmente. Um programa, o *Bayesian inference using Gibbs sampling* (BUGS) (GILKS; THOMAS; SPIEGELHALTER, 1994), em português, inferência Bayesiana usando o amostrador de Gibbs, utiliza a inferência Bayesiana

usando MCMC com base no amostrador de Gibbs. BUGS é usado nos *softwares* WinBUGS (LUNN *et al.*, 2000), OpenBUGS (THOMAS *et al.*, 2006) e Just another Gibbs sampler (JAGS) (PLUMMER, 2003). O JAGS pode ser utilizado no *software* R (R Core Team, 2018) através dos pacotes *rjags* (PLUMMER, 2016) e *R2jags* (SU; YAJIMA, 2015), já o OpenBUGS e o WinBUGS são utilizados no R com os pacotes *R2OpenBUGS* (STURTZ; LIGGES; GELMAN, 2010) e *R2WinBUGS* (STURTZ; LIGGES; GELMAN, 2005), respectivamente. O JAGS possui algumas vantagens em relação aos outros membros da família BUGS, e a principal é que o JAGS está escrito em C++, uma linguagem de programação orientada a objeto, com uma interface simples de módulos que podem ser escritos pelo usuário para estender as suas funcionalidades. O OpenBUGS e o WinBUGS têm sua própria linguagem e é baseada em *Component Pascal*, uma linguagem de programação pouco difundida na comunidade.

O Stan (Stan Development Team, 2014) também é um *software* que pode ser utilizado para implementação do MCMC e utiliza o algoritmo No-U-Turn Sampler (NUTS) (HOFFMAN; GELMAN, 2014), uma variação do algoritmo Monte Carlo Hamiltoniano (DUANE *et al.*, 1987). O Stan pode ser utilizado no R através do pacote *rstan* (Stan Development Team, 2018).

Um ponto que deve ser levado em consideração é o esforço computacional. No caso da estimação do modelo de resposta gradual por métodos MCMC, o processo é mais custoso computacionalmente, e conseqüentemente é mais demorado em comparação com o método da máxima verossimilhança marginal.

No capítulo subsequente, será apresentado resultados de um estudo de simulação com o objetivo de comparar diferentes cenários para o modelo de resposta gradual unidimensional e multiunidimensional, considerando diferentes métodos de estimação dos parâmetros do modelo. Além disso, será apresentado também uma aplicação à dados da síndrome de *burnout* em enfermeiras.

RESULTADOS

5.1 Abordagem unidimensional

5.1.1 *Um estudo de simulação*

Um estudo de simulação foi realizado para comparar a performance dos métodos de máxima verossimilhança marginal (MVM) e amostrador de Gibbs, na estimação dos parâmetros do modelo de resposta gradual unidimensional.

A ideia é comparar, com base em estatísticas apropriadas, as estimativas dos parâmetros dos itens e dos indivíduos com os valores dos parâmetros gerados. As estatísticas utilizadas são a raiz quadrada do erro quadrático médio (REQM), o viés e a correlação, que serão descritos mais adiante.

Neste estudo, dois fatores são considerados: o tamanho do teste (8, 12, 16 e 22 itens) e o tamanho da amostra (75, 150, 300, 500, 750 e 1000 indivíduos), além dos 2 métodos de estimação. Assim, as combinações resultam em 48 diferentes condições (2 métodos de estimação \times 4 tamanhos de teste \times 6 tamanhos de amostra).

5.1.1.1 *Simulação dos dados*

Para cada uma das 24 combinações do tamanho do teste e do tamanho da amostra, 20 conjuntos de dados diferentes foram gerados, isto é, o número de repetição R é 20. O parâmetro do indivíduo, θ_j , $j = 1, \dots, n$, foi gerado de uma distribuição normal padrão, $N(0, 1)$. Assim, para determinado tamanho amostral e a cada repetição, um novo conjunto de parâmetros do indivíduo foi simulado. Os parâmetros de discriminação e dificuldade dos itens usados, respectivamente a_i , $i = 1, \dots, I$ e b_{ik} , $k = 1, \dots, 6$, foram estimativas dos parâmetros de discriminação e de dificuldade do modelo de resposta gradual unidimensional aplicados a dados reais de enfermeiras, que são os dados descritos no capítulo 2, considerando uma escala de 7 pontos.

Segundo Kieftenbeld e Natesan (2012), parâmetros de itens comuns entre repetições ajudam a minimizar os efeitos do erro aleatório nas estimativas dos parâmetros. Portanto, as respostas foram geradas utilizando os mesmos parâmetros dos itens em todas as repetições. Os valores podem ser observados na Tabela 5, em que um teste simulado de tamanho L contém os primeiros L itens.

Tabela 5 – Parâmetros dos itens usados na simulação dos dados na abordagem unidimensional.

Itens	a	b_1	b_2	b_3	b_4	b_5	b_6
1	1,87	-3,24	-0,60	0,42	1,76	2,49	4,69
2	2,25	-4,07	-1,89	-0,71	0,60	1,28	3,83
3	2,11	-3,21	-1,19	-0,07	1,22	1,96	4,21
4	0,28	0,11	1,58	2,09	3,14	3,93	4,99
5	1,12	-0,27	0,64	1,07	1,86	2,39	4,06
6	0,93	0,28	1,62	2,43	3,42	4,38	5,69
7	0,54	0,16	1,81	2,20	3,09	3,88	4,46
8	2,58	-1,09	1,02	1,65	2,88	3,65	5,98
9	0,50	-1,48	0,12	0,62	1,42	2,04	2,83
10	0,98	0,50	1,48	1,96	2,95	3,49	4,56
11	0,94	-0,27	1,00	1,50	2,37	2,81	3,56
12	1,47	-2,10	0,21	1,05	2,25	3,32	4,43
13	1,75	-1,75	0,03	0,79	1,92	2,59	4,10
14	1,24	-2,61	-1,28	-0,59	0,34	0,88	2,31
15	0,74	1,10	2,30	2,85	3,58	4,03	4,95
16	1,21	-0,80	1,10	1,81	2,65	3,33	5,04
17	0,66	-0,72	1,17	1,73	2,50	3,11	3,99
18	0,29	-0,67	1,10	1,68	2,86	3,58	4,99
19	0,76	-1,18	0,76	1,43	2,53	3,36	4,51
20	2,24	-1,44	1,03	1,88	2,97	3,90	5,73
21	0,64	-1,03	0,54	1,04	1,92	2,85	4,98
22	0,70	-0,46	0,81	1,43	2,19	2,97	5,05

A partir dos valores simulados dos parâmetros dos indivíduos e dos itens, calcula-se a probabilidade do indivíduo responder, em cada item, a determinada categoria ou outra maior, de acordo com o modelo de resposta gradual unidimensional. Com base na probabilidade do indivíduo responder cada categoria, determina-se a resposta do indivíduo a cada item do teste, isto é, a resposta do indivíduo será designada como sendo a categoria com maior probabilidade de resposta de acordo com o modelo de resposta gradual unidimensional.

5.1.1.2 Estimação dos parâmetros

O pacote ltm do *software* R foi utilizado para obter estimativas dos parâmetros do modelo de resposta gradual sob o método da máxima verossimilhança marginal. Para obter as estimativas dos parâmetros pelo método Bayesiano amostrador de Gibbs, o JAGS e seus pacotes no *software*

R foram utilizados. Os códigos utilizados no JAGS para ajustar o modelo de resposta gradual unidimensional estão apresentados no Apêndice B.

No método Bayesiano, considerou-se a distribuição Normal truncada $[0, +\infty)$ com média 1 e desvio padrão 2,5 como distribuição a priori do parâmetro de discriminação, a_i , $i = 1, \dots, I$ e a distribuição Normal com média 0 e desvio padrão 2,5 como a priori para os parâmetros de dificuldade, b_{ik} , $k = 1, \dots, 6$, sujeito a restrição $b_{i1} \leq b_{i2} \leq \dots \leq b_{i6}$. Para os traços latentes dos indivíduos considerou-se uma distribuição a priori normal padrão. Essas prioris foram escolhidas com base na análise de sensibilidade da priori realizada em [Silva, Bazán e Huggins-Manley \(2017\)](#). Além disso, a cada repetição os valores iniciais dos parâmetros do i -ésimo item foram fixados em $a_i = 1,0$, $b_{i1} = -1,5$, $b_{i2} = -1,0$, $b_{i3} = -0,5$, $b_{i4} = 0,5$, $b_{i5} = 1,0$ e $b_{i6} = 1,5$. As primeiras 1000 iterações foram consideradas o período de aquecimento da cadeia (*burn-in*), e 10.000 iterações foram utilizadas para estimar os parâmetros.

Para comparar a performance dos modelos nos diferentes cenários apresentados considera-se a estatística REQM, definida como

$$\text{REQM}_r(\pi) = \sqrt{\frac{1}{L} \sum_{i=1}^L (\hat{\pi}_{ir} - \pi_i)^2},$$

em que $\hat{\pi}_{ir}$ é a estimativa do parâmetro π_i no item i na repetição r , $r = 1, \dots, R$, com $R = 20$. Neste caso, π representa os parâmetros do item, isto é, o parâmetro de discriminação a ou um dos parâmetros de dificuldade b_1, \dots, b_6 . Após obter os R REQM, calcula-se o seu valor médio para cada parâmetro. Outra estatística considerada é o viés, dado por

$$\text{vies}(\pi_i) = \hat{\pi}_i - \pi_i,$$

em que $\hat{\pi}_i = \frac{1}{R} \sum_{r=1}^R \hat{\pi}_{ir}$.

No caso do parâmetro do indivíduo θ , o REQM é calculado como

$$\text{REQM}_r(\theta) = \sqrt{\frac{1}{n} \sum_{j=1}^n (\hat{\theta}_{jr} - \theta_{jr})^2},$$

em que $\hat{\theta}_{jr}$ e θ_{jr} representam, respectivamente, o valor estimado e simulado do j -ésimo indivíduo na repetição r , $r = 1, \dots, R$, com $R = 20$. Após obter os R REQM, calcula-se o seu valor médio. O viés para o parâmetro do indivíduo é dado por

$$\text{vies}(\theta_j) = \frac{1}{R} \sum_{r=1}^R (\hat{\theta}_{jr} - \theta_{jr}).$$

Considera-se também a correlação média entre os valores estimados e seus correspondentes valores originais, calculada dentro de cada repetição para cada tipo de parâmetro $(a, b_1, \dots, b_6, \theta)$.

Os resultados do estudo de simulação estão apresentados nas Tabelas 12, 13 e 14 no Apêndice C. A convergência das cadeias dos resultados simulados pelos métodos de MCMC foram checadas através da estatística \hat{R} , do fator de redução de escala potencial, de Gelman e Rubin (1992) e do tamanho efetivo da amostra em cada repetição. Os resultados indicam que as cadeias alcançam o equilíbrio satisfatoriamente.

Em relação aos parâmetros do item, de uma maneira geral, nota-se que o método da máxima verossimilhança marginal produz os melhores resultados na estimação do parâmetro de discriminação a se levado em consideração o viés médio das estimativas, porém, quando o REQM médio é observado, em geral, o amostrador de Gibbs apresenta os melhores resultados para este parâmetro. Já as estimativas dos parâmetros de dificuldade b_1, b_2, b_3, b_4, b_5 e b_6 , apresentam, em geral, os melhores resultados do REQM e do viés no amostrador de Gibbs. É possível observar, entretanto, que o viés médio do teste com 22 itens considerando o método da máxima verossimilhança marginal possui melhores resultados do que o amostrador de Gibbs, em geral.

Na Tabela 12 é possível observar que os REQM's médios dos parâmetros de discriminação, a , diminuem à medida que o tamanho da amostra e do teste aumentam quando considerado o método amostrador de Gibbs. Nos resultados obtidos para a amostra de tamanho 75, por exemplo, nota-se que o amostrador de Gibbs possui REQM's médios de 0,39 no teste com 16 itens. Em amostras maiores, este valor diminui, sendo igual a 0,12 quando a amostra é de 1000 observações para o teste com a mesma quantidade de itens.

O viés médio, apresentado na tabela 13, indica que o método da máxima verossimilhança marginal produz os melhores resultados para o parâmetro a . Por exemplo, para a amostra de tamanho 300, o viés médio é de $-0,02$ para o teste com 12 itens pelo método da máxima verossimilhança marginal e de $0,15$ pelo método MCMC amostrador de Gibbs. Além disso, o método da máxima verossimilhança marginal em geral possui o menor viés médio para testes com a menor quantidade de itens, isto é, à medida que o tamanho do teste aumenta, o valor do viés médio aumenta, em geral.

É possível notar nas Tabelas 12 e 13 que o REQM médio e o viés médio dos parâmetros b_1, b_2, b_3, b_4, b_5 e b_6 em geral são menores pelo método MCMC via amostrador de Gibbs. Por exemplo, na amostra de tamanho 500 e no teste de tamanho 8 o REQM médio do parâmetro b_5 é $0,55$ pelo amostrador de Gibbs e $2,90$ pelo método da máxima verossimilhança marginal. Entretanto, quando o viés médio é levado em consideração, observa-se que a performance do método da máxima verossimilhança marginal na estimação dos parâmetros de dificuldade, em geral, é superior, quando observado o tamanho do teste com 22 itens. Por exemplo, na amostra de tamanho 150 e para o parâmetro b_3 , o viés médio pelo amostrador de Gibbs é $-0,54$, enquanto

para o método da máxima verossimilhança marginal é $-0,04$.

A correlação média, apresentada na Tabela 14, indica que, em geral, à medida que o tamanho amostral aumenta, a correlação média aumenta. O método MCMC via amostrador de Gibbs apresenta correlação média alta em grande parte dos parâmetros de dificuldade, e a correlação média no parâmetro de discriminação é grande, principalmente em grandes amostras. A correlação média nos parâmetros de dificuldade pelo método da máxima verossimilhança marginal em geral são inferiores ao outro método. Na amostra de tamanho 750 e teste de tamanho 16, por exemplo, a correlação média no parâmetro b_4 é de 0,25 pelo método da máxima verossimilhança marginal e 0,96 pelo amostrador de Gibbs.

Os resultados do parâmetro θ em geral apresentam melhores resultados para o amostrador de Gibbs. Em grandes amostras, os resultados dos REQM's dos métodos clássico e Bayesiano se aproximam, mas os resultados do amostrador de Gibbs ainda são um pouco superiores. Por exemplo, na amostra de tamanho 1000 e tamanho do teste igual a 12 itens, o REQM médio do parâmetro θ é 0,40 no método da máxima verossimilhança marginal e 0,38 pelo amostrador de Gibbs. As correlações médias nesse parâmetro, em geral, também são próximas e altas em ambos os métodos.

5.1.2 Uma aplicação aos dados de enfermeiras

5.1.2.1 Amostragem dos dados

Devido ao custo computacional e ao tempo, não é possível utilizar na análise da TRI informações de todas as enfermeiras que participaram do projeto RN4CAST. Assim, será considerado apenas enfermeiras que trabalham na Bélgica. Nesse país, 2643 enfermeiras responderam todos os itens do MBI, de um total de 3186 enfermeiras. Com base no estudo de simulação, nota-se que trabalhar com uma amostra de tamanho 500 é razoável, uma vez que trabalhar com uma amostra maior do que essa pode não compensar por conta de custos adicionais, de acordo com Kieftenbeld e Natesan (2012). Por esse motivo, uma amostra com 498 enfermeiras da Bélgica será utilizada para ajustar o modelo de resposta gradual.

O processo de escolha das 498 enfermeiras se dá com base na amostragem estratificada, que consiste na divisão de uma população em grupos (chamados estratos), e pode ser visto mais detalhadamente em Bolfarine e Wilton (2005). Inicialmente, considera-se que os hospitais da Bélgica são os estratos, e dentro de cada hospital, as unidades de enfermagem são estratos também. A ideia de utilizar esse tipo de amostragem é que a amostra contenha proporcionalmente o número de enfermeiras nos hospitais e nas unidades de enfermagem da população das enfermeiras da Bélgica que participaram do projeto RN4CAST.

A alocação da amostra pelos estratos será realizada utilizando a alocação proporcional, isto é, a amostra de tamanho $n = 498$ é distribuída proporcionalmente ao tamanho dos estratos

$$n_h = n \frac{N_h}{N}$$

em que N_h é o tamanho do estrato h , $h = 1, \dots, H$ e N é tamanho do universo, ou seja, $N = \sum_{h=1}^H N_h$.

Ao utilizar o processo de amostragem estratificada com alocação proporcional, obtém-se uma amostra representativa da população de enfermeiras da Bélgica do projeto RN4CAST, considerando a quantidade de enfermeiras nos hospitais e nos tipos de unidades de enfermagem. As Tabelas 6 e 7 são as frequências, em porcentagem, das respostas da população de enfermeiras da Bélgica e das enfermeiras da amostra considerada, respectivamente, para as categorias de cada item. Nota-se que, em geral, a proporção de enfermeiras que respondem as categorias dos itens são próximos na população de enfermeiras da Bélgica e na amostra considerada.

Nota-se na Tabela 7 que, em geral, em grande parte dos itens referentes a exaustão emocional e a despersonalização, as enfermeiras responderam com maior frequência as categorias 0 (nunca) e 1 (algumas vezes por ano ou menos). Por exemplo, no item 8 (“Sinto-me esgotado devido ao meu trabalho”) 34,3% das enfermeiras responderam a categoria 0 e 27,7% das enfermeiras responderam a categoria 1. Já nos itens referentes a realização pessoal reduzida, as maiores frequências de resposta são nas categorias 5 (algumas vezes por semana) e 6 (todos os dias). Por exemplo, no item 4 (“Consigo compreender facilmente como os meus pacientes se sentem acerca das coisas”), 53,8% das enfermeiras responderam a categoria 6 e 27,5% responderam a categoria 5. Entre os itens referentes a exaustão emocional e a despersonalização, a maior frequência de resposta da categoria 6 é no item 14 (“Sinto que estou a trabalhar demasiado no meu trabalho”), sendo de 13,5% das enfermeiras. O item 2 (“No fim do dia de trabalho, sinto-me exausto”), referente a exaustão emocional possui 26,1% de respostas na categoria 5. O item 15 (“De fato, não me interessa o que acontece a alguns pacientes”), referente a despersonalização, destaca-se pelo fato de 75,1% das enfermeiras responderem a categoria 0.

A síndrome de *burnout* é indicada por maiores pontuações da exaustão emocional e despersonalização e menores pontuações da realização pessoal reduzida. Aqui, para a modelagem dos dados com o modelo de resposta gradual, o código para a realização pessoal reduzida será invertido para propósitos interpretativos (a indicação de *burnout* para as três medidas é um valor grande), isto é, para os itens referentes a realização pessoal reduzida, a escala *Likert* é dada por 0 – todos os dias; 1 – algumas vezes por semana; 2 – uma vez por semana; 3 – algumas vezes por mês; 4 – uma vez por mês; 5 – algumas vezes por ano ou menos; 6 – nunca.

5.1.2.2 Modelagem dos dados

Nesta seção será apresentada uma aplicação do modelo de resposta gradual aos dados de enfermeira via algoritmo amostrador de Gibbs e via o método da máxima verossimilhança marginal. Considerou-se $n = 498$ enfermeiras que trabalham em hospitais da Bélgica provenientes

Tabela 6 – Frequências (em percentagem) de resposta da população de enfermeiras da Bélgica para as categorias de respostas por item.

	Categorias						
	0	1	2	3	4	5	6
Item 1	10,7	28,0	15,7	20,9	7,9	12,6	4,2
Item 2	7,8	16,7	13,1	19,7	9,5	23,2	10,0
Item 3	11,6	20,9	15,3	19,4	9,5	16,7	6,5
Item 4	0,8	1,5	1,8	6,7	5,7	30,8	52,7
Item 5	44,2	18,0	7,6	10,7	5,5	9,9	4,1
Item 6	54,5	25,6	8,2	6,1	3,0	1,8	0,8
Item 7	1,1	1,4	2,2	4,6	5,0	31,4	54,3
Item 8	34,1	27,5	8,8	11,1	5,2	9,7	3,6
Item 9	7,2	6,8	6,1	15,9	11,6	31,6	20,8
Item 10	59,8	18,2	6,5	7,0	3,1	3,7	1,7
Item 11	43,2	24,5	9,1	9,3	4,0	5,0	4,9
Item 12	3,3	5,5	8,5	17,4	12,6	35,6	17,1
Item 13	24,3	26,4	12,4	16,0	6,9	9,6	4,4
Item 14	9,6	16,5	10,8	19,6	8,9	20,5	14,1
Item 15	74,2	15,4	3,7	3,0	1,3	1,3	1,1
Item 16	35,1	33,7	11,2	9,6	4,1	4,8	1,5
Item 17	1,8	2,8	4,0	8,7	9,1	40,5	33,1
Item 18	0,6	2,1	3,0	11,1	9,8	40,3	33,1
Item 19	2,0	2,6	5,2	14,0	12,4	38,6	25,2
Item 20	29,7	34,4	10,4	11,0	5,6	6,4	2,5
Item 21	1,3	5,0	8,1	15,0	11,4	32,8	26,4
Item 22	40,2	29,6	9,3	9,4	5,0	5,1	1,4

tes do projeto RN4CAST, como apresentado no Capítulo 2. A estimação dos parâmetros via o algoritmo amostrador de Gibbs foi realizada utilizando os pacotes *rjags* e *R2jags* no *software* R e a estimação pelo método da máxima verossimilhança marginal via a regra de quadratura *Gauss-Hermite* foi realizada utilizando o pacote *ltm*, também no *software* R.

Para cada grupo de itens das dimensões de *burnout*, considerou-se um modelo de resposta gradual unidimensional. Assim, a exaustão emocional, a despersonalização e a realização pessoal reduzida das $n = 498$ enfermeiras são estimadas com base nas respostas das enfermeiras a cada grupo de itens do MBI, sendo, respectivamente, 9 itens do MBI (itens 1, 2, 3, 6, 8, 13, 14, 16 e 20), 5 itens do MBI (itens 5, 10, 11, 15 e 22) e 8 itens do MBI (itens 4, 7, 9, 12, 17, 18, 19 e 21). Isto é, para cada dimensão de *burnout*, considerou-se um modelo de resposta gradual unidimensional e estimou-se o traço latente das enfermeiras e os parâmetros dos respectivos itens desta dimensão.

Em termos da síndrome de *burnout*, agora Y_{ij} representa a resposta da j -ésima enfermeira no i -ésimo item no MBI. Cada item i , $i = 1, \dots, I$, em que I pode ser 9, 5 ou 8, tem $m_i = 7$

Tabela 7 – Frequências (em porcentagem) de resposta da amostra de 498 enfermeiras da Bélgica para as categorias de respostas por item.

	Categorias						
	0	1	2	3	4	5	6
Item 1	11,6	27,1	17,7	18,1	8,8	13,1	3,6
Item 2	9,0	14,3	10,7	23,3	8,8	26,1	7,8
Item 3	11,6	19,9	13,9	20,7	12,0	15,7	6,2
Item 4	1,2	1,2	1,4	5,6	9,3	27,5	53,8
Item 5	45,6	17,5	7,0	11,9	4,4	9,4	4,2
Item 6	54,2	25,1	7,7	8,6	3,2	0,8	0,4
Item 7	1,2	1,6	1,4	4,6	4,6	34,3	52,2
Item 8	34,3	27,7	9,0	11,5	4,6	10,1	2,8
Item 9	8,9	5,8	9,2	16,7	11,0	26,5	21,9
Item 10	60,5	18,5	8,2	7,6	2,2	1,6	1,4
Item 11	43,6	24,7	10,7	8,2	3,0	3,2	6,6
Item 12	3,6	5,0	8,4	17,1	13,7	33,9	18,3
Item 13	23,5	26,9	12,3	16,7	6,2	9,6	4,8
Item 14	10,4	15,3	12,4	19,5	7,4	21,5	13,5
Item 15	75,1	13,5	5,4	3,0	1,4	0,6	1,0
Item 16	36,4	32,5	12,7	8,6	3,4	4,6	1,8
Item 17	3,2	2,8	3,8	9,3	6,4	38,6	35,9
Item 18	0,6	2,0	2,8	11,9	10,4	39,4	32,9
Item 19	2,8	1,6	5,0	16,7	12,9	34,3	26,7
Item 20	30,1	32,9	11,3	11,9	5,2	6,6	2,0
Item 21	0,8	6,0	9,0	15,3	9,4	34,8	24,7
Item 22	41,4	31,1	6,8	10,5	5,0	3,6	1,6

categorias (0 a 6), mas considerou-se que $Y_{ij} = 1, \dots, 7$ (a categoria 0 é representada por $Y_{ij} = 1$, a categoria 1 é representada por $Y_{ij} = 2$, e assim por diante) e $j = 1, \dots, n$ é o número de enfermeiras, e nesse caso $n = 498$.

Na abordagem Bayesiana empregando o algoritmo amostrador de Gibbs, as prioris dos parâmetros do item e do traço latente do indivíduo são as mesmas usadas no estudo de simulação, isto é, $\theta_j \sim \text{Normal}(0, 1)$ para $j = 1, \dots, n$, $a_i \sim N(1, (2,5)^2)T[0, +\infty)$, $b_{ik} \sim \text{Normal}(0; (2,5)^2)$ e $b_{i(k+1)} \sim \text{Normal}(0; 2^2)I(b_{ik})$ para $i = 1, \dots, I$ e $k = 0, \dots, m_i$, em que $I(b_{ik})$ representa que necessariamente $b_{i1} \leq b_{i2} \leq \dots \leq b_{im_i}$. Nos modelos considerados para cada dimensão de *burnout*, utilizou-se 3 cadeias, cada uma com 100.000 iterações e com as primeiras 10.000 iterações descartadas (*burn in*).

A Tabela 8 apresenta as estimavas dos parâmetros dos itens no modelo de resposta gradual unidimensional via algoritmo amostrador de Gibbs considerando os grupos de itens de cada dimensão de *burnout*. Nota-se que a maioria das estimativas dos parâmetros de discriminação dos itens, a , é maior do que 1, exceto os itens 6 (“Trabalhar com pessoas o dia todo é, de

fato, um esforço para mim”) e o item 22 (“Sinto que os pacientes me culpam por alguns dos seus problemas”), em que $\hat{a} = 0,84$ e $\hat{a} = 0,78$, respectivamente. Isso indica que, nesses itens, enfermeiras de diferentes níveis de exaustão emocional e despersonalização, respectivamente, não são bem discriminadas em relação a probabilidade de responder determinada categoria quanto nos outros itens. O item com maior valor na estimativa do parâmetro de discriminação é o item 10 (“Tornei-me mais insensível em relação às pessoas, desde que comecei este trabalho”), em que $\hat{a} = 2,87$, relacionado com a despersonalização da enfermeira.

Ainda é possível notar na Tabela 8 que os maiores valores das estimativas médias dos parâmetros de dificuldade, $b_{\text{média}}$, são os itens 6 (“Trabalhar com pessoas o dia todo é, de fato, um esforço para mim”), em que $b_{\text{média}}$ é 3,18, relacionado à exaustão emocional, 15 (“De fato, não me interessa o que acontece a alguns pacientes”) e 22 (“Sinto que alguns pacientes me culpam por alguns dos meus problemas”), com $b_{\text{média}}$ iguais a 2,80 e 2,49 respectivamente, relacionados à despersonalização, e 4 (“Consigo compreender facilmente como os meus pacientes se sentem a cerca das coisas”) e 7 (“Lido muito eficazmente com os problemas dos meus pacientes”), em que $b_{\text{média}}$ são iguais a 2,54 e 2,35 respectivamente, relacionados à realização pessoal reduzida. Os itens que possuem as menores dificuldades médias são 14 (“Sinto que estou a trabalhar demasiado no meu emprego”), 2 (“No fim do dia de trabalho, sinto-me exausto”) e 3 (“Sinto-me fatigado quando acordo de manhã e tenho que enfrentar mais um dia de trabalho”), como $b_{\text{média}}$ iguais a -0,12, -0,10 e 0,17 respectivamente, que estão relacionados com a exaustão emocional. Dentre os itens relacionados com a despersonalização, o item 11 (“Preocupo-me que este trabalho me esteja a “endurecer” emocionalmente.”) é o que possui menor valor da média dos parâmetros de dificuldade, com $b_{\text{média}} = 1,29$. O item 9 (“Sinto que estou a influenciar positivamente a vida de outras pessoas com o meu trabalho”), que está relacionado com a realização pessoal reduzida, é o que possui menor $b_{\text{média}}$, sendo igual 1,00.

Vincenzi *et al.* (2018) apresenta uma interpretação das médias dos parâmetros de dificuldade, em que o itens que possuem altos valores de $b_{\text{média}}$ requerem um nível alto do traço latente para ser respondido satisfatoriamente à medida que ocorre uma das dimensões da síndrome de *burnout*, neste caso. Já os itens que possuem baixos valores de $b_{\text{média}}$, não requerem um nível muito alto do traço latente para que o indivíduo responda às maiores categorias.

As Figuras 8, 9, 10 e 11 apresentam as curvas características das categorias dos itens 2, 6, 9 e 22, respectivamente. As curvas características das categorias do item 2, que possui um valor alto da estimativa do parâmetro de discriminação, mostram que enfermeiras de diferentes valores do traço latente, neste caso, a exaustão emocional, possuem diferentes probabilidades de responder determinada categoria, e por isso são bem discriminadas. Por exemplo, enfermeiras que possuem o valor do traço latente entre -4 e -2 a maior probabilidade é de responder a categoria 1 (nunca), e enfermeiras com valor do traço latente de -2 a -1, aproximadamente, possui alta probabilidade de responder a categoria 2 (algumas vezes por ano ou menos), e assim por diante, ao passo que, enfermeiras com valores altos do traço latente, entre 2 e 4, possuem

Tabela 8 – Estimativas dos parâmetros dos itens do modelo de resposta gradual unidimensional nos dados do MBI via amostrador Gibbs.

Itens	a	b_1	b_2	b_3	b_4	b_5	b_6	b_{media}
Exaustão Emocional								
1	2,11	-1,58	-0,42	0,14	0,79	1,21	2,36	0,42
2	2,86	-1,61	-0,87	-0,49	0,20	0,48	1,68	-0,10
3	2,39	-1,51	-0,63	-0,18	0,48	0,93	1,91	0,17
6	0,84	0,11	1,61	2,33	3,73	5,09	6,20	3,18
8	2,51	-0,52	0,29	0,59	1,07	1,34	2,39	0,86
13	1,78	-1,02	-0,02	0,42	1,11	1,47	2,39	0,73
14	1,52	-1,93	-0,97	-0,43	0,32	0,61	1,69	-0,12
16	1,25	-0,65	0,72	1,41	2,11	2,55	3,73	1,65
20	2,35	-0,69	0,32	0,73	1,30	1,68	2,66	1,00
Despersonalização								
5	1,06	-0,24	0,58	0,95	1,70	2,07	3,40	1,41
10	2,87	0,27	0,90	1,29	1,89	2,22	2,63	1,53
11	1,55	-0,27	0,70	1,22	1,74	1,99	2,33	1,29
15	1,27	1,08	1,97	2,60	3,21	3,74	4,19	2,80
22	0,78	-0,56	1,31	1,85	2,97	3,92	5,42	2,49
Realização Pessoal Reduzida								
4	1,13	0,20	1,74	2,25	3,00	3,58	4,47	2,54
7	1,57	0,12	1,51	2,01	2,63	3,31	4,51	2,35
9	1,18	-1,29	0,17	0,75	1,57	2,11	2,71	1,00
12	1,24	-1,60	0,20	0,66	1,59	2,24	3,20	1,05
17	2,23	-0,52	0,77	1,20	1,81	2,16	2,70	1,35
18	2,07	-0,67	0,74	1,17	2,04	2,51	3,34	1,52
19	2,17	-0,86	0,41	0,89	1,56	2,11	2,58	1,12
21	1,65	-0,94	0,37	0,86	1,57	2,15	3,33	1,22

maior probabilidade de responder a categoria 7 (todos os dias). Já as curvas características das categorias dos itens 6 e 22, que são itens que possuem valores baixos da estimativa do parâmetro de discriminação, mostram que não há muita diferença na probabilidade de enfermeiras de diferentes traços latentes responderem uma determinada categoria. Por exemplo, no item 22, enfermeiras com traço latente, neste caso a despersonalização, entre -4 e 0 possuem maior probabilidade de responderem a categoria 1 (nunca), isto é, enfermeiras com diferentes níveis de despersonalização provavelmente respondem a mesma categoria neste item.

Como os itens 6 e 22, em média, possuem valores altos das estimativas dos parâmetros de dificuldade, as curvas características das categorias destes itens estão levemente deslocadas para a direita. Assim, provavelmente as enfermeiras que respondem as categorias maiores possuem também os maiores valores do traço latente. Os itens 2 e 9, que possuem, em média, valores baixos das estimativas dos parâmetros de dificuldade, apresentam as curvas características das

categorias deslocadas mais uniformemente no eixo da abscissa.

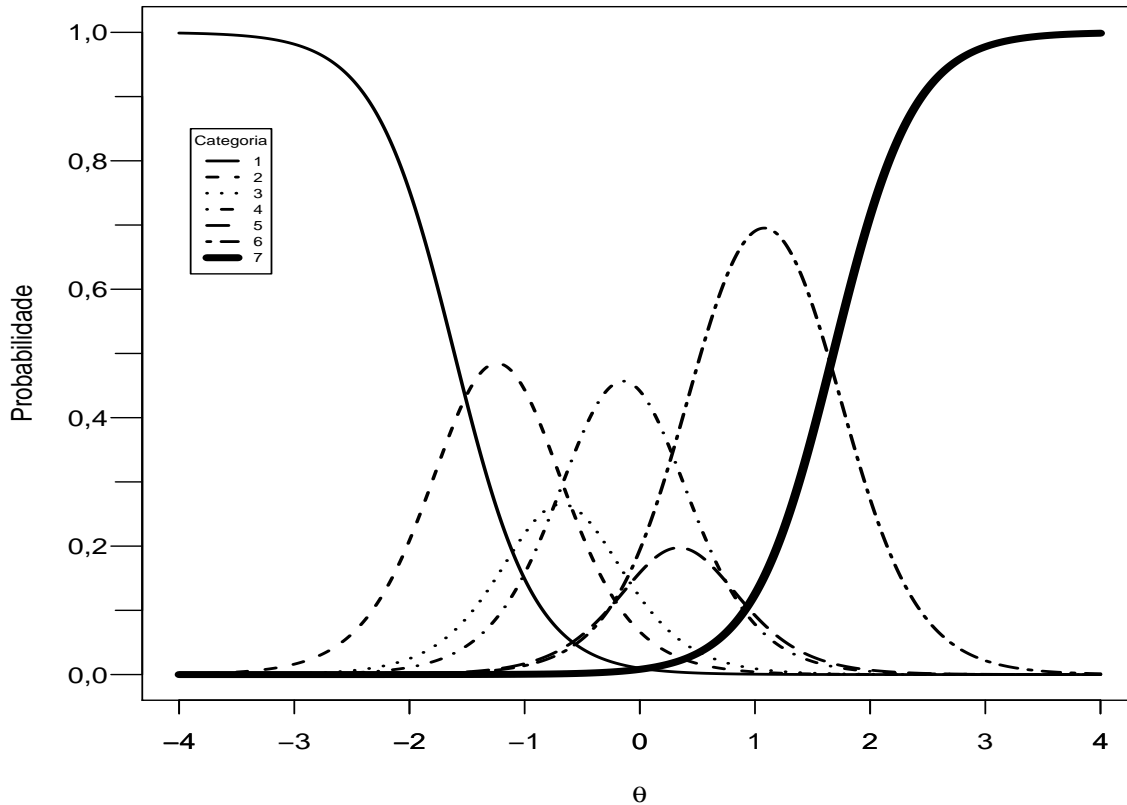


Figura 8 – Curvas características das categorias do item 2 - No fim do dia de trabalho, sinto-me exausto - relacionado a exaustão emocional da enfermeira. Neste item, $\hat{a} = 2,86$, $\hat{b}_1 = -1,61$, $\hat{b}_2 = -0,87$, $\hat{b}_3 = -0,49$, $\hat{b}_4 = 0,20$, $\hat{b}_5 = 0,48$ e $\hat{b}_6 = 1,68$.

Fonte: Elaborada pelo autor.

As Figuras 12, 13 e 14 são histogramas que mostram a frequência das estimativas dos traços latentes relacionados com a exaustão emocional, despersonalização e realização pessoal reduzida, respectivamente. Nota-se na Figura 12 que grande parte dos traços latentes da exaustão emocional das enfermeiras foi estimado entre $-0,75$ e $0,75$, e parece haver uma quantidade próxima de enfermeiras com traços latentes estimados de $-0,75$ a $-2,25$ e de $0,75$ a $2,25$. Poucas enfermeiras possuem o traço latente estimado entre $-2,25$ e $-2,75$ e entre $2,25$ e $3,75$.

O histograma da Figura 13 mostra que a maior parte das estimativas dos traços latentes relacionado a despersonalização das enfermeiras está entre $-1,25$ e $0,25$. Não houve enfermeiras com traço latente estimado menor do que $-1,25$ e nem maior do que $2,75$, existindo uma quantidade razoável de enfermeiras com traço latente estimado entre $0,25$ e $1,25$ e poucas enfermeiras com traço latente estimado entre $1,25$ e $2,75$.

No histograma referente ao traço latente relacionado com a realização pessoal reduzida, na Figura 14, nota-se que grande parte das enfermeiras possui traços latentes estimados entre

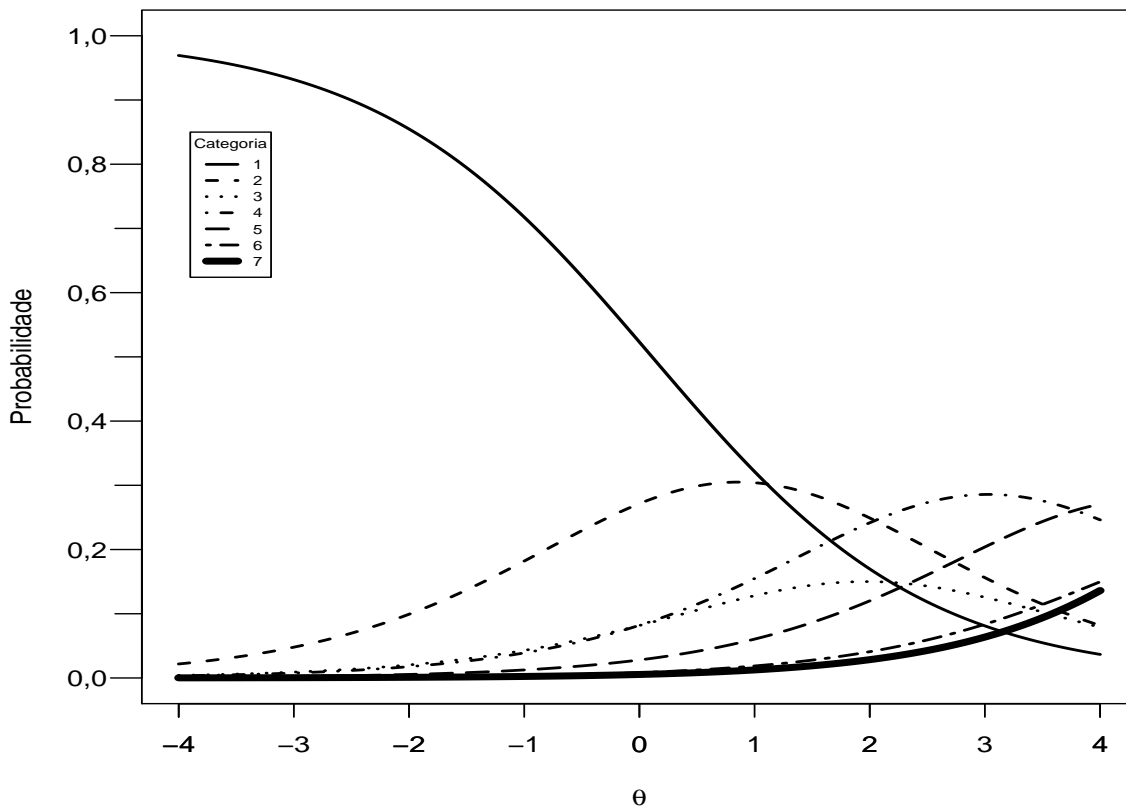


Figura 9 – Curvas características das categorias do item 6 - Trabalhar com pessoas o dia todo é, de fato, um esforço para mim - relacionado a exaustão emocional da enfermeira. Neste item, $\hat{a} = 0,84$, $\hat{b}_1 = 0,11$, $\hat{b}_2 = 1,61$, $\hat{b}_3 = 2,33$, $\hat{b}_4 = 3,73$, $\hat{b}_5 = 5,09$ e $\hat{b}_6 = 6,20$.

Fonte: Elaborada pelo autor.

$-0,75$ e $0,75$, uma quantidade razoável de enfermeiras possui traço latente estimado de $-0,75$ a $-2,25$ e entre $0,75$ e $1,75$, além disso, não existe enfermeira com traço latente estimado menor que $-2,25$ e poucas enfermeiras possui traço latente estimado entre $1,75$ e $3,75$.

A Tabela 9 apresenta as estimativas dos parâmetros dos itens no modelo de resposta gradual unidimensional via o método da máxima verossimilhança marginal considerando cada dimensão de *burnout*. Nota-se que o item 10 (“Tornei-me mais insensível em relação às pessoas, desde que comecei este trabalho”), relacionado a despersonalização e o item 2 (“No fim do dia de trabalho, sinto-me exausta”), relacionado a exaustão emocional possuem os maiores valores das estimativas dos parâmetros de discriminação, sendo $\hat{a} = 2,89$ e $\hat{a} = 2,88$, respectivamente. As menores estimativas dos parâmetros de discriminação são dos itens 6 (“Trabalhar com pessoas o dia todo é, de fato, um esforço para mim”), relacionado a exaustão emocional, e o item 22 (“Sinto que os pacientes me culpam por alguns dos seus problemas”), relacionado a despersonalização, sendo $\hat{a} = 0,62$ em ambos os itens. Já as estimativas dos parâmetros de dificuldade que mais chamam atenção são as do item 6, relacionado com a exaustão emocional, uma vez que possuem

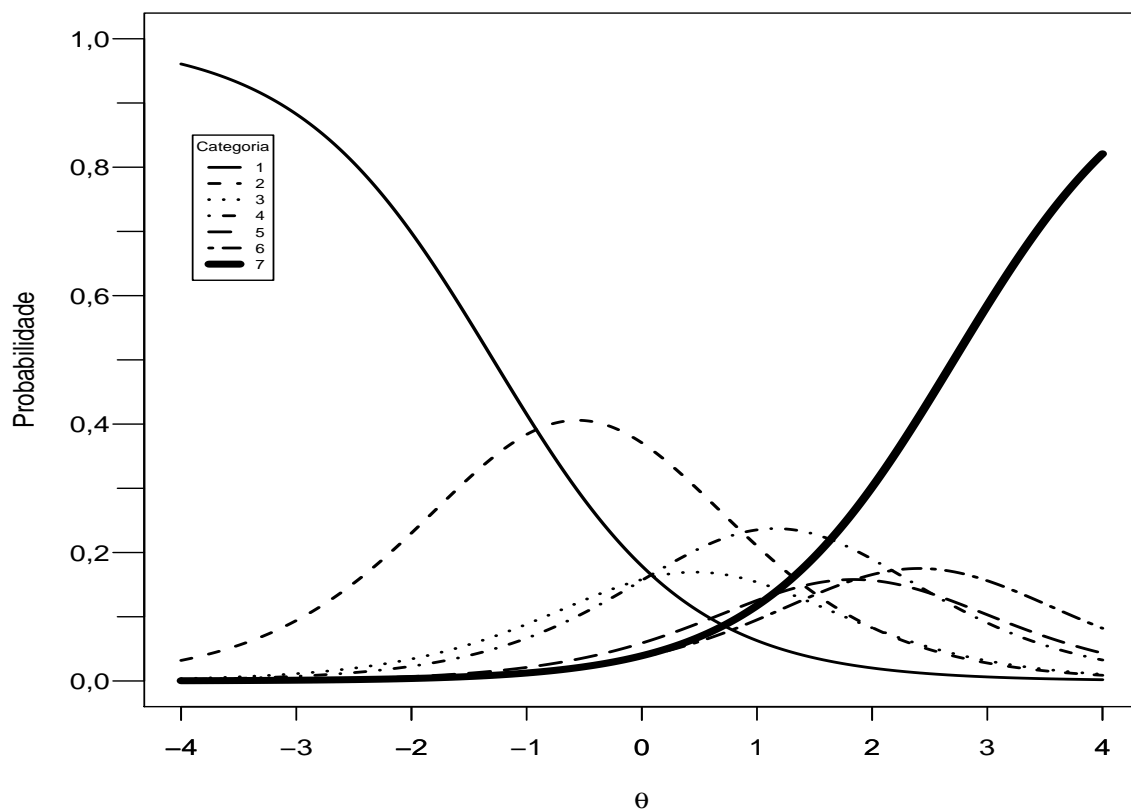


Figura 10 – Curvas características das categorias do item 9 - Sinto que estou a influenciar positivamente a vida de outras pessoas com o meu trabalho - relacionado a realização pessoal reduzida da enfermeira. Neste item, $\hat{a} = 1,18$, $\hat{b}_1 = -1,29$, $\hat{b}_2 = 0,17$, $\hat{b}_3 = 0,75$, $\hat{b}_4 = 1,57$, $\hat{b}_5 = 2,11$ e $\hat{b}_6 = 2,71$.

Fonte: Elaborada pelo autor.

valores elevados. O item 14 (“Sinto que estou a trabalhar demasiado no meu trabalho”), por sua vez, possui estimativas baixas para os parâmetros de dificuldade.

As Figuras 15, 16 e 17 apresentam curvas características das categorias dos itens, curvas de informação do item e função de informação do teste relacionados à exaustão emocional, despersonalização e realização pessoal reduzida, respectivamente, de acordo com a estimação dos parâmetros do modelo de resposta gradual via método da máxima verossimilhança marginal. Todas as figuras foram geradas utilizando o pacote ltm no R. A curva de informação do item mostra quão bem e precisamente cada item mede o traço latente em vários níveis desta variável. Determinados itens podem fornecer mais informações em níveis baixos do traço latente, enquanto outros podem fornecer mais informações em níveis mais altos. Já a função de informação do teste agrega as curvas de informação dos itens através de todos os itens, isto é, esta curva mostra como o teste, em geral, mede o traço latente em diferentes níveis.

Na Figura 15, as curvas características das categorias dos itens 2 e 6 são apresentadas. Como a estimativa do parâmetro de discriminação do item 2 possui um valor alto, nota-se

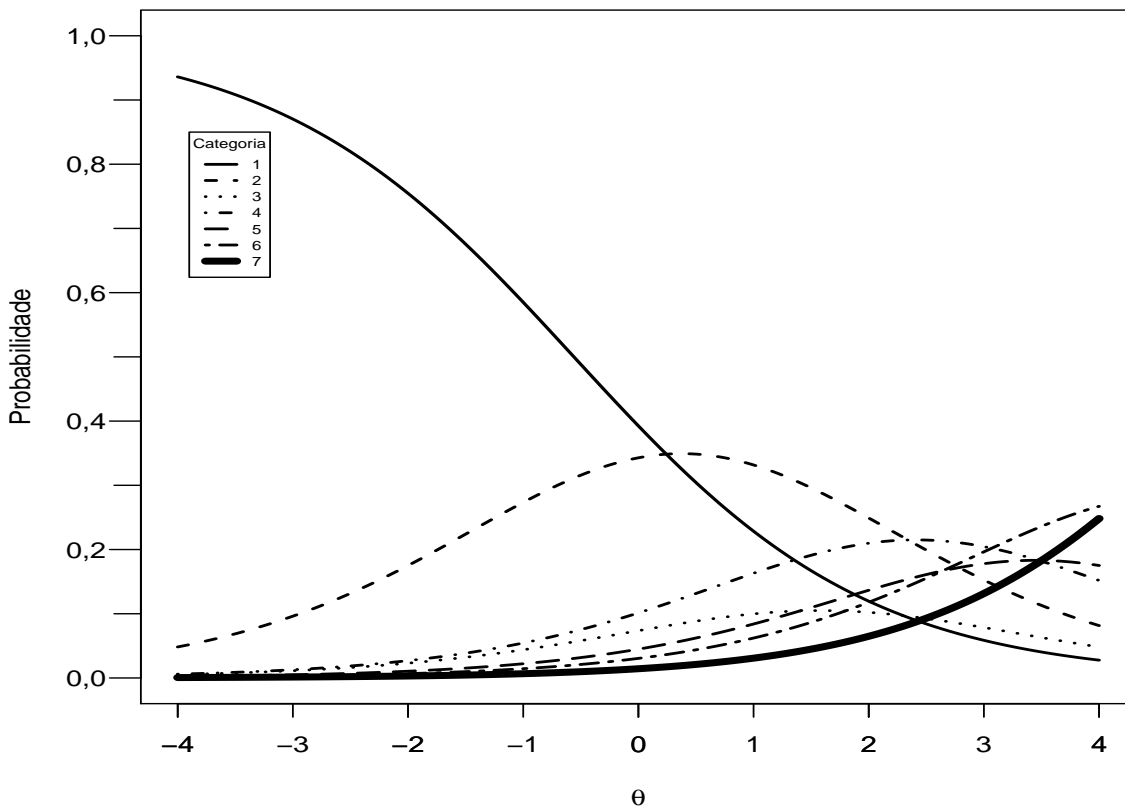


Figura 11 – Curvas características das categorias do item 22 - Sinto que alguns pacientes me culpam por alguns dos meus problemas - relacionado a despersonalização da enfermeira. Neste item, $\hat{a} = 0,78$, $\hat{b}_1 = -0,56$, $\hat{b}_2 = 1,31$, $\hat{b}_3 = 1,85$, $\hat{b}_4 = 2,97$, $\hat{b}_5 = 3,92$ e $\hat{b}_6 = 5,42$.

Fonte: Elaborada pelo autor.

que as curvas das categorias são bem discriminadas em diferentes níveis do traço latente. Em contrapartida, o item 6 possui um valor baixo do parâmetro de discriminação e, por isso as curvas das categorias nesse item não são bem discriminadas em diferentes traços latentes. Além disso, nota-se no item 6 que existe uma baixa probabilidade de resposta nas maiores categorias, visto que as estimativas dos parâmetros de dificuldade são altas neste item. A curva de informação do item mostra que o item 2 é o item que melhor mede o traço latente em diferentes níveis, e o item 6 é o item com pior desempenho. A função de informação do teste mostra que o conjunto dos 9 itens relacionados à exaustão emocional fornece 58,3% da informação total para os níveis mais altos do traço latente e 38,5% para os níveis mais baixos.

As curvas características das categorias dos itens 5 e 10, na Figura 16, são curvas com valores baixo e alto, respectivamente, da estimativa do parâmetro de discriminação do item. Nesta figura, a curva de informação do item 10 é a que mais fornece informação nos diferentes níveis da despersonalização, e a do item 22 é a que fornece menos informação. Além disso, 73,4% da informação total do conjunto dos 5 itens relacionados à despersonalização é dada

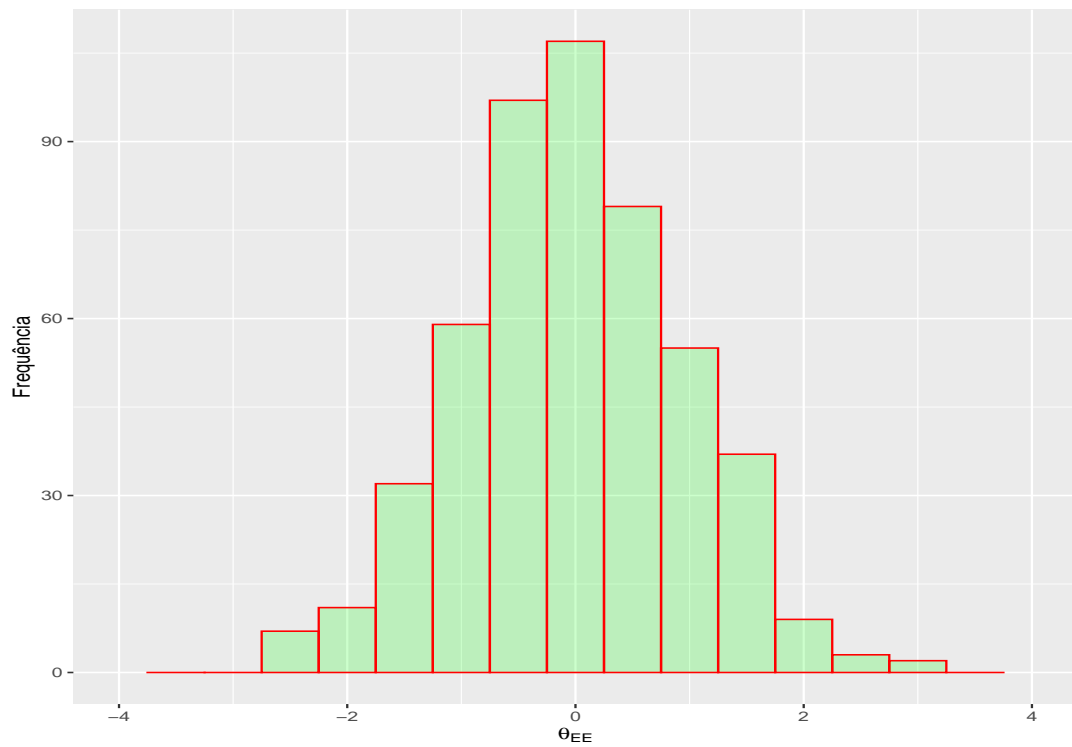


Figura 12 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a exaustão emocional do modelo de resposta gradual unidimensional.

Fonte: Elaborada pelo autor.



Figura 13 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a despersonalização do modelo de resposta gradual unidimensional.

Fonte: Elaborada pelo autor.

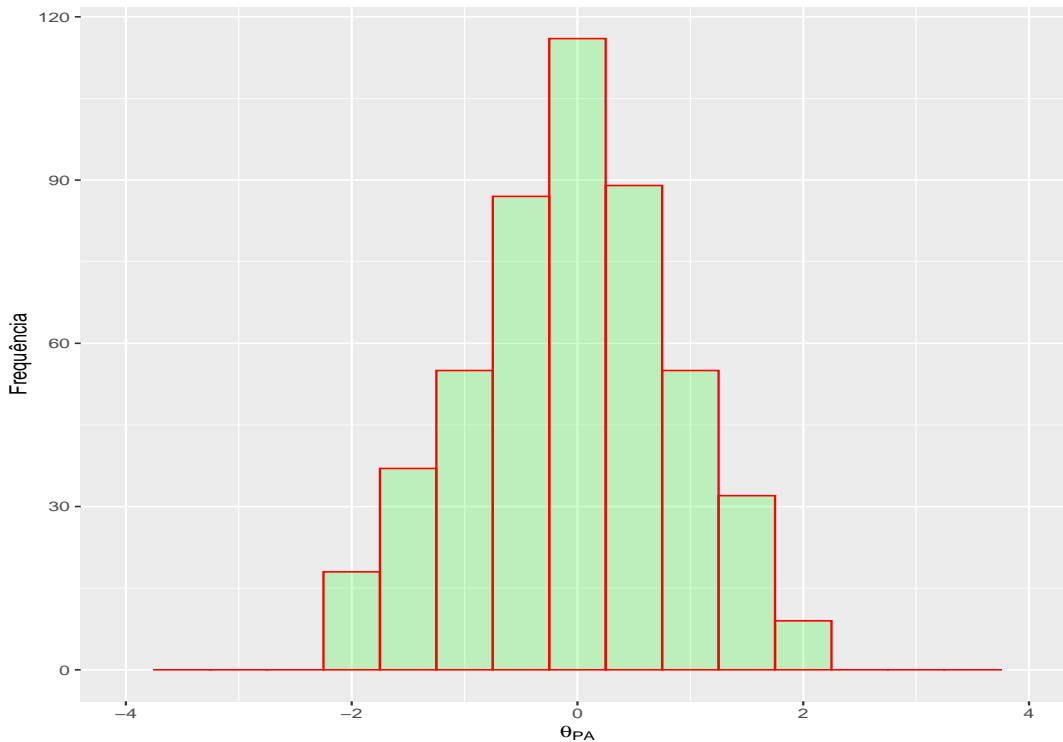


Figura 14 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a realização pessoal reduzida do modelo de resposta gradual unidimensional.

Fonte: Elaborada pelo autor.

pelos níveis mais altos do traço latente, enquanto apenas 15,6% é dada pelos níveis mais baixos.

Os itens 4 e 19 apresentam valores mais altos e mais baixos, respectivamente, das estimativas dos parâmetros de dificuldade, e é possível observar a diferença nas curvas característica das categorias destes itens na Figura 17. O item 19 é o item que fornece mais informação em diferentes níveis da realização pessoal reduzida, e o item 4 é o que fornece menos informação. O conjunto de 8 itens relacionados à realização pessoal reduzida fornece 61,3% da informação total para os níveis mais altos do traço latente, e 28,2% para os níveis mais baixos.

Tabela 9 – Estimativas dos parâmetros dos itens do modelo de resposta gradual unidimensional nos dados do MBI via método da máxima verossimilhança marginal.

Itens	a	b_1	b_2	b_3	b_4	b_5	b_6	b_{media}
Exaustão Emocional								
1	2,13	-1,53	-0,38	0,18	0,81	1,23	2,37	0,45
2	2,88	-1,57	-0,83	-0,46	0,23	0,50	1,69	-0,07
3	2,39	-1,46	-0,59	-0,14	0,51	0,96	1,93	0,20
6	0,62	0,23	2,22	3,18	5,16	7,35	9,15	4,55
8	2,51	-0,49	0,33	0,62	1,10	1,36	2,40	0,89
13	1,80	-0,98	0,01	0,45	1,13	1,48	2,40	0,75
14	1,54	-1,86	-0,91	-0,39	0,35	0,63	1,69	-0,08
16	1,23	-0,62	0,77	1,46	2,17	2,60	3,78	1,69
20	2,40	-0,65	0,36	0,76	1,33	1,70	2,66	1,03
Despersonalização								
5	1,04	-0,20	0,64	1,00	1,76	2,12	3,46	1,46
10	2,89	0,31	0,94	1,33	1,93	2,26	2,63	1,57
11	1,54	-0,24	0,75	1,27	1,78	2,03	2,35	1,32
15	1,13	1,22	2,21	2,91	3,60	4,20	4,62	3,13
22	0,62	-0,62	1,69	2,35	3,77	4,99	6,98	3,19
Realização Pessoal Reduzida								
4	0,91	0,20	1,85	2,79	3,92	4,46	5,25	3,08
7	1,49	0,11	1,68	2,07	2,69	3,00	3,63	2,20
9	1,30	-1,26	0,03	0,51	1,27	1,81	2,29	0,78
12	1,35	-1,45	0,15	0,72	1,58	2,25	3,00	1,04
17	1,86	-0,46	0,93	1,23	1,82	2,18	2,59	1,38
18	1,61	-0,63	0,86	1,37	2,38	2,94	3,94	1,81
19	2,29	-0,76	0,42	0,86	1,67	2,14	2,38	1,12
21	1,33	-1,08	0,43	0,84	1,65	2,46	4,23	1,42

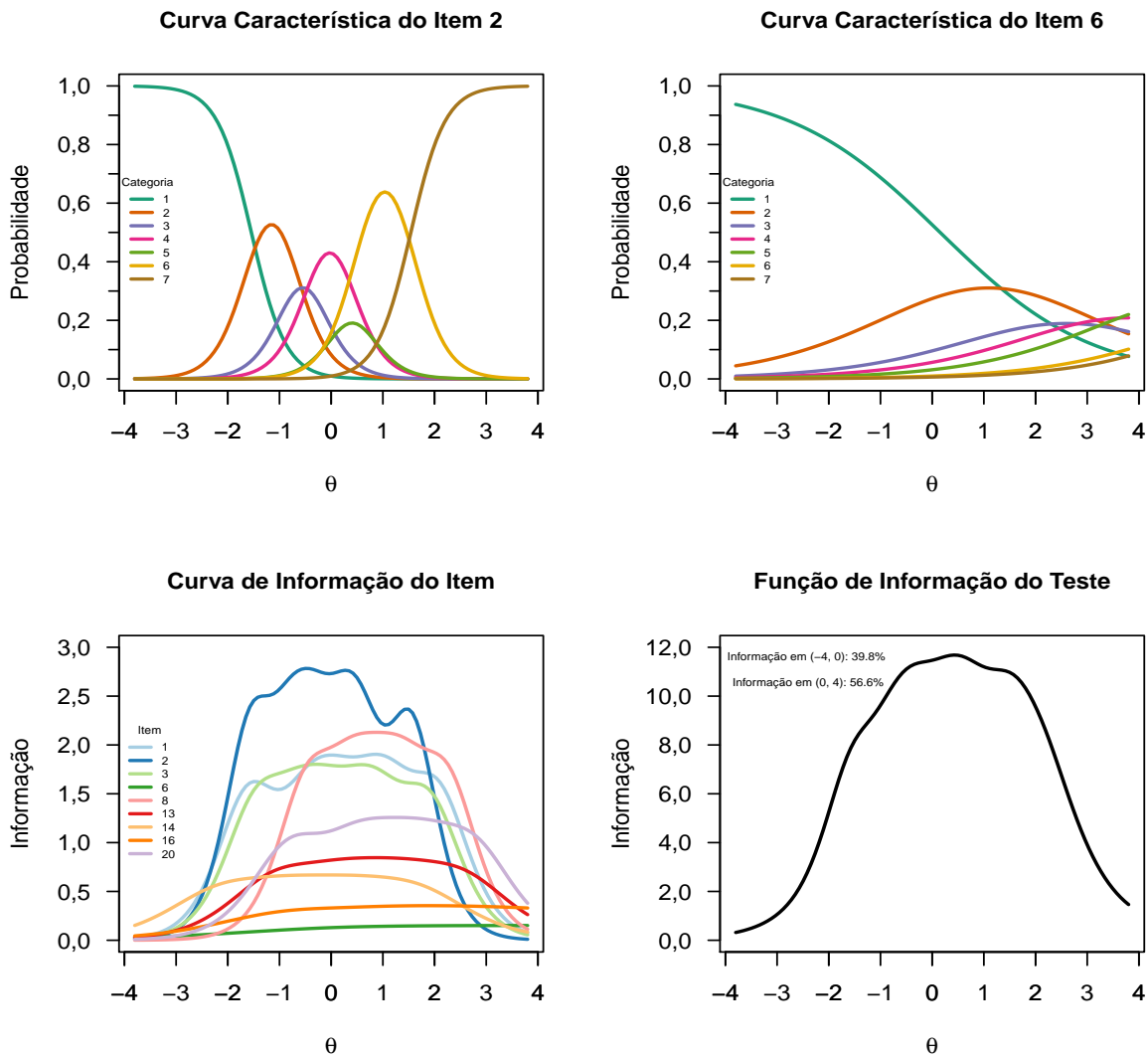


Figura 15 – Curvas características das categorias dos itens 2 e 6, curvas de informação do item e função de informação do teste dos itens relacionados à exaustão emocional das enfermeiras de acordo com a estimação dos parâmetros do modelo de resposta gradual via método da máxima verossimilhança marginal.

Fonte: Elaborada pelo autor.

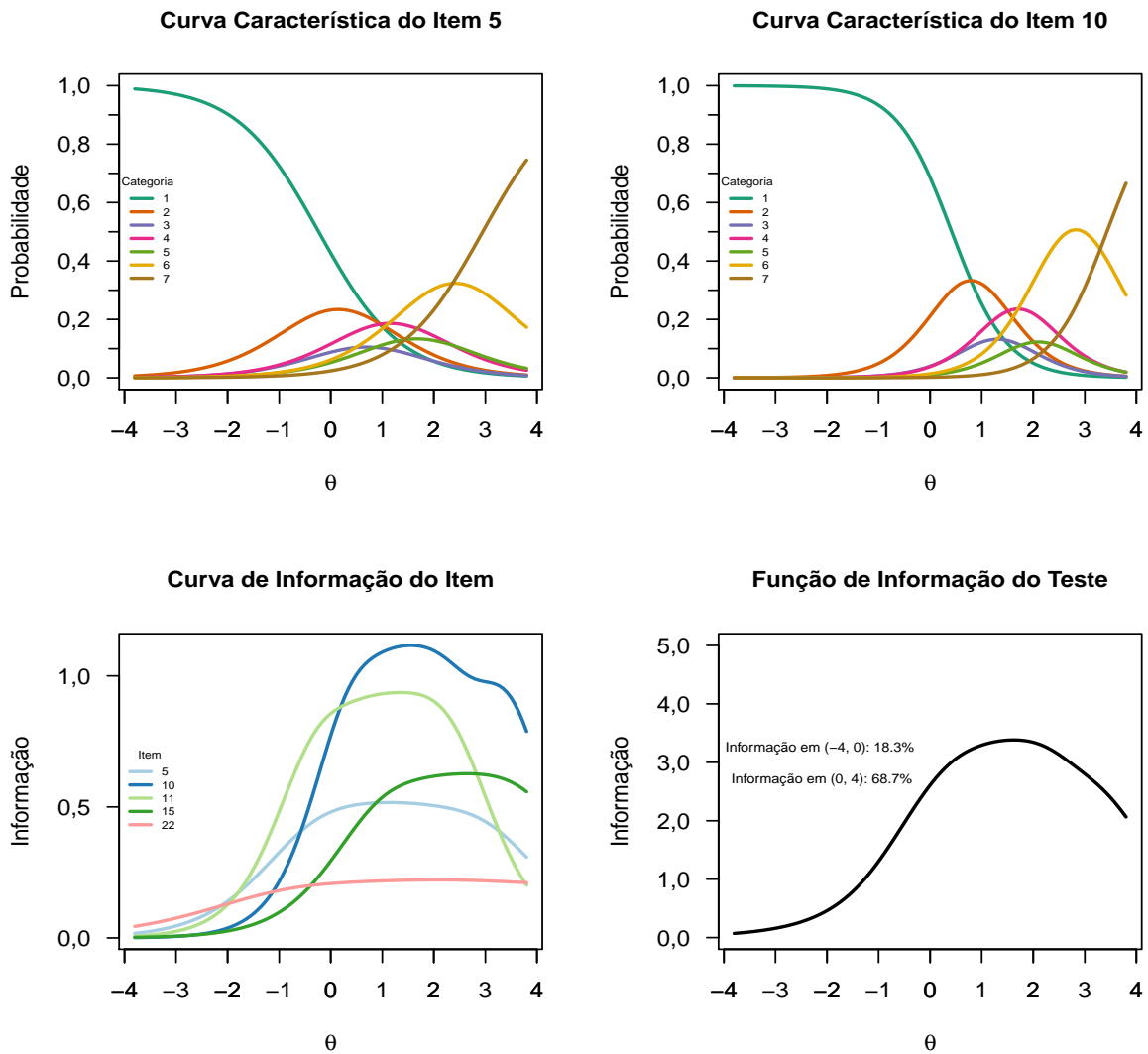


Figura 16 – Curvas características das categorias dos itens 5 e 10, curvas de informação do item e função de informação do teste dos itens relacionados à despersonalização das enfermeiras de acordo com a estimação dos parâmetros do modelo de resposta gradual via método da máxima verossimilhança marginal.

Fonte: Elaborada pelo autor.

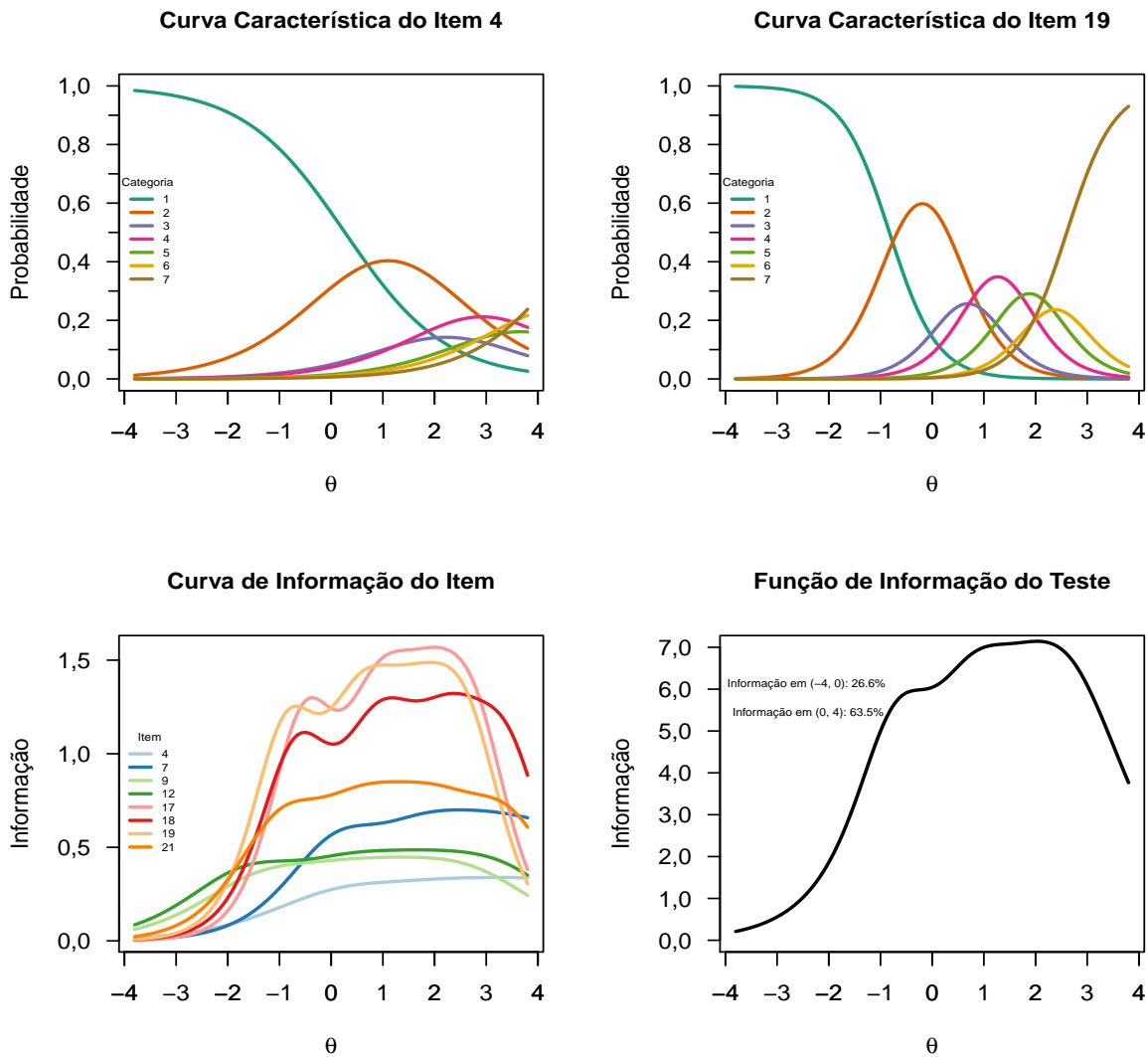


Figura 17 – Curvas características das categorias dos itens 4 e 19, curvas de informação do item e função de informação do teste dos itens relacionados à realização pessoal reduzida das enfermeiras de acordo com a estimação dos parâmetros do modelo de resposta gradual via método da máxima verossimilhança marginal.

Fonte: Elaborada pelo autor.

5.2 Abordagem multiunidimensional

5.2.1 Um estudo de simulação

A ideia é gerar resultados a fim de comparar a performance nos diferentes cenários para ajuste do modelo de resposta gradual multiunidimensional com estimação via método Bayesiano amostrador de Gibbs.

Assim como na abordagem unidimensional, vamos comparar as estimativas dos parâmetros dos itens e dos indivíduos com os valores dos parâmetros gerados. Além disso, também iremos analisar os resultados das estimativas da correlação entre os traços latentes. As estatísticas utilizadas são as mesmas descritas na abordagem unidimensional.

Neste estudo, três fatores são considerados: o tamanho do teste (8, 12, 16 e 22 itens), o tamanho da amostra (75, 150, 300, 500 e 750 indivíduos) e a correlação entre os traços latentes (com intensidade das associações como fraca, média e forte). Assim, as combinações resultam em 60 diferentes condições (4 tamanhos de teste \times 5 tamanhos de amostra \times 3 níveis de correlação).

5.2.1.1 Simulação dos dados

Para cada uma das 60 combinações do tamanho do teste, do tamanho da amostra e dos níveis de correlação, 20 conjuntos de dados diferentes foram gerados, isto é, o número de repetição R é 20. Consideramos três subtestes para cada teste, e para cada tamanho amostral consideramos uma partição diferente dos subtestes, sendo respectivamente (3, 3, 2), (4, 4, 4), (6, 5, 5) e (8, 7, 7) para os tamanhos amostrais 8, 12, 16 e 22. Como em cada teste levamos em consideração três traços latentes dos indivíduos, o parâmetro do indivíduo, θ_j , $j = 1, \dots, n$, foi gerado de uma distribuição normal multivariada, $N(\mu_\theta, \Sigma_\theta)$, em que $\mu_\theta = (0, 0, 0)$ e $\Sigma_\theta = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}$, com ρ_{12} , ρ_{13} e ρ_{23} sendo iguais a, respectivamente, 0,2, 0,1 e 0,3 (ρ_{fr}) para associação fraca, 0,6, 0,4 e 0,5 (ρ_{me}) para associação média e 0,7, 0,9 e 0,8 (ρ_{fo}) para associação forte. Assim, para determinado tamanho amostral e a cada repetição, um novo conjunto de parâmetros do indivíduo foi simulado. Os parâmetros de discriminação e dificuldade dos itens usados, respectivamente a_{hi} , $h = 1, \dots, H$, com $H = 3$ e $i = 1, \dots, I_h$ e b_{ik} , $k = 1, \dots, 6$, foram estimativas dos parâmetros de discriminação e de dificuldade do modelo de resposta gradual multiunidimensional aplicados a dados reais de enfermeiras, que são os dados descritos no capítulo 2, considerando uma escala de 7 pontos.

As respostas foram geradas utilizando os mesmos parâmetros dos itens em todas as repetições. Os valores podem ser observados na Tabela 10, em que um teste simulado de tamanho L contem os primeiros L itens, e o tamanho de cada subteste de um teste de tamanho L foram considerados de forma respectiva, a depender de cada partição comentada nesta seção.

A partir dos valores simulados dos parâmetros dos indivíduos e dos itens, calcula-se a

Tabela 10 – Parâmetros dos itens usados na simulação dos dados na abordagem multiunidimensional.

Itens	a	b_1	b_2	b_3	b_4	b_5	b_6
1	2,02	-1,64	-0,44	0,13	0,78	1,21	2,38
2	2,66	-1,69	-0,92	-0,52	0,19	0,47	1,69
3	2,32	-1,56	-0,66	-0,20	0,46	0,92	1,91
4	0,87	0,08	1,56	2,27	3,64	4,99	6,07
5	2,50	-0,55	0,28	0,58	1,06	1,33	2,37
6	1,80	-1,05	-0,04	0,40	1,09	1,44	2,36
7	1,47	-2,00	-1,01	-0,46	0,30	0,60	1,71
8	1,26	-0,67	0,70	1,38	2,08	2,52	3,68
9	2,39	-0,71	0,30	0,71	1,27	1,65	2,62
10	1,24	-0,24	0,51	0,84	1,51	1,85	3,03
11	2,29	0,27	0,95	1,38	2,03	2,41	2,87
12	1,50	-0,31	0,69	1,22	1,75	2,02	2,37
13	1,30	1,05	1,92	2,54	3,15	3,67	4,12
14	0,81	-0,58	1,26	1,78	2,88	3,80	5,28
15	1,01	0,12	1,63	2,47	3,43	3,92	4,65
16	1,46	0,05	1,62	2,01	2,62	2,96	3,62
17	1,29	-1,35	-0,04	0,44	1,21	1,76	2,25
18	1,50	-1,43	0,08	0,62	1,42	2,05	2,75
19	1,91	-0,51	0,85	1,15	1,72	2,08	2,50
20	1,48	-0,72	0,82	1,35	2,40	2,98	4,04
21	2,28	-0,82	0,35	0,79	1,60	2,07	2,34
22	1,33	-1,15	0,36	0,77	1,58	2,39	4,11

probabilidade do indivíduo responder, em cada item, a determinada categoria ou outra maior, de acordo com o modelo de resposta gradual multiunidimensional. Com base na probabilidade do indivíduo responder cada categoria, determina-se a resposta do indivíduo a cada item do teste, isto é, a resposta do indivíduo será designada como sendo a categoria com maior probabilidade de resposta de acordo com o modelo de resposta gradual multiunidimensional.

5.2.1.2 Estimação dos parâmetros

Para obter as estimativas dos parâmetros do modelo de resposta gradual multiunidimensional pelo método Bayesiano amostrador de Gibbs foi utilizado o JAGS e seus respectivos pacotes no *software* R. O código utilizado no JAGS para ajustar o modelo de resposta gradual multiunidimensional está apresentado no Apêndice B.

Considerou-se a distribuição Normal truncada $[0, +\infty)$ com média 1 e desvio padrão 2,5 como distribuição a priori do parâmetro de discriminação, a_{hi} , $h = 1, \dots, H$ com $H = 3$ e $i = 1, \dots, I_h$ e a distribuição Normal com média 0 e desvio padrão 2,5 como a priori para os parâmetros de dificuldade, b_{ik} , $k = 1, \dots, 6$, sujeito a restrição $b_{i1} \leq b_{i2} \leq \dots \leq b_{i6}$. Para os traços latentes dos indivíduos considerou-se uma distribuição a priori normal multivariada com

vetor de média $(0,0,0)^\top$ e a matriz de variância e covariância considerada foi com variâncias iguais a 1 e as covariâncias iguais aos coeficientes de correlação usados na geração dos dados. Além disso, a cada repetição os valores iniciais dos parâmetros do i -ésimo item foram fixados em $a_i = 1,0$, $b_{i1} = -1,5$, $b_{i2} = -1,0$, $b_{i3} = -0,5$, $b_{i4} = 0,5$, $b_{i5} = 1,0$ e $b_{i6} = 1,5$. As primeiras 1000 iterações foram consideradas o período de aquecimento da cadeia (*burn in*), e 10.000 iterações foram utilizadas para estimar os parâmetros.

A comparação para avaliar a performance dos modelos nos diferentes cenários apresentados será realizada pelas mesmas estatísticas descritas na seção da abordagem unidimensional: REQM, viés e correlação.

No caso do parâmetro do indivíduo θ_h , o REQM é calculado como

$$\text{REQM}_r(\theta_h) = \sqrt{\frac{1}{n} \sum_{j=1}^n (\hat{\theta}_{hjr} - \theta_{hjr})^2},$$

em que $\hat{\theta}_{hjr}$ e θ_{hjr} representam, respectivamente, o valor estimado e simulado do j -ésimo indivíduo, na dimensão h e na repetição r , $r = 1, \dots, R$, com $R = 20$. Após obter os R REQM, calcula-se o seu valor médio. O viés para o parâmetro do indivíduo é dado por

$$\text{vies}(\theta_{h,j}) = \frac{1}{R} \sum_{r=1}^R (\hat{\theta}_{hjr} - \theta_{hjr}).$$

Os resultados do estudo de simulação considerando o modelo de resposta gradual multiunidimensional via amostrador de Gibbs estão apresentados nas Tabelas 15, 16 e 17 (para dados gerados com correlação fraca), 18, 19 e 20 (para dados gerados com correlação média) e 21, 22 e 23 (para dados gerados com correlação forte) no Apêndice C.

Ao observar o REQM e o viés do parâmetro de discriminação, a , nota-se que à medida que o tamanho amostral e do teste e, consequentemente do subteste, aumentam, os resultados em geral tendem a melhorar. Por exemplo, na Tabela 15, quando o tamanho amostral é 75 e no subteste S1 do teste de tamanho 12, o REQM médio é de 0,75. Já quando o tamanho amostral é 500, o REQM médio é de 0,41. Além disso, os diferentes níveis de correlação entre os traços latentes parecem afetar pouco na estimação do parâmetro de discriminação quando o tamanho amostral e do teste são pequenos, e parecem não afetar à medida que o tamanho amostral e do teste aumentam. Por exemplo, nas Tabelas 16, 19 e 22 observa-se quando a amostra é de tamanho 150 e o subteste S2 para o tamanho do teste 8 que o viés médio para o parâmetro a é de $-0,92$ quando a correlação é fraca, $-0,52$ quando a correlação é média e $-0,43$ quando a correlação é forte. Já quando o tamanho amostral é 500 e o subteste é S1 no tamanho do teste 22, o viés médio é de 0,02 quando a correlação é fraca e de 0,03 quando as correlações são média e forte.

Os resultados para os parâmetros de dificuldade b_1, b_2, b_3, b_4, b_5 e b_6 são similares aos do parâmetro de discriminação a , isto é, à medida que o tamanho amostral e do teste aumentam,

o REQM médio e o viés médio em geral tendem a diminuir. Por exemplo, na Tabela 22, para o parâmetro b_1 quando a amostra é de tamanho 75 e para S1 quando o tamanho do teste é 22 o viés médio é de $-0,53$ e quando o tamanho amostral é 750, o viés médio é de $-0,07$. Além disso, nos parâmetros de dificuldade também é possível observar que os diferentes níveis de correlação entre os traços latentes parecem afetar pouco na estimação do parâmetro de discriminação quando o tamanho amostral e do teste são pequenos, e parecem não afetar à medida que o tamanho amostral e do teste aumentam. Nas Tabelas 15, 18 e 21, por exemplo, quando o tamanho amostral é 75 e para o subteste S2 no tamanho do teste 12, o REQM médio do parâmetro b_3 é 0,35 quando a correlação é fraca, 0,26 quando a correlação é média e 0,31 quando a correlação é forte. Para este mesmo parâmetro, quando a amostra é de tamanho 500, e para o subteste S2 do tamanho do teste 22, o REQM médio é de 0,20 quando a correlação é fraca, 0,21 quando a correlação é média e 0,23 quando a correlação é forte.

É possível observar nas Tabelas 17, 20 e 23 que as correlações médias entre parâmetros dos itens a , b_1 , b_2 , b_3 , b_4 , b_5 e b_6 e suas estimativas em geral são baixas quando o tamanho amostral e do teste são pequenos. À medida que esses tamanhos aumentam, a correlação média também aumenta. Por exemplo na Tabela 20, para o parâmetro b_2 , o tamanho do teste 75 e o subteste S1 no tamanho do teste 8, a correlação média é de 0,79 e quando o tamanho amostral é 500, a correlação média é 0,95. Além disso, é possível notar que para o parâmetro de discriminação do item a , quando o tamanho do teste é pequeno a correlação média é muito baixa. À medida que o tamanho do teste aumenta, a correlação média entre o parâmetro a e suas estimativas aumenta.

Ao observar os resultados do REQM médio para parâmetro do indivíduo, θ , nota-se que, em geral, há uma diferença nos resultados nos diferentes níveis de correlação dos traços latentes. Quando a correlação é fraca, observa-se que o REQM médio em geral diminui à medida que o tamanho amostral aumenta, em todos os tamanhos dos testes. Quando a correlação entre os traços latentes é média, observa-se que quando o tamanho do teste é pequeno, à medida que o tamanho amostra aumenta, o REQM médio em geral não diminui, porém quando o tamanho do teste é maior, o REQM médio tende a ter uma leve melhora à medida que o tamanho amostral cresce. Por exemplo, na Tabela 18, quando o tamanho amostral é 75 e o subteste é S2 no tamanho do teste 12, o REQM médio de θ é 0,90, e quando o tamanho amostral é 500, o REQM médio é de 1,12. Já quando para o subteste S3 no tamanho do teste 22, quando o tamanho amostral é 75, o REQM médio é de 0,51 e quando o tamanho amostral é 500, o REQM é 0,41. Já quando a correlação é forte, nota-se que à medida que o tamanho amostral aumenta, o REQM muda muito pouco quando o tamanho do teste é pequeno, e há uma mudança maior quando o tamanho do teste aumenta.

A Tabela 24 apresenta os resultados das correlações médias entre os traços latentes estimados nos diferentes tamanhos amostrais (75, 150, 300, 500, 750 e 1000), tamanhos dos testes (8, 12, 16 e 22) e níveis de correlação (fraca, média e forte). Os parâmetros ρ_{12} , ρ_{23} e ρ_{13}

são respectivamente iguais a 0,2, 0,3 e 0,1 na correlação fraca, 0,6, 0,5 e 0,4 na correlação média e 0,7, 0,8 e 0,9 na correlação forte. É possível notar que as correlações médias estimadas não são próximas ao valor da correlação quando o tamanho do teste é pequeno em todos os diferentes níveis de correlação. À medida que o tamanho do teste aumenta, as correlações médias estimadas se aproximam do valor da correlação. Por exemplo, quando a correlação é fraca, o tamanho amostral é 500 e o tamanho do teste é 8, as correlações médias estimadas de ρ_{12} , ρ_{23} e ρ_{13} são respectivamente 0,99, 0,97 e 0,97. Já quando o tamanho do teste é 22, as correlações médias estimadas são respectivamente 0,24, 0,36 e 0,12.

5.2.2 Uma aplicação aos dados de enfermeiras

5.2.2.1 Amostragem dos dados

Na abordagem multiunidimensional foram utilizados os mesmos dados da abordagem unidimensional. O processo de amostragem dos dados das enfermeiras do projeto RN4CAST estão explicados na seção 5.1.2.1. Assim, consideramos uma amostra de 498 enfermeiras provenientes da Bélgica para ajustar o modelo de resposta gradual multiunidimensional.

5.2.2.2 Modelagem dos dados

Nesta seção será apresentada uma aplicação do modelo de resposta gradual multiunidimensional aos dados de enfermeira via algoritmo amostrador de Gibbs. Considerou-se $n = 498$ enfermeiras que trabalham em hospitais da Bélgica provenientes do projeto RN4CAST, como apresentado no Capítulo 2. A estimação dos parâmetros via algoritmo amostrador de Gibbs foi feita pelos pacotes `rjags` e `R2jags` no *software* R.

Considerou-se cada dimensão da síndrome de *burnout* como uma dimensão do modelo de resposta gradual multiunidimensional. Assim, a exaustão emocional, a despersonalização e a realização pessoal reduzida das $n = 498$ enfermeiras são estimadas com base nas respostas das enfermeiras aos itens do MBI, sendo que os subtestes possuem 9 itens do MBI (itens 1, 2, 3, 6, 8, 13, 14, 16 e 20), 5 itens do MBI (itens 5, 10, 11, 15 e 22) e 8 itens do MBI (itens 4, 7, 9, 12, 17, 18, 19 e 21). Isto é, estimou-se os traços latentes das enfermeiras, a relação entre eles e os parâmetros dos respectivos itens das dimensões.

Em termos da síndrome de *burnout*, agora Y_{hij} representa a resposta da j -ésima enfermeira no i -ésimo item pertencente ao h -ésimo subteste no MBI. Cada item i , $i = 1, \dots, I_h$, em que I_h pode ser 9, 5 ou 8 para $h = 1, h = 2$ e $h = 3$ respectivamente, tem $m_i = 7$ categorias (0 a 6) mas considerou-se que $Y_{hij} = 1, \dots, 7$ (a categoria 0 é representada por $Y_{hij} = 1$, a categoria 1 é representada por $Y_{hij} = 2$, e assim por diante) e $j = 1, \dots, n$ é o número de enfermeiras e, nesse caso, $n = 498$.

As *prioris* dos parâmetros do item e do traço latente do indivíduo são as mesmas utilizadas na estimação dos parâmetros dos dados simulados. Isto é, a distribuição Normal truncada $[0, +\infty)$

com média 1 e desvio padrão 2,5 como distribuição a priori do parâmetro de discriminação, a_{hi} , $h = 1, \dots, H$ com $H = 3$ e $i = 1, \dots, I_h$ e a distribuição Normal com média 0 e desvio padrão 2,5 como a priori para os parâmetros de dificuldade, b_{ik} , $k = 1, \dots, 6$, sujeito a restrição $b_{i1} \leq b_{i2} \leq \dots \leq b_{i6}$. Para os traços latentes dos indivíduos considerou-se uma distribuição a priori normal multivariada com vetor de média $(0, 0, 0)^\top$ e a matriz de variância e covariância considerada foi com variâncias iguais a 1 e as covariâncias iguais as correlações entre os traços latentes do modelo de resposta gradual unidimensional estimados pelo método da máxima verossimilhança marginal, isto é, 0,5 entre exaustão emocional e despersonalização e 0,4 entre exaustão emocional e realização pessoal reduzida e entre despersonalização e realização pessoal reduzida. Utilizou-se 3 cadeias, cada uma com 100.000 iterações e com as primeiras 10.000 iterações descartadas (*burn in*).

A Tabela 11 apresenta os resultados das estimativas dos parâmetros dos itens via amostrador de Gibbs do modelo de resposta gradual multiunidimensional nos dados das enfermeiras. Nota-se que as estimativas dos parâmetros de discriminação são maiores do que 1 em todos os itens, exceto nos itens 14 (“Sinto que estou a trabalhar demasiado no meu emprego”), relacionado à exaustão emocional, com $\hat{a} = 0,91$ e 4 (“Consigo compreender facilmente como os meus pacientes se sentem acerca das coisas”), relacionado à realização pessoal reduzida, com $\hat{a} = 0,87$. Os itens com maiores estimativas nos parâmetros de discriminação a são os itens 2 (“No fim do dia do trabalho, sinto-me exausto”), relacionado à exaustão emocional, e 5 (“Sinto que trato alguns pacientes como se fossem objetos impessoais”), relacionado à despersonalização, com $\hat{a} = 2,83$ e $\hat{a} = 2,60$, respectivamente.

Já os resultados relacionados aos parâmetros de dificuldade mostram que as maiores estimativas deste parâmetro, considerando a dificuldade média, são nos itens 4 (“Consigo compreender facilmente como os meus pacientes se sentem acerca das coisas”), relacionado à realização pessoal reduzida, 15 (“De fato, não me interessa o que acontece a alguns pacientes”), relacionado à despersonalização, e 13 (“Sinto-me muito frustrado com o meu trabalho”), relacionado à exaustão emocional. Os itens com menor dificuldade média são 7 (“Lido muito eficazmente com os problemas dos meus pacientes”), relacionado à realização pessoal reduzida, 5 (“Sinto que trato alguns pacientes como se fossem objetos impessoais.”), relacionado à despersonalização e 2 (“No fim do dia do trabalho, sinto-me exausto”), relacionado à exaustão emocional.

As Figuras 18, 19 e 20 apresentam as curvas características das categorias dos itens 2 - relacionado à exaustão emocional, 5 - relacionado à despersonalização e 4 - relacionado à realização pessoal reduzida, respectivamente. O item 2 apresenta grande discriminação estimada e baixa dificuldade estimada. Já o item 4 apresenta baixa discriminação estimada e alta dificuldade estimada. O item 5 apresenta valor baixo da estimativa do parâmetro de discriminação.

As correlações estimadas entre os traços latentes são 0,70 entre a exaustão emocional e a despersonalização, 0,49 entre a exaustão emocional e a realização pessoal reduzida e 0,59 entre a despersonalização e a realização pessoal reduzida. As Figuras 21, 22 e 23 apresentam

Tabela 11 – Estimativas dos parâmetros dos itens do modelo de resposta gradual multiunidimensional nos dados do MBI via amostrador Gibbs.

Itens	a	b_1	b_2	b_3	b_4	b_5	b_6	b_{media}
Exaustão Emocional								
1	2,35	-1,70	-0,49	0,04	0,77	1,11	2,05	0,30
2	2,83	-1,67	-0,84	-0,40	0,24	0,50	1,53	-0,11
3	2,34	-1,52	-0,70	-0,17	0,53	0,87	1,86	0,15
6	1,62	-1,11	-0,04	0,49	1,04	1,56	2,49	0,74
8	1,10	-0,88	0,70	1,36	2,16	2,77	4,22	1,72
13	1,45	0,99	1,90	2,20	2,77	3,09	3,73	2,45
14	0,91	-0,69	0,91	1,56	2,43	3,04	4,98	2,04
16	1,54	0,11	1,52	2,01	2,64	3,34	4,56	2,36
20	1,90	-0,72	0,76	1,20	2,12	2,61	3,50	1,58
Despersonalização								
5	2,60	-0,51	0,26	0,60	1,07	1,42	2,20	0,84
10	1,52	-0,28	0,42	0,73	1,27	1,66	2,62	1,07
11	1,60	0,40	1,20	1,52	2,09	2,39	3,77	1,90
15	1,10	0,19	1,76	2,28	3,05	3,64	4,55	2,58
22	1,65	-0,96	0,36	0,85	1,56	2,14	3,31	1,21
Realização Pessoal Reduzida								
4	0,87	0,04	1,56	2,45	3,41	4,59	5,62	2,95
7	1,37	-2,27	-1,15	-0,63	0,22	0,63	1,65	-0,26
9	2,01	-0,84	0,35	0,78	1,36	1,86	2,70	1,04
12	1,55	-0,30	0,62	1,06	1,65	1,99	2,44	1,24
17	1,16	-1,33	0,16	0,75	1,57	2,13	2,73	1,00
18	1,35	-1,56	0,17	0,62	1,50	2,11	3,01	0,98
19	2,20	-0,55	0,77	1,19	1,81	2,17	2,72	1,35
21	2,16	-0,88	0,40	0,88	1,55	2,11	2,58	1,11

os histogramas das estimativas dos traços latentes da exaustão emocional, despersonalização e realização pessoal reduzida, respectivamente.

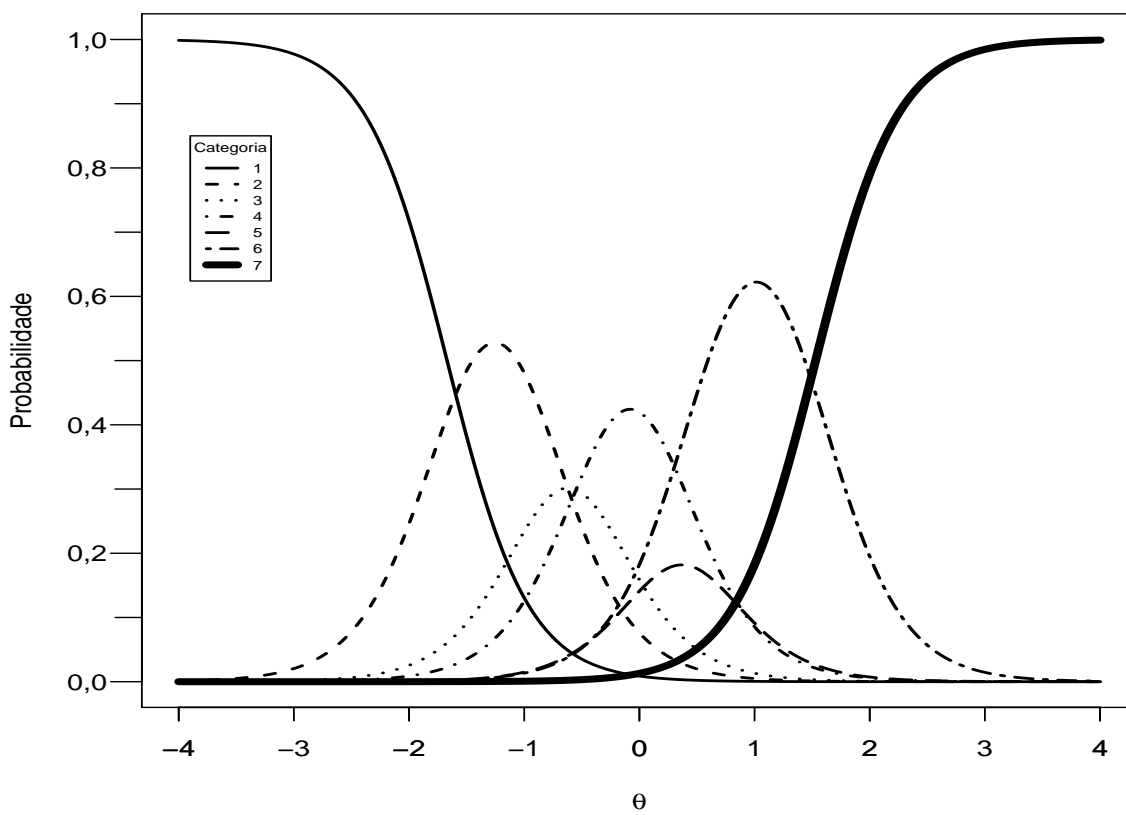


Figura 18 – Curvas características das categorias do item 2 - No fim do dia do trabalho, sinto-me exausto - relacionado à exaustão emocional da enfermeira. Neste item, $\hat{a} = 2,83$, $\hat{b}_1 = -1,67$, $\hat{b}_2 = -0,84$, $\hat{b}_3 = -0,40$, $\hat{b}_4 = 0,24$, $\hat{b}_5 = 0,50$ e $\hat{b}_6 = 1,53$.

Fonte: Elaborada pelo autor.

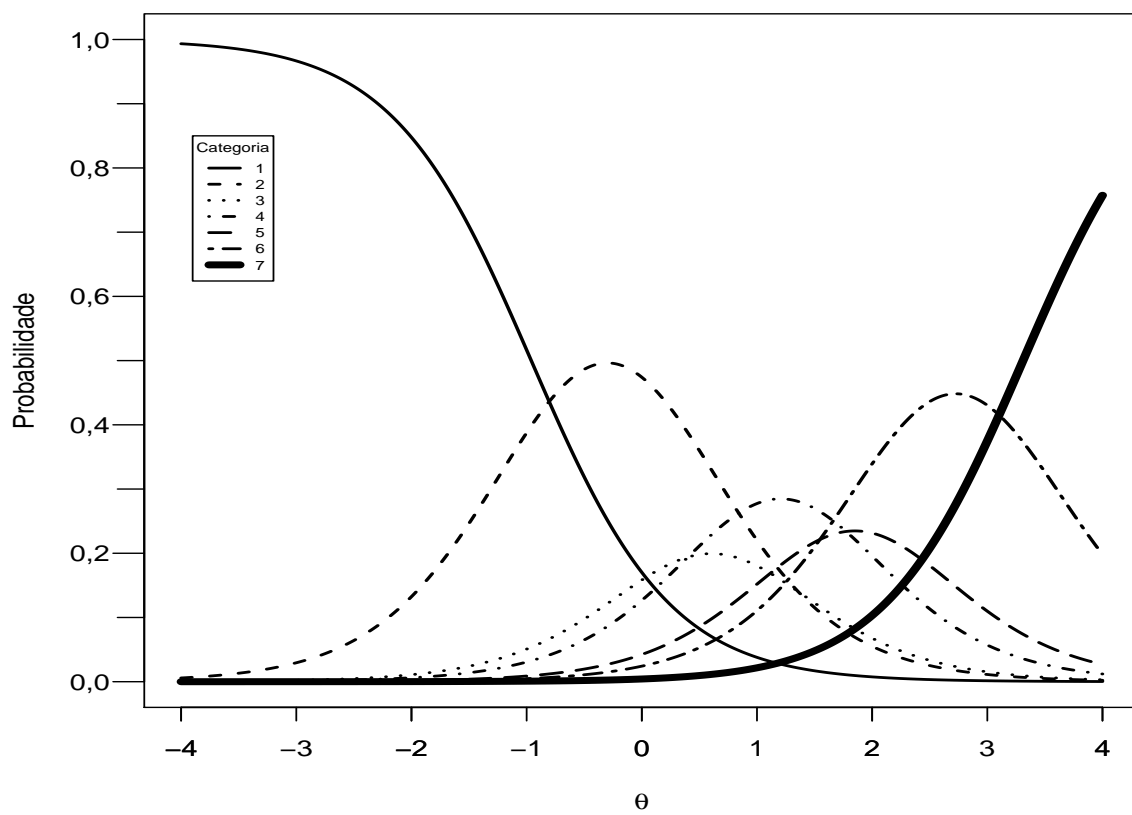


Figura 19 – Curvas características das categorias do item 5 - Sinto que trato alguns pacientes como se fossem objetos impessoais - relacionado à despersonalização da enfermeira. Neste item, $\hat{a} = 2,60$, $\hat{b}_1 = -0,51$, $\hat{b}_2 = 0,26$, $\hat{b}_3 = 0,60$, $\hat{b}_4 = 1,07$, $\hat{b}_5 = 1,42$ e $\hat{b}_6 = 2,20$.

Fonte: Elaborada pelo autor.

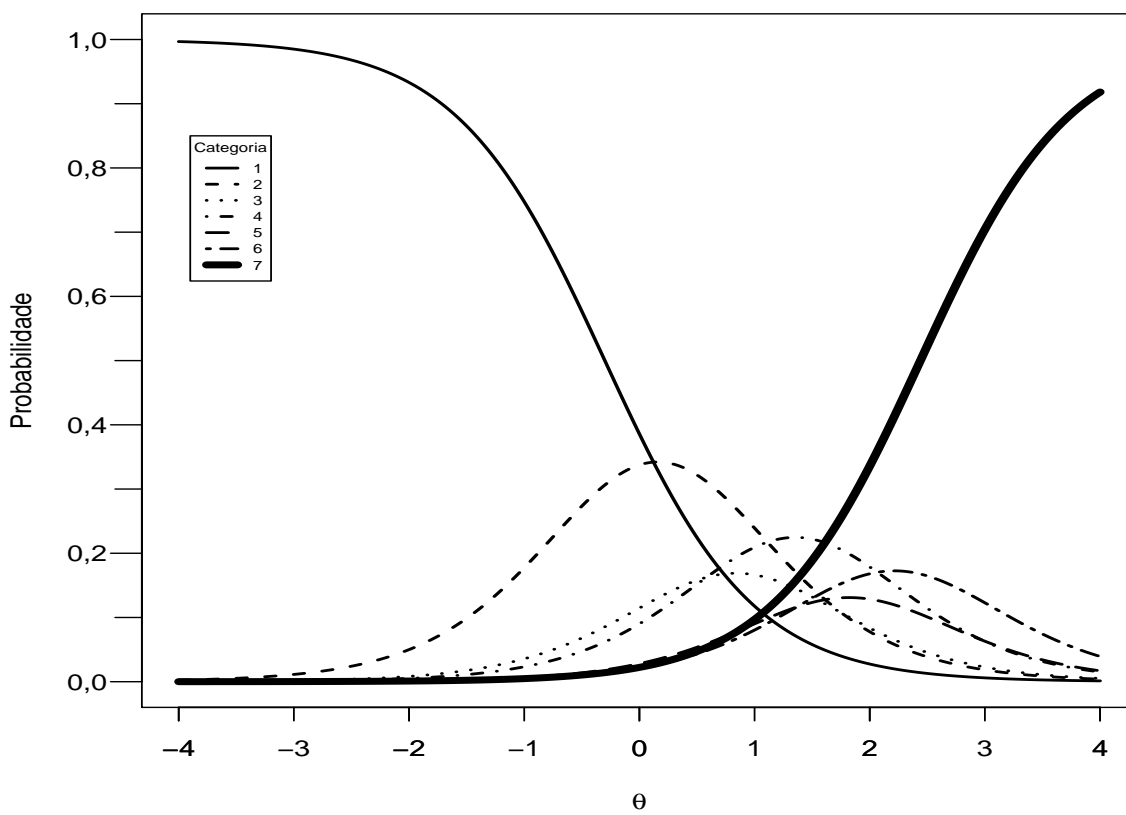


Figura 20 – Curvas características das categorias do item 4 - Consigo compreender facilmente como os meus pacientes se sentem acerca das coisas - relacionado à realização pessoal reduzida da enfermeira. Neste item, $\hat{a} = 0,87$, $\hat{b}_1 = 0,04$, $\hat{b}_2 = 1,56$, $\hat{b}_3 = 2,45$, $\hat{b}_4 = 3,41$, $\hat{b}_5 = 4,59$ e $\hat{b}_6 = 5,62$.

Fonte: Elaborada pelo autor.

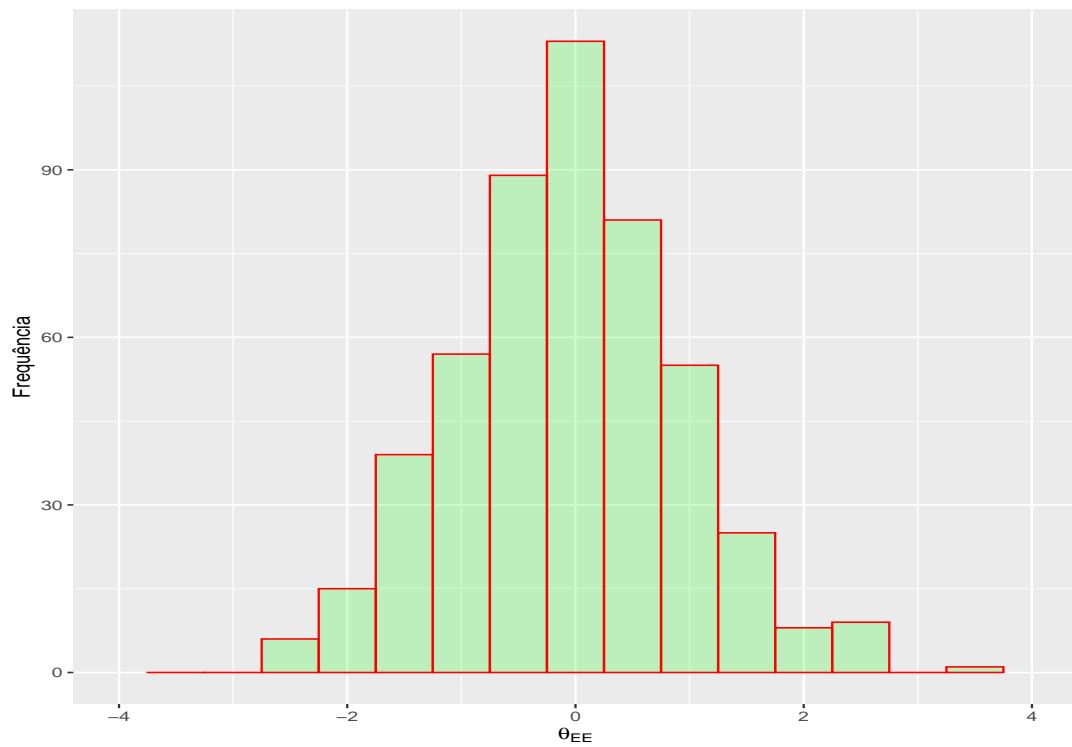


Figura 21 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a exaustão emocional do modelo de resposta gradual multiunidimensional.

Fonte: Elaborada pelo autor.

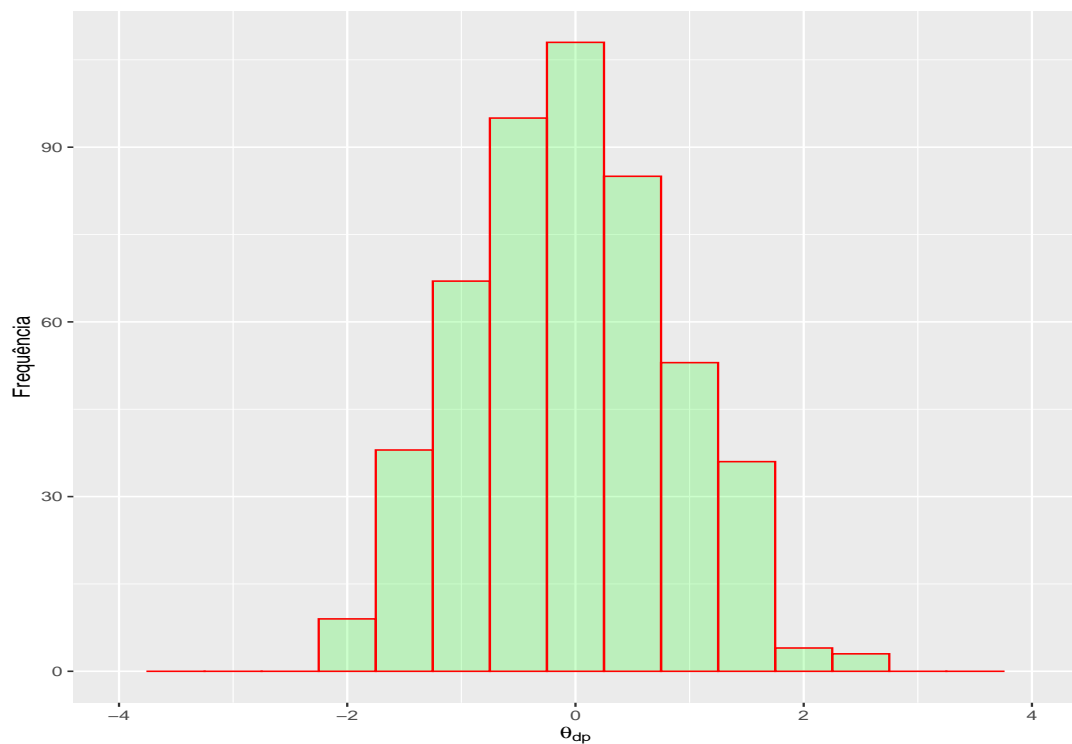


Figura 22 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a despersonalização do modelo de resposta gradual multiunidimensional.

Fonte: Elaborada pelo autor.

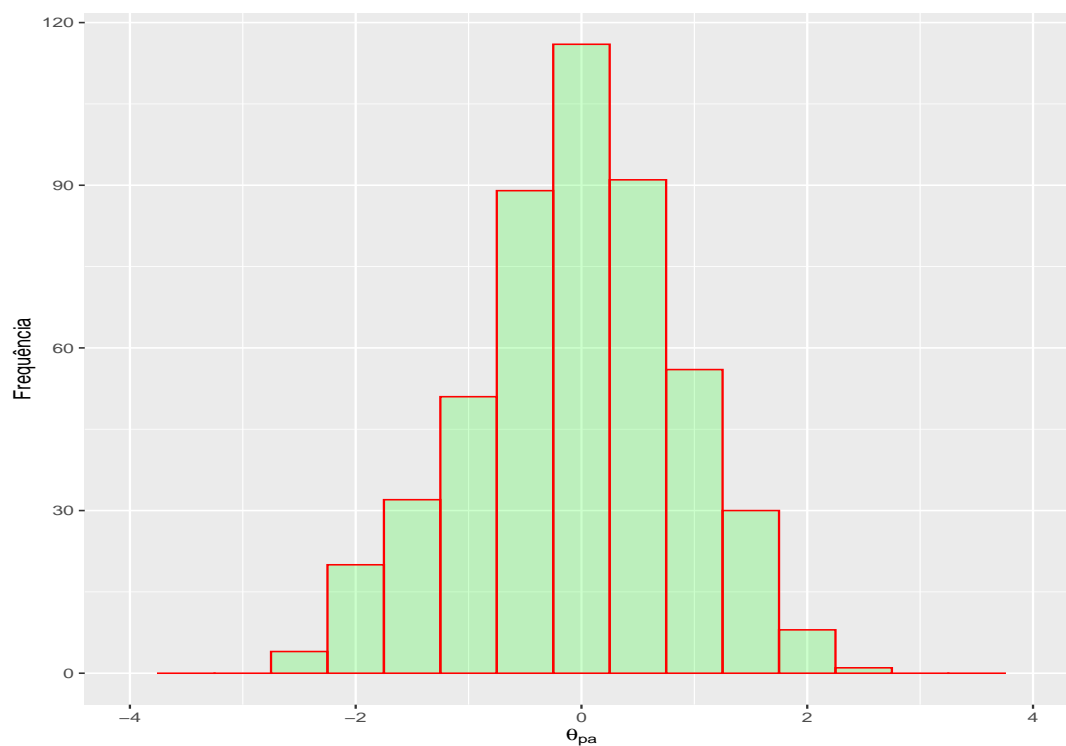


Figura 23 – Histograma dos valores estimados dos traços latentes das enfermeiras relacionados com a realização pessoal reduzida do modelo de resposta gradual multiunidimensional.

Fonte: Elaborada pelo autor.

DISCUSSÃO

Este trabalho se insere no contexto da Teoria de Resposta ao Item, em que se considera o modelo de resposta gradual unidimensional e o modelo de resposta gradual multiunidimensional. Esse é um modelo para dados ordinais amplamente conhecido na literatura. O estudo deste tipo de modelo é de grande importância tanto para o campo teórico quanto para o prático. Atualmente diversas áreas do conhecimento, como educacional, psicológica e comportamental, utilizam variáveis ordinais e modelos que possam incorporá-las.

O presente trabalho teve como objetivo estudar o modelo de resposta gradual, motivado pelos dados da síndrome de *burnout* em enfermeiras, e para isso realizou-se um estudo de simulação para comparar métodos de estimação para os parâmetros do modelo. No estudo de simulação realizado, considerou-se diferentes cenários para estimação dos parâmetros dos itens e dos traços latentes dos modelos de resposta gradual unidimensional e multiunidimensional. A comparação foi realizada levando-se em conta métodos das abordagens clássica e Bayesiana, que são o método da máxima verossimilhança marginal e o método Monte Carlo via Cadeia de Markov amostrador de Gibbs. Além disso, também levou-se em conta diferentes tamanhos amostrais, tamanhos do teste e diferentes níveis de correlação entre os traços latentes, no caso do modelo multiunidimensional.

O estudo de simulação revela que, de uma maneira geral, o método Monte Carlo via Cadeia de Markov amostrador de Gibbs apresenta os melhores resultados do erro quadrático médio quando o parâmetro de discriminação a do modelo de resposta gradual unidimensional é observado, porém o método da máxima verossimilhança marginal apresenta bons resultados quando o viés médio é verificado. O algoritmo amostrador de Gibbs em geral produz os melhores resultados para os parâmetros de dificuldade.

No que tange a estimação do traço latente do indivíduo no modelo de resposta gradual unidimensional, o estudo de simulação revelou que os métodos apresentam resultados próximos em grande amostras. De uma maneira geral, nota-se que o método amostrador de Gibbs apresenta

os melhores resultados.

Os resultados do estudo de simulação para o modelo de resposta gradual multiunidimensional mostram que, em geral, a estimação dos parâmetros dos itens e dos indivíduos pelo método amostrador de Gibbs tendem a melhorar à medida que o tamanho do teste e o tamanho amostral aumentam.

Na aplicação do modelo de resposta gradual unidimensional aos dados relacionados à síndrome de *burnout* em enfermeiras, considerando tanto o método Monte Carlo via Cadeia de Markov amostrador de Gibbs quanto o método da máxima verossimilhança marginal na estimação dos parâmetros do modelo, é possível avaliar os itens do questionário MBI, além das estimativas dos traços latentes das enfermeiras em cada dimensão de *burnout*. Já na aplicação do modelo de resposta gradual multiunidimensional, também é possível avaliar os itens do questionário MBI, além das estimativas dos traços latentes das enfermeiras, agora considerando a existência de associação entre as dimensões da síndrome de *burnout*.

Destacamos a importância e relevância de estudos e planejamentos relacionados a profissionais da saúde que podem dar suporte a ações em nível político e público, o que pode ter impacto positivo nos resultados e segurança do paciente. No contexto atual da pandemia de COVID-19, é possível notar mais claramente a necessidade do bem-estar dos profissionais de saúde. Por isso a investigação, prevenção e tratamento de síndromes psicológicas, como, por exemplo, a síndrome de *burnout* são estratégias para melhorar o sistema de saúde.

A proposta futura para este trabalho é considerar outros cenários do estudo de simulação, como, por exemplo, diferentes tamanhos amostrais e outros métodos de estimação, como, por exemplo, o algoritmo No-U-Turn Sampler (NUTS), que é uma extensão do algoritmo Monte Carlo Hamiltoniano (HOFFMAN; GELMAN, 2014). Também podemos verificar, através de critérios de comparação de modelos, qual modelo seria melhor ajustado para o conjunto de dados da síndrome de *burnout*. Outra ideia é estudar a Teoria de Resposta ao Item adaptado em experimentos da área de Aprendizagem de Máquina, como apresentado em [Martínez-Plumed et al. \(2019\)](#) e [Martínez-Plumed et al. \(2016\)](#).

REFERÊNCIAS

- ACKERMAN, T. A.; GIERL, M. J.; WALKER, C. M. Using multidimensional item response theory to evaluate educational and psychological tests. **Educational Measurement: Issues and Practice**, v. 22, n. 3, p. 37–51, 2003. Citado na página 3.
- ANDERSEN, E. B. A goodness of fit test for the Rasch model. **Psychometrika**, v. 38, n. 1, p. 123–140, 1973. Citado na página 4.
- _____. **Discrete statistical models with social science applications**. Amsterdam: North-Holland Publishing Company, 1980. Citado na página 29.
- ANDRADE, D. F. de; TAVARES, H. R.; VALLE, R. da C. **Teoria da Resposta ao Item: conceitos e aplicações**. São Paulo: ABE - Associação Brasileira de Estatística, 2000. Citado 4 vezes nas páginas 1, 3, 15 e 29.
- ANDRICH, D. A rating formulation for ordered response categories. **Psychometrika**, v. 43, n. 4, p. 561–573, 1978. Citado 2 vezes nas páginas 2 e 21.
- ANKENMANN, R. D.; STONE, C. A. A Monte Carlo study of marginal maximum likelihood parameter estimates for the graded model. Artigo apresentado na reunião anual do NCME - The National Council on Measurement in Education, San Francisco - CA, 1992. Citado na página 2.
- AZEVEDO, C. L. N. **Métodos de estimação na teoria de resposta ao item**. Dissertação (Mestrado) — Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2003. Citado na página 29.
- BAKER, F. B. Equating tests under the graded response model. **Applied Psychological Measurement**, v. 16, n. 1, p. 87–96, 1992. Citado na página 2.
- BERGER, J. O. **Statistical decision theory and Bayesian analysis**. New York: Springer Science & Business Media, 1985. Citado na página 35.
- BERNARDO, J. M.; SMITH, A. F. **Bayesian theory**. New York: Wiley series in probability and mathematical, 2000. Citado na página 35.
- BIRNBAUM, A. Some latent trait models and their use in inferring an examinee's ability. In: LORD, F. M.; NOVICK, M. R. (Ed.). **Statistical theories of mental test scores**. Reading, Mass.: Addison-Wesley, 1968. Citado na página 2.
- BOCK, R. D. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. **Psychometrika**, v. 37, n. 1, p. 29–51, 1972. Citado 2 vezes nas páginas 2 e 21.
- BOCK, R. D.; AITKIN, M. Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. **Psychometrika**, v. 46, n. 4, p. 443–459, 1981. Citado 3 vezes nas páginas 4, 5 e 29.

- BOCK, R. D.; LIEBERMAN, M. Fitting a response model for n dichotomously scored items. **Psychometrika**, v. 35, n. 2, p. 179–197, 1970. Citado 2 vezes nas páginas 4 e 29.
- BOCK, R. D.; ZIMOWSKI, M. F. Multiple group IRT. In: LINDEN, W. J. van der; HAMBLETON, R. K. (Ed.). **Handbook of modern item response theory**. New York: Springer, 1997. p. 433–448. Citado na página 32.
- BOLFARINE, H.; WILTON, B. **Elementos de amostragem**. São Paulo: Blucher, 2005. Citado na página 43.
- BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. New York: Addison-Wesley, 1973. Citado na página 35.
- CARLIN, B. P.; LOUIS, T. A. **Bayes and empirical Bayes methods for data analysis**. Second edition. New York: Chapman & Hall, 2001. Citado na página 35.
- CASABIANCA, J. M.; LEWIS, C. IRT item parameter recovery with marginal maximum likelihood estimation using loglinear smoothing models. **Journal of Educational and Behavioral Statistics**, v. 40, n. 6, p. 547–578, 2015. Citado na página 4.
- CHANG, M.-I. **A Comparison of two MCMC algorithms for estimating the 2PL IRT models**. Tese (Doutorado) — Department of Counseling, Quantitative Methods and Special Education, Southern Illinois University, Carbondale, 2017. Citado 2 vezes nas páginas 5 e 35.
- COHEN, A. S.; KIM, S.-H.; BAKER, F. B. Detection of differential item functioning in the graded response model. **Applied psychological measurement**, v. 17, n. 4, p. 335–350, 1993. Citado na página 2.
- CORNFIELD, J. The bayesian outlook and its application. **Biometrics**, v. 25, p. 617–657, 1969. Citado na página 35.
- CURTIS, S. M. Bugs code for item response theory. **Journal of Statistical Software**, Foundation for Open Access Statistics, v. 36, n. 1, p. 1–34, 2010. Citado na página 36.
- DEMARS, C. E. Recovery of graded response and partial credit parameters in MULTILOG and PARSCALE. **Artigo apresentado no encontro anual da American Educational Research Association**, 2002. Citado na página 2.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the royal statistical society**, v. 39, n. 1, p. 1–38, 1977. Citado 2 vezes nas páginas 4 e 31.
- DRASGOW, F. An evaluation of marginal maximum likelihood estimation for the two-parameter logistic model. **Applied Psychological Measurement**, v. 13, n. 1, p. 77–90, 1989. Citado na página 4.
- DRASGOW, F.; PARSONS, C. K. Application of unidimensional item response theory models to multidimensional data. **Applied Psychological Measurement**, v. 7, n. 2, p. 189–199, 1983. Citado na página 3.
- DUANE, S.; KENNEDY, A. D.; PENDLETON, B. J.; ROWETH, D. Hybrid Monte Carlo. **Physics letters B**, v. 195, n. 2, p. 216–222, 1987. Citado na página 37.

EHLERS, R. S. **Inferência bayesiana**. São Carlos: Departamento de Matemática Aplicada e Estatística/ICMC, Universidade de São Paulo, 2011. Citado na página 34.

EVANS, M.; SWARTZ, T. Methods for approximating integrals in statistics with special emphasis on bayesian integration problems. **Statistical science**, v. 10, n. 3, p. 254–272, 1995. Citado na página 35.

FERRANDO, P. J. A graded response model for measuring person reliability. **British Journal of Mathematical and Statistical Psychology**, v. 62, n. 3, p. 641–662, 2009. Citado na página 2.

FOX, J.-P. **Bayesian item response modeling: Theory and applications**. New York: Springer, 2010. Citado na página 5.

FREUDENBERGER, H. J. Staff burn-out. **Journal of social issues**, v. 30, n. 1, p. 159–165, 1974. Citado na página 8.

GAMERMAN, D.; MIGON, H. dos S. **Inferência estatística: uma abordagem integrada**. Rio de Janeiro: Instituto de Matemática, Universidade Federal do Rio de Janeiro, 1993. Citado na página 35.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian data analysis**. Second edition. New York: Chapman & Hall, 2004. Citado na página 35.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical science**, v. 7, n. 4, p. 457–511, 1992. Citado na página 42.

GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 6, n. 6, p. 721–741, 1984. Citado 2 vezes nas páginas 5 e 83.

GILKS, W. R.; THOMAS, A.; SPIEGELHALTER, D. J. A language and program for complex Bayesian modelling. **Journal of the Royal Statistical Society**, v. 43, n. 1, p. 169–177, 1994. Citado na página 36.

HAMBLETON, R. K.; SWAMINATHAN, H.; ROGERS, H. J. **Fundamentals of item response theory**. Newbury Park: Sage Publications, 1991. Citado na página 15.

HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. **Biometrika**, v. 57, n. 1, p. 97–109, 1970. Citado 2 vezes nas páginas 35 e 83.

HILDEBRAND, F. B. **Introduction to numerical analysis**. Second edition. New York: Dover Publications, 1987. Citado na página 30.

HOFFMAN, M. D.; GELMAN, A. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. **Journal of Machine Learning Research**, v. 15, n. 1, p. 1593–1623, 2014. Citado 2 vezes nas páginas 37 e 72.

HOLZINGER, K. J.; SWINEFORD, F. The bi-factor method. **Psychometrika**, Springer, v. 2, n. 1, p. 41–54, 1937. Citado na página 3.

ISAACSON, E.; KELLER, H. B. **Analysis of Numerical Methods**. New York: Wiley & Sons, 1966. Citado na página 32.

- KACMAR, K. M.; FARMER, W. L.; ZIVNUSKA, S.; WITT, L. Applying multidimensional item response theory analysis to a measure of meta-perspective performance. **Electronic Journal of Business Research Methods**, v. 4, n. 1, p. 23–30, 2006. Citado na página 3.
- KELDERMAN, H.; RIJKES, C. P. Loglinear multidimensional IRT models for polytomously scored items. **Psychometrika**, v. 59, n. 2, p. 149–176, 1994. Citado na página 3.
- KIEFTENBELD, V.; NATESAN, P. Recovery of graded response model parameters: A comparison of marginal maximum likelihood and Markov chain Monte Carlo estimation. **Applied Psychological Measurement**, v. 36, n. 5, p. 399–419, 2012. Citado 5 vezes nas páginas 2, 5, 36, 40 e 43.
- KIM, J.-S.; BOLT, D. M. Estimating item response theory models using Markov chain Monte Carlo methods. **Educational Measurement: Issues and Practice**, v. 26, n. 4, p. 38–51, 2007. Citado na página 5.
- KIM, S.-H. An evaluation of a Markov chain Monte Carlo method for the Rasch model. **Applied Psychological Measurement**, v. 25, n. 2, p. 163–176, 2001. Citado na página 5.
- KOCH, W. R. Likert scaling using the graded response latent trait model. **Applied Psychological Measurement**, v. 7, n. 1, p. 15–32, 1983. Citado na página 2.
- KOSE, I. A.; DEMIRTASLI, N. C. Comparison of unidimensional and multidimensional models based on item response theory in terms of both variables of test length and sample size. **Procedia-Social and Behavioral Sciences**, v. 46, p. 135–140, 2012. Citado na página 3.
- KUO, T.-C.; SHENG, Y. A comparison of estimation methods for a multi-unidimensional graded response IRT model. **Frontiers in psychology**, v. 7:880, p. 1–12, 2016. Citado 2 vezes nas páginas 2 e 3.
- LAHUIS, D. M.; CLARK, P.; O'BRIEN, E. An examination of item response theory item fit indices for the graded response model. **Organizational research methods**, v. 14, n. 1, p. 10–23, 2011. Citado na página 2.
- LAKE, E. T. Development of the practice environment scale of the nursing work index. **Research in nursing & health**, v. 25, n. 3, p. 176–188, 2002. Citado na página 10.
- LI, Y.; JIAO, H.; LISSITZ, R. W. Applying multidimensional item response theory models in validating test dimensionality: An example of k–12 large-scale science assessment. **Journal of Applied Testing Technology**, v. 13, n. 2, 2012. Citado na página 3.
- LIKERT, R. A technique for the measurement of attitudes. **Archives of psychology**, 1932. Citado 2 vezes nas páginas 2 e 9.
- LORD, F. M. The relation of the reliability of multiple-choice tests to the distribution of item difficulties. **Psychometrika**, v. 17, n. 2, p. 181–194, 1952. Citado na página 1.
- _____. **A theory of test scores**. Richmond: Psychometric corporation, 1952. Citado na página 1.
- _____. The relation of test score to the trait underlying the test. **Educational and Psychological Measurement**, v. 13, n. 4, p. 517–549, 1953. Citado na página 1.
- _____. **Applications of item response theory to practical testing problems**. Hilldale: Lawrence Erlbaum Associates, 1980. Citado na página 15.

LUNN, D. J.; THOMAS, A.; BEST, N.; SPIEGELHALTER, D. WinBUGS - a Bayesian modeling framework: concepts, structure, and extensibility. **Statistics and computing**, v. 10, n. 4, p. 325–337, 2000. Citado na página 37.

MARCOULIDES, K. M. Careful with those priors: A note on Bayesian estimation in two-parameter logistic item response theory models. **Measurement: Interdisciplinary Research and Perspectives**, v. 16, n. 2, p. 92–99, 2018. Citado na página 5.

MARTELLI, I. **Multidimensional item response theory models with general and specific latent traits for ordinal data**. Tese (Doutorado) — Alma Mater Studiorum - Università di Bologna, Bologna, 2014. Citado 3 vezes nas páginas 3, 25 e 27.

MARTELLI, I.; MATTEUCCI, M.; MIGNANI, S. Bayesian estimation of a multidimensional additive graded response model for correlated traits. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 45, n. 5, p. 1636–1654, 2016. Citado na página 3.

MARTÍNEZ-PLUMED, F.; PRUDÊNCIO, R. B.; MARTÍNEZ-USÓ, A.; HERNÁNDEZ-ORALLO, J. Making sense of item response theory in machine learning. In: IOS PRESS. **Proceedings of the Twenty-second European Conference on Artificial Intelligence**. [S.l.], 2016. p. 1140–1148. Citado na página 72.

_____. Item response theory in ai: Analysing machine learning classifiers at the instance level. **Artificial Intelligence**, Elsevier, v. 271, p. 18–42, 2019. Citado na página 72.

MASLACH, C.; JACKSON, S. E.; LEITER, M. P. The Maslach burnout inventory manual. In: ZALAQUETT, C. P.; WOOD, R. J. (Ed.). **Evaluating stress: A book of resources**. London: The Scarecrow Press, 1996. Citado 2 vezes nas páginas 8 e 9.

MASTERS, G. N. A Rasch model for partial credit scoring. **Psychometrika**, v. 47, n. 2, p. 149–174, 1982. Citado 2 vezes nas páginas 2 e 21.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. **The journal of chemical physics**, v. 21, n. 6, p. 1087–1092, 1953. Citado na página 35.

MISLEVY, R. J. Bayes modal estimation in item response models. **Psychometrika**, v. 51, n. 2, p. 177–195, 1986. Citado 2 vezes nas páginas 4 e 35.

MURAKI, E. A generalized partial credit model: Application of an em algorithm. **Applied Psychological Measurement**, v. 16, n. 2, p. 159–176, 1992. Citado 2 vezes nas páginas 2 e 21.

NATESAN, P.; NANDAKUMAR, R.; MINKA, T.; RUBRIGHT, J. D. Bayesian prior choice in IRT estimation using MCMC and variational Bayes. **Frontiers in psychology**, v. 7, p. 1422, 2016. Citado na página 5.

OGASAWARA, H. Marginal maximum likelihood estimation of item response theory (IRT) equating coefficients for the common-examinee design. **Japanese Psychological Research**, v. 43, n. 2, p. 72–82, 2001. Citado na página 4.

O’HAGAN, A.; FORSTER, J. J. **Kendall’s advanced theory of statistics**. London: Arnold, 2004. v. 2B: Bayesian inference. Citado na página 35.

OSHIMA, T.; RAJU, N. S.; FLOWERS, C. P. Development and demonstration of multidimensional IRT-based internal measures of differential functioning of items and tests. **Journal of Educational Measurement**, v. 34, n. 3, p. 253–272, 1997. Citado na página 3.

PATZ, R. J.; JUNKER, B. W. A straightforward approach to Markov chain Monte Carlo methods for item response models. **Journal of educational and behavioral Statistics**, v. 24, n. 2, p. 146–178, 1999. Citado na página 5.

PLUMMER, M. JAGS: A program for analysis of bayesian graphical models using gibbs sampling. In: HORNIK, K.; LEISCH, F.; ZEILEIS, A. (Ed.). **Proceedings of the 3rd international workshop on distributed statistical computing**. Viena, Austria: [s.n.], 2003. v. 124. Citado na página 37.

_____. rjags: Bayesian graphical models using MCMC. 2016. R package version 4-6. Disponível em: <<https://CRAN.R-project.org/package=rjags>>. Citado na página 37.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2018. Disponível em: <<https://www.R-project.org/>>. Citado 2 vezes nas páginas 34 e 37.

RAO, C. R. **Linear statistical inference and its applications**. New York: Wiley, 1973. Citado na página 32.

RASCH, G. Probabilistic models for some intelligence and achievement tests. **Danish Institute for Educational Research**, 1960. Citado 2 vezes nas páginas 1 e 2.

RECKASE, M. D. The difficulty of test items that measure more than one ability. **Applied psychological measurement**, v. 9, n. 4, p. 401–412, 1985. Citado na página 3.

_____. The past and future of multidimensional item response theory. **Applied Psychological Measurement**, v. 21, n. 1, p. 25–36, 1997. Citado na página 2.

_____. **Multidimensional item response theory models**. New York: Springer, 2009. Citado 4 vezes nas páginas 13, 21, 20 e 24.

REISE, S. P. The rediscovery of bifactor measurement models. **Multivariate behavioral research**, Taylor & Francis, v. 47, n. 5, p. 667–696, 2012. Citado na página 3.

REISE, S. P.; MORIZOT, J.; HAYS, R. D. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. **Quality of Life Research**, Springer, v. 16, n. 1, p. 19–31, 2007. Citado na página 3.

REISE, S. P.; REVICKI, D. A. **Handbook of item response theory modeling: Applications to typical performance assessment**. New York: Routledge, 2014. Citado na página 1.

REISE, S. P.; WIDAMAN, K. F.; PUGH, R. H. Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. **The American Psychological Association**, v. 114, n. 3, p. 552, 1993. Citado na página 2.

REISE, S. P.; YU, J. Parameter recovery in the graded response model using multilog. **Journal of Educational Measurement**, v. 27, n. 2, p. 133–144, 1990. Citado na página 2.

RIBAS, C. d. C. S. d. **Síndrome de burnout em profissionais de saúde: uma abordagem bioé-tica num estudo preliminar**. Dissertação (Mestrado) — Faculdade de Medicina - Universidade do Porto, Porto, 2011. Citado na página 8.

RIZOPOULOS, D. ltm: An R package for latent variable modelling and item response theory analyses. **Journal of Statistical Software**, v. 17, n. 5, p. 1–25, 2006. Disponível em: <<http://www.jstatsoft.org/v17/i05/>>. Citado na página 34.

SAHU, S. K. Bayesian estimation and model choice in item response models. **Journal of Statistical Computation and Simulation**, v. 72, n. 3, p. 217–232, 2002. Citado na página 5.

SAMEJIMA, F. Estimation of latent ability using a pattern of graded scores. **Psychometrika - Monograph Supplement**, v. 34, n. 1, p. 1–97, 1969. Citado 2 vezes nas páginas 2 e 21.

SAMEJIMA, F. A. Graded response model. In: LINDEN, W. J. van der; HAMBLETION, R. K. (Ed.). **Handbook of Modern Item Response Theory**. New York: Springer, 1997. p. 85–100. Citado na página 32.

SERMEUS, W.; AIKEN, L. H.; HEEDE, K. Van den; RAFFERTY, A. M.; GRIFFITHS, P.; MORENO-CASBAS, M. T.; BUSSE, R.; LINDQVIST, R.; SCOTT, A. P.; BRUYNEEL, L.; BRZOSTEK, T.; KINNUNEN, J.; SCHUBERT, M.; SCHOONHOVEN, L.; ZIKOS, D. Nurse forecasting in europe (RN4CAST): rationale, design and methodology. **BMC nursing**, v. 10, n. 1, p. 6, 2011. Citado 2 vezes nas páginas 7 e 8.

SHENG, Y.; WIKLE, C. K. Comparing multiunidimensional and unidimensional item response theory models. **Educational and Psychological Measurement**, v. 67, n. 6, p. 899–919, 2007. Citado na página 3.

_____. Bayesian multidimensional irt models with a hierarchical structure. **Educational and psychological measurement**, Sage Publications Sage CA: Los Angeles, CA, v. 68, n. 3, p. 413–430, 2008. Citado na página 3.

SILVA, M. A. da; BAZÁN, J. L.; HUGGINS-MANLEY, A. C. Sensitivity analysis and choosing between alternative polytomous IRT models using bayesian model comparison criteria. **Communications in Statistics - Simulation and Computation**, 2017. Citado 3 vezes nas páginas 2, 36 e 41.

Stan Development Team. Stan: A C++ library for probability and sampling. 2014. Disponível em: <<http://mc-stan.org>>. Citado na página 37.

_____. RStan: the R interface to Stan. 2018. R package version 2.17.3. Disponível em: <<http://mc-stan.org/>>. Citado na página 37.

STURTZ, S.; LIGGES, U.; GELMAN, A. R2WinBUGS: A package for running WinBUGS from R. **Journal of Statistical Software**, v. 12, n. 3, p. 1–16, 2005. Disponível em: <<http://www.jstatsoft.org>>. Citado na página 37.

_____. R2OpenBUGS: a package for running OpenBUGS from R. URL <http://cran.rproject.org/web/packages/R2OpenBUGS/vignettes/R2OpenBUGS.pdf>, 2010. Citado na página 37.

SU, Y.-S.; YAJIMA, M. R2jags: Using r to run 'JAGS'. 2015. R package version 0.5-7. Disponível em: <<https://CRAN.R-project.org/package=R2jags>>. Citado na página 37.

SVETINA, D.; VALDIVIA, A.; UNDERHILL, S.; DAI, S.; WANG, X. Parameter recovery in multidimensional item response theory models under complexity and nonnormality. **Applied Psychological Measurement**, v. 41, n. 7, p. 530–544, 2017. Citado na página 3.

SWAMINATHAN, H.; GIFFORD, J. A. Bayesian estimation in the Rasch model. **Journal of Educational Statistics**, v. 7, n. 3, p. 175–191, 1982. Citado na página 4.

_____. Bayesian estimation in the two-parameter logistic model. **Psychometrika**, v. 50, n. 3, p. 349–364, 1985. Citado na página 4.

_____. Bayesian estimation in the three-parameter logistic model. **Psychometrika**, v. 51, n. 4, p. 589–601, 1986. Citado na página 4.

THOMAS, A.; HARA, B.; LIGGES, U.; STURTZ, S. Making BUGS Open. **R news**, v. 6, p. 12–17, 2006. Citado na página 37.

TOLAND, M. D.; SULIS, I.; GIAMBONA, F.; PORCU, M.; CAMPBELL, J. M. Introduction to bifactor polytomous item response theory analysis. **Journal of school psychology**, Elsevier, v. 60, p. 41–63, 2017. Citado na página 3.

TORRE, J. D. L.; PATZ, R. J. Making the most of what we have: A practical application of multidimensional item response theory in test scoring. **Journal of Educational and Behavioral Statistics**, v. 30, n. 3, p. 295–311, 2005. Citado na página 3.

TSUTAKAWA, R. K.; LIN, H. Y. Bayesian estimation of item response curves. **Psychometrika**, v. 51, n. 2, p. 251–267, 1986. Citado na página 4.

VINCENZI, S. L.; POSSAN, E.; ANDRADE, D. F. de; PITUCO, M. M.; SANTOS, T. de O.; JASSE, E. P. Assessment of environmental sustainability perception through item response theory: A case study in brazil. **Journal of Cleaner Production**, Elsevier, v. 170, p. 1369–1386, 2018. Citado 2 vezes nas páginas 22 e 47.

WALKER, C. M.; BERETVAS, S. N. Comparing multidimensional and unidimensional proficiency classifications: Multidimensional IRT as a diagnostic aid. **Journal of Educational Measurement**, v. 40, n. 3, p. 255–275, 2003. Citado na página 3.

WETZEL, E.; HELL, B. Multidimensional item response theory models in vocational interest measurement: An illustration using the aist-r. **Journal of Psychoeducational Assessment**, v. 32, n. 4, p. 342–355, 2014. Citado na página 3.

WOLLACK, J. A.; BOLT, D. M.; COHEN, A. S.; LEE, Y.-S. Recovery of item parameters in the nominal response model: A comparison of marginal maximum likelihood estimation and markov chain monte carlo estimation. **Applied psychological measurement**, v. 26, n. 3, p. 339–352, 2002. Citado na página 35.

WRIGHT, B. D. Sample-free test calibration and person measurement. **Proceedings of the 1967 invitational conference on testing problems**, p. 85–101, 1968. Citado na página 2.

ZHANG, B.; STONE, C. A. Evaluating item fit for multidimensional item response models. **Educational and Psychological Measurement**, v. 68, n. 2, p. 181–196, 2008. Citado 2 vezes nas páginas 3 e 24.

ZHOU, X. Hierarchical item response models for analyzing public opinion. **Political Analysis**, Cambridge University Press, v. 27, n. 4, p. 481–502, 2019. Citado na página 3.

ZHU, X.; STONE, C. A. Bayesian comparison of alternative graded response models for performance assessment applications. **Educational and Psychological Measurement**, v. 72, n. 5, p. 774–799, 2012. Citado na página 36.

MÉTODO DE MONTE CARLO VIA CADEIAS DE MARKOV

Métodos MCMC são uma classe de algoritmos para amostragem a partir de uma distribuição de probabilidade (por exemplo, a distribuição a posteriori) com base na construção de uma cadeia de Markov. Monte Carlo é uma técnica de simulação e cadeia de Markov é uma sequência de variáveis aleatórias, $\{X_0, X_1, X_2, \dots\}$, com distribuição $P(X_t|X_{t-1})$, de forma que, a distribuição de X_t dados seus valores anteriores $\{X_0, X_1, X_2, \dots, X_{t-1}\}$ depende apenas do seu estado anterior X_{t-1} , isto é, $P(X_t|X_0, X_1, X_2, \dots, X_{t-1}) = P(X_t|X_{t-1})$.

Os métodos MCMC requerem que a cadeia seja:

homogênea: as probabilidades de transição de um estado para outro são invariantes;

irredutível: cada estado pode ser atingido a partir de qualquer outro em um número finito de iterações;

aperiódica: não haja estados absorventes, isto é, não haja estados que se a cadeia entrar neste estado, o processo nunca irá deixá-lo.

A estacionariedade é um importante conceito da cadeia de Markov, que permite definir a probabilidade de cada estado de um sistema em um tempo aleatório. Dessa forma, em cada estado da cadeia de Markov, amostras aleatórias de parâmetros do modelo são geradas a partir da distribuição das amostras geradas no estado anterior. Como as amostras iniciais podem ser afetadas pelos valores iniciais, elas são descartadas, como se formassem uma amostra de aquecimento, no chamado *burn-in* ou *warm-up*. Após o estágio de *burn-in*, a qualidade da amostra torna-se aproximadamente estável.

O amostrador de Gibbs ([GEMAN; GEMAN, 1984](#)) e o algoritmo Metropolis-Hastings ([HASTINGS, 1970](#)) são métodos MCMC clássicos que utilizam a noção de passeio aleatório,

isto indica que, a cada passo, a direção do movimento é aleatória. A ideia é simular um passeio aleatório no espaço paramétrico que converge para uma distribuição estacionária, que é a distribuição de interesse no problema. Esses algoritmos são revisados a seguir.

A.1 O algoritmo Metropolis-Hastings

A ideia do algoritmo Metropolis-Hastings é gerar valores aleatórios de uma distribuição alvo por meio de uma distribuição auxiliar, de forma que gerar valores de uma distribuição auxiliar é mais fácil. Esse mecanismo garante a convergência da cadeia para a distribuição a posteriori. Assim, para gerar valores de uma distribuição a posteriori $P(\boldsymbol{\alpha}|\mathbf{x})$, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)^\top$, uma distribuição condicional auxiliar $h(\cdot|\cdot)$, chamada de *distribuição proposta*, é utilizada. O algoritmo Metropolis-Hastings é descrito no Algoritmo 1

Algoritmo 1: Metropolis-Hastings

1. Especifique valores iniciais para os parâmetros $\boldsymbol{\alpha}^{(0)} = (\alpha_1^{(0)}, \alpha_2^{(0)}, \dots, \alpha_p^{(0)})^\top$
2. Para $t = 1, \dots, N$
 - a) Gere um novo valor $\boldsymbol{\alpha}^{*(t)}$ da distribuição $h(\boldsymbol{\alpha}^{*(t)}|\boldsymbol{\alpha}^{(t-1)})$
 - b) Calcule a probabilidade de aceitação $\gamma = \min \left\{ \frac{P(\boldsymbol{\alpha}=\boldsymbol{\alpha}^{*(t)}|\mathbf{x})h(\boldsymbol{\alpha}^{(t-1)}|\boldsymbol{\alpha}^{*(t)})}{P(\boldsymbol{\alpha}=\boldsymbol{\alpha}^{(t-1)}|\mathbf{x})h(\boldsymbol{\alpha}^{*(t)}|\boldsymbol{\alpha}^{(t-1)})}, 1 \right\}$
 - c) Gere $u \sim \text{Uniforme}(0, 1)$
 - d) Defina:

$$\boldsymbol{\alpha}^{(t)} = \begin{cases} \boldsymbol{\alpha}^{*(t)} & \text{se } u \leq \gamma \\ \boldsymbol{\alpha}^{(t-1)} & \text{se } u > \gamma \end{cases} .$$

Uma característica importante do algoritmo é que para o cálculo da probabilidade de aceitação, γ , apenas é necessário conhecer a posteriori $P(\boldsymbol{\alpha}|\mathbf{x})$ parcialmente, ou seja, é possível calcular γ utilizando (4.11), uma vez que pode não ser possível calcular a posteriori por (4.10).

A.2 O amostrador de Gibbs

O amostrador de Gibbs não utiliza o mecanismo de aceitação-rejeição, de forma que a cadeia sempre irá se mover para um novo valor. A ação de se mover de um estado para o outro é realizada de acordo com *distribuições condicionais completas* $P(\alpha_i|\boldsymbol{\alpha}_{-i})$, em que $\boldsymbol{\alpha}_{-i} = (\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_p)^\top$. A distribuição condicional completa é a distribuição da i -ésima componente de $\boldsymbol{\alpha}$ condicionada em todas as outras componentes, e é obtida a partir da distribuição conjunta

$$P(\alpha_i|\boldsymbol{\alpha}_{-i}) = \frac{P(\boldsymbol{\alpha})}{\int P(\boldsymbol{\alpha})d\alpha_i} . \quad (\text{A.1})$$

Para obter a distribuição condicional completa da posteriori do modelo de resposta gradual, considera-se que

$$\boldsymbol{\alpha} = (\boldsymbol{\theta}, \boldsymbol{\xi})^\top$$

isto é

$$\boldsymbol{\alpha} = (\theta_1, \theta_2, \dots, \theta_n, a_1, a_2, \dots, a_I, b_{11}, \dots, b_{1m_1}, \dots, b_{I1}, \dots, b_{Im_I})^\top$$

e a distribuição condicional completa definida em (A.1) é dada por

$$P(\alpha_i | \mathbf{Y}_{\dots}, \boldsymbol{\alpha}_{-i}) = \frac{P(\boldsymbol{\alpha} | \mathbf{y}_{\dots})}{\int P(\boldsymbol{\alpha} | \mathbf{y}_{\dots}) d\alpha_i}.$$

Assim, para gerar valores da distribuição a posteriori (4.13), pode-se utilizar as distribuições condicionais completas, caso forem completamente conhecidas. O amostrador de Gibbs utiliza esta técnica, e é definido no Algoritmo 2.

Cada iteração se completa após p movimentos nas componentes do vetor de parâmetros $\boldsymbol{\alpha}$, e após a convergência, os valores resultantes formam uma amostra de $P(\boldsymbol{\alpha} | \mathbf{Y}_{\dots})$, isto é, $P(\boldsymbol{\theta}, \boldsymbol{\xi} | \mathbf{Y}_{\dots})$. O amostrador de Gibbs é um caso especial do algoritmo Metropolis-Hastings, em que os elementos de $\boldsymbol{\alpha}$ são atualizados tomando a distribuição condicional completa como a distribuição proposta e a probabilidade de aceitação igual a 1.

Algoritmo 2: Amostrador de Gibbs

1. Especifique valores iniciais para os parâmetros

$$\boldsymbol{\alpha}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, \boldsymbol{\theta}_2^{(0)}, \dots, \boldsymbol{\theta}_n^{(0)}, a_1^{(0)}, a_2^{(0)}, \dots, a_I^{(0)}, b_{11}^{(0)}, \dots, b_{1m_1}^{(0)}, \dots, b_{I1}^{(0)}, \dots, b_{Im_I}^{(0)})^\top$$

2. Para $t = 1, \dots, N$

- a) Obtenha um novo valor $\boldsymbol{\alpha}^{(t)}$ a partir de $\boldsymbol{\alpha}^{(t-1)}$ através da geração sucessiva dos valores

$$\boldsymbol{\theta}_1^{(t)} \sim P(\boldsymbol{\theta}_1 | \mathbf{y}_{\dots}, \boldsymbol{\theta}_2^{(t-1)}, \dots, \boldsymbol{\theta}_n^{(t-1)}, a_1^{(t-1)}, \dots, a_I^{(t-1)}, b_{11}^{(t-1)}, \dots, b_{1m_1}^{(t-1)}, \dots, b_{I1}^{(t-1)}, \dots, b_{Im_I}^{(t-1)})$$

$$\vdots$$

$$\boldsymbol{\theta}_n^{(t)} \sim P(\boldsymbol{\theta}_n | \mathbf{y}_{\dots}, \boldsymbol{\theta}_1^{(t-1)}, \dots, \boldsymbol{\theta}_{n-1}^{(t-1)}, a_1^{(t-1)}, \dots, a_I^{(t-1)}, b_{11}^{(t-1)}, \dots, b_{1m_1}^{(t-1)}, \dots, b_{I1}^{(t-1)}, \dots, b_{Im_I}^{(t-1)})$$

$$a_1^{(t)} \sim P(a_1 | \mathbf{y}_{\dots}, \boldsymbol{\theta}_1^{(t-1)}, \dots, \boldsymbol{\theta}_n^{(t-1)}, a_2^{(t-1)}, \dots, a_I^{(t-1)}, b_{11}^{(t-1)}, \dots, b_{1m_1}^{(t-1)}, \dots, b_{I1}^{(t-1)}, \dots, b_{Im_I}^{(t-1)})$$

$$\vdots$$

$$a_I^{(t)} \sim P(a_I | \mathbf{y}_{\dots}, \boldsymbol{\theta}_1^{(t-1)}, \dots, \boldsymbol{\theta}_n^{(t-1)}, a_1^{(t-1)}, \dots, a_{I-1}^{(t-1)}, b_{11}^{(t-1)}, \dots, b_{1m_1}^{(t-1)}, \dots, b_{I1}^{(t-1)}, \dots, b_{Im_I}^{(t-1)})$$

$$b_{11}^{(t)} \sim P(b_{11} | \mathbf{y}_{\dots}, \boldsymbol{\theta}_1^{(t-1)}, \dots, \boldsymbol{\theta}_n^{(t-1)}, a_1^{(t-1)}, \dots, a_{I-1}^{(t-1)}, b_{12}^{(t-1)}, \dots, b_{1m_1}^{(t-1)}, \dots, b_{I1}^{(t-1)}, \dots, b_{Im_I}^{(t-1)})$$

$$\vdots$$

$$b_{Im_I}^{(t)} \sim P(b_{Im_I} | \mathbf{y}_{\dots}, \boldsymbol{\theta}_1^{(t-1)}, \dots, \boldsymbol{\theta}_n^{(t-1)}, a_1^{(t-1)}, \dots, a_{I-1}^{(t-1)}, b_{12}^{(t-1)}, \dots, b_{1m_1}^{(t-1)}, \dots, b_{I1}^{(t-1)}, \dots, b_{I(m_I-1)}^{(t-1)})$$

CÓDIGOS

Código-fonte 1: Código JAGS para o modelo de resposta gradual unidimensional

```
1 model {
2
3 #prioris
4
5     #discriminação
6
7     for(j in 1:n_item){
8         alpha[j] ~ dnorm(1,0.16)T(0,)
9     }
10
11    #dificuldade
12
13    for(j in 1:n_item){
14        for(k in 1:(K-1)){
15            bc[j,k] ~ dnorm(0,0.16)
16        }
17        kappa[j,1:(K-1)] <- sort(bc[j,1:(K-1)])
18    }
19
20    #traco latente
21
22    for(i in 1:n_nurse){
23        theta[i] ~ dnorm(0.0,1.0)
24    }
25
26 #probabilidades acumuladas
27
28
29    for(i in 1:n_nurse){
```

```

30         for(j in 1:n_item){
31             for(k in 1:(K-1)){
32                 Pmas[i,j,k] <- ilogit(alpha[j]*(
theta[i]-kappa[j,k]))
33             }
34             #Pmas[i,j,1] <- 1.0
35             Pmas[i,j,K] <- 0.0
36         }
37
38 #probabilidades das categorias
39
40         for(j in 1:n_item){
41             prob[i,j,1] <- 1 - Pmas[i,j,1]
42             for(k in 2:K){
43                 prob[i,j,k] <- Pmas[i,j,(k-1)] -
Pmas[i,j,k]
44             }
45         }
46
47
48         for(j in 1:n_item){
49             Y[i,j] ~ dcat(prob[i,j,1:K])
50         }
51     }
52 }
53
54 }

```

Código-fonte 2: Código JAGS para o modelo de resposta gradual multiunidimensional

```

1 model {
2     Pr.theta[1:m,1:m] <- inverse(S1[,])
3 #priors
4
5     #discriminação
6
7     for(j in 1:p){
8         alpha[j] ~ dnorm(1,0.16)T(0,)
9     }
10
11     #dificuldade
12
13     for(j in 1:p){
14         for(k in 1:(K-1)){
15             kappa.star[j,k] ~ dnorm(0,0.16)
16         }
17         kappa[j,1:(K-1)] <- sort(kappa.star[j,1:(K-1)])

```

```
18     }
19
20     #traco latente
21
22     for(i in 1:n){
23         theta[i,1:m] ~ dmnorm(m1[],Pr.theta[,])
24     }
25
26 #probabilidades acumuladas
27
28
29     for(i in 1:n){
30         for(j in 1:p){
31             for(k in 1:(K-1)){
32                 P[i,j,k] <- ilogit(alpha[j]*(
33 theta[i,vv[j]] - kappa[j,k]))
34             }
35             #Pmas[i,j,1] <- 1.0
36             P[i,j,K] <- 0.0
37         }
38 #probabilidades das categorias
39
40
41         for(j in 1:p){
42             prob[i,j,1] <- 1 - P[i,j,1]
43             for(k in 2:K){
44                 prob[i,j,k] <- P[i,j,(k-1)] -
45 P[i,j,k]
46             }
47         }
48
49         for(j in 1:p){
50             Y[i,j] ~ dcat(prob[i,j,1:K])
51         }
52
53     }
54
55 }
```

RESULTADOS DO ESTUDO DE SIMULAÇÃO

Neste apêndice estão apresentadas as tabelas referentes ao resultado do estudo de simulação descrito no capítulo 5.

Tabela 12 – REQM médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) sob diferentes métodos de estimação, tamanhos amostrais e tamanhos do teste

n	Par	MVM				Gibbs			
		8	12	16	22	8	12	16	22
75	a	0,57	0,52	0,44	0,45	0,40	0,41	0,39	0,38
	b_1	1,47	1,41	1,31	1,12	0,65	0,63	0,65	0,66
	b_2	1,55	1,44	1,45	1,42	0,64	0,70	0,69	0,80
	b_3	1,66	1,50	1,56	1,60	0,69	0,74	0,75	0,85
	b_4	2,04	1,89	1,98	1,90	0,94	0,93	0,95	1,07
	b_5	2,33	2,17	2,22	2,10	1,12	1,08	1,09	1,21
	b_6	2,31	2,20	2,18	2,17	1,83	1,66	1,66	1,78
	θ	0,60	0,56	0,67	0,68	0,44	0,47	0,43	0,50
150	a	0,29	0,29	0,26	0,25	0,32	0,31	0,31	0,33
	b_1	2,05	1,56	1,50	1,36	0,47	0,45	0,43	0,45
	b_2	1,28	1,04	0,94	0,88	0,49	0,45	0,48	0,56
	b_3	1,29	1,09	0,98	1,00	0,54	0,50	0,55	0,63
	b_4	1,85	1,61	1,49	1,47	0,73	0,69	0,72	0,83
	b_5	2,23	1,92	1,79	1,71	0,89	0,85	0,86	1,00
	b_6	2,51	2,11	1,98	1,98	1,58	1,41	1,37	1,53
	θ	0,42	0,39	0,35	0,32	0,42	0,41	0,39	0,41
300	a	0,55	0,49	0,50	0,44	0,27	0,24	0,21	0,23
	b_1	1,80	1,51	1,30	1,12	0,35	0,31	0,30	0,32
	b_2	1,75	1,42	1,75	2,10	0,36	0,33	0,34	0,39
	b_3	2,09	1,73	2,08	2,37	0,41	0,38	0,39	0,44
	b_4	2,82	2,52	2,78	2,73	0,53	0,51	0,51	0,60
	b_5	3,11	2,82	3,05	2,88	0,68	0,63	0,62	0,74
	b_6	3,86	3,34	3,36	2,92	1,27	1,14	1,08	1,22
	θ	0,65	0,63	0,82	0,87	0,42	0,38	0,35	0,36
500	a	0,42	0,34	0,18	0,14	0,19	0,17	0,17	0,17
	b_1	1,61	1,27	1,61	1,48	0,27	0,26	0,24	0,21
	b_2	1,66	1,66	1,05	0,95	0,29	0,26	0,25	0,26
	b_3	1,74	1,80	1,10	1,09	0,32	0,31	0,28	0,31
	b_4	2,60	2,52	1,72	1,71	0,44	0,42	0,39	0,43
	b_5	2,90	2,78	2,06	2,01	0,55	0,51	0,50	0,55
	b_6	3,19	2,90	2,54	2,48	1,11	0,96	0,91	1,00
	θ	0,52	0,59	0,40	0,36	0,41	0,38	0,34	0,33
750	a	0,15	0,12	0,12	0,12	0,14	0,12	0,13	0,14
	b_1	2,26	1,95	1,73	1,46	0,20	0,21	0,20	0,18
	b_2	1,23	1,01	0,92	0,82	0,21	0,20	0,21	0,23
	b_3	1,19	0,98	0,91	0,96	0,27	0,23	0,24	0,27
	b_4	2,15	1,69	1,54	1,57	0,37	0,32	0,32	0,38
	b_5	2,60	2,12	1,93	2,04	0,46	0,40	0,41	0,48
	b_6	3,37	2,74	2,59	2,65	1,00	0,83	0,78	0,87
	θ	0,42	0,37	0,34	0,31	0,41	0,38	0,34	0,32
1000	a	0,30	0,22	0,27	0,32	0,14	0,11	0,12	0,12
	b_1	1,68	1,53	1,14	0,89	0,18	0,15	0,17	0,16
	b_2	1,07	0,91	0,76	0,64	0,16	0,18	0,18	0,19
	b_3	1,27	1,02	0,99	1,03	0,19	0,21	0,21	0,23
	b_4	2,15	1,75	1,61	1,73	0,27	0,29	0,29	0,32
	b_5	2,85	2,35	2,15	2,23	0,36	0,36	0,36	0,40
	b_6	3,62	3,07	2,94	2,73	0,88	0,77	0,71	0,79
	θ	0,44	0,40	0,37	0,39	0,40	0,38	0,34	0,32

Tabela 13 – Viés médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) sob diferentes métodos de estimação, tamanhos amostrais e tamanhos do teste

n	Par	MVM				Gibbs			
		8	12	16	22	8	12	16	22
75	a	0,00	-0,01	0,02	0,05	0,23	0,21	0,24	0,24
	b_1	-0,51	-0,33	-0,24	0,02	-0,03	-0,24	-0,28	-0,38
	b_2	0,09	0,09	0,31	0,39	-0,44	-0,56	-0,54	-0,68
	b_3	0,50	0,39	0,60	0,57	-0,52	-0,60	-0,61	-0,75
	b_4	0,82	0,69	0,89	0,62	-0,75	-0,78	-0,80	-0,96
	b_5	0,89	0,74	0,95	0,61	-0,90	-0,89	-0,91	-1,07
	b_6	0,53	0,49	0,69	0,24	-1,65	-1,38	-1,43	-1,57
	θ	0,28	0,23	0,46	0,49	-0,14	-0,27	-0,27	-0,37
150	a	0,06	-0,01	0,03	0,04	0,20	0,17	0,21	0,22
	b_1	-1,40	-0,88	-0,85	-0,65	0,02	-0,10	-0,12	-0,19
	b_2	-0,55	-0,35	-0,29	-0,26	-0,30	-0,33	-0,35	-0,46
	b_3	0,00	0,01	0,01	-0,04	-0,36	-0,38	-0,43	-0,54
	b_4	0,51	0,35	0,35	0,14	-0,55	-0,53	-0,59	-0,73
	b_5	0,70	0,45	0,47	0,14	-0,68	-0,64	-0,70	-0,86
	b_6	0,65	0,41	0,45	-0,02	-1,38	-1,10	-1,15	-1,30
	θ	0,02	0,02	0,03	0,00	-0,08	-0,14	-0,16	-0,25
300	a	-0,11	-0,02	-0,03	-0,14	0,14	0,15	0,15	0,16
	b_1	-0,24	0,06	0,15	0,17	-0,02	-0,04	-0,06	-0,08
	b_2	0,56	0,62	0,90	0,84	-0,22	-0,22	-0,22	-0,29
	b_3	0,87	0,87	1,16	1,00	-0,27	-0,27	-0,27	-0,35
	b_4	1,41	1,35	1,61	1,13	-0,39	-0,38	-0,39	-0,50
	b_5	1,51	1,42	1,70	1,11	-0,52	-0,47	-0,49	-0,60
	b_6	1,70	1,40	1,68	0,83	-1,04	-0,85	-0,83	-0,95
	θ	0,41	0,45	0,68	0,70	-0,06	-0,09	-0,09	-0,15
500	a	0,01	0,00	0,04	-0,02	0,11	0,09	0,12	0,11
	b_1	-0,72	-0,31	-0,80	-0,63	-0,02	-0,02	-0,03	-0,03
	b_2	-0,03	0,25	-0,12	-0,15	-0,15	-0,15	-0,15	-0,18
	b_3	0,55	0,66	0,21	0,06	-0,19	-0,19	-0,20	-0,23
	b_4	1,23	1,22	0,63	0,32	-0,29	-0,27	-0,29	-0,33
	b_5	1,35	1,28	0,78	0,37	-0,39	-0,34	-0,35	-0,42
	b_6	1,43	1,18	1,01	0,41	-0,82	-0,65	-0,63	-0,72
	θ	0,20	0,36	0,11	0,08	-0,03	-0,04	-0,06	-0,08
750	a	0,01	-0,02	0,00	-0,02	0,08	0,05	0,08	0,09
	b_1	-1,52	-1,08	-0,95	-0,68	0,00	-0,04	-0,03	-0,03
	b_2	-0,53	-0,39	-0,31	-0,26	-0,12	-0,11	-0,12	-0,16
	b_3	-0,04	-0,05	-0,01	-0,04	-0,16	-0,14	-0,16	-0,20
	b_4	0,63	0,37	0,37	0,20	-0,23	-0,19	-0,22	-0,28
	b_5	0,88	0,56	0,55	0,30	-0,31	-0,23	-0,27	-0,36
	b_6	1,43	0,90	0,93	0,52	-0,69	-0,50	-0,50	-0,58
	θ	0,02	0,00	0,02	0,01	-0,03	-0,04	-0,04	-0,07
1000	a	-0,10	-0,05	0,03	0,15	0,05	0,05	0,07	0,08
	b_1	-1,07	-0,78	-0,51	-0,25	0,00	-0,02	-0,03	-0,01
	b_2	-0,33	-0,25	-0,12	-0,10	-0,06	-0,09	-0,10	-0,12
	b_3	0,10	0,05	0,15	0,10	-0,09	-0,12	-0,13	-0,16
	b_4	0,62	0,42	0,47	0,31	-0,15	-0,18	-0,19	-0,23
	b_5	0,95	0,65	0,68	0,43	-0,21	-0,22	-0,23	-0,29
	b_6	1,58	1,13	1,14	0,52	-0,54	-0,45	-0,40	-0,52
	θ	0,09	0,06	0,09	0,10	-0,01	-0,03	-0,04	-0,05

Tabela 14 – Correlação média das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) sob diferentes métodos de estimação, tamanhos amostrais e tamanhos do teste

n	Par	MVM				Gibbs			
		8	12	16	22	8	12	16	22
75	a	0,75	0,77	0,83	0,83	0,91	0,89	0,91	0,93
	b_1	0,85	0,84	0,84	0,83	0,94	0,93	0,93	0,92
	b_2	0,73	0,71	0,69	0,65	0,94	0,94	0,94	0,93
	b_3	0,48	0,47	0,46	0,43	0,91	0,90	0,91	0,90
	b_4	0,12	0,12	0,15	0,16	0,83	0,83	0,84	0,83
	b_5	-0,04	-0,01	0,04	0,07	0,81	0,78	0,80	0,78
	b_6	0,10	0,26	0,28	0,28	0,26	0,17	0,29	0,32
	θ	0,83	0,84	0,81	0,81	0,91	0,92	0,94	0,94
150	a	0,94	0,93	0,93	0,94	0,95	0,95	0,95	0,95
	b_1	0,97	0,96	0,95	0,94	0,97	0,97	0,97	0,96
	b_2	0,89	0,88	0,88	0,88	0,96	0,97	0,97	0,96
	b_3	0,64	0,65	0,67	0,67	0,93	0,94	0,94	0,94
	b_4	0,19	0,25	0,31	0,33	0,89	0,87	0,89	0,88
	b_5	0,00	0,09	0,17	0,20	0,85	0,82	0,84	0,83
	b_6	0,06	0,30	0,36	0,35	0,21	0,24	0,42	0,43
	θ	0,91	0,92	0,93	0,95	0,91	0,92	0,94	0,95
300	a	0,78	0,78	0,73	0,79	0,96	0,97	0,97	0,98
	b_1	0,67	0,69	0,70	0,72	0,98	0,98	0,98	0,98
	b_2	0,66	0,67	0,63	0,56	0,98	0,98	0,98	0,98
	b_3	0,43	0,43	0,41	0,38	0,97	0,96	0,97	0,96
	b_4	0,06	0,06	0,09	0,13	0,94	0,93	0,94	0,93
	b_5	-0,09	-0,05	0,00	0,06	0,92	0,91	0,92	0,90
	b_6	-0,04	0,16	0,20	0,25	0,36	0,42	0,55	0,53
	θ	0,87	0,89	0,89	0,89	0,90	0,93	0,94	0,95
500	a	0,86	0,88	0,95	0,97	0,98	0,98	0,98	0,99
	b_1	0,89	0,89	0,90	0,91	0,98	0,99	0,99	0,99
	b_2	0,82	0,78	0,79	0,79	0,99	0,99	0,99	0,99
	b_3	0,57	0,54	0,57	0,57	0,98	0,97	0,98	0,98
	b_4	0,16	0,14	0,21	0,24	0,94	0,94	0,95	0,95
	b_5	0,03	0,05	0,13	0,16	0,93	0,92	0,93	0,92
	b_6	0,10	0,31	0,34	0,34	0,34	0,46	0,58	0,59
	θ	0,87	0,89	0,91	0,92	0,91	0,92	0,94	0,95
750	a	0,98	0,99	0,98	0,98	0,99	0,99	0,99	0,99
	b_1	0,98	0,97	0,96	0,95	0,99	0,99	0,99	0,99
	b_2	0,89	0,90	0,89	0,89	0,99	0,99	0,99	0,99
	b_3	0,67	0,68	0,70	0,69	0,98	0,98	0,98	0,98
	b_4	0,22	0,25	0,30	0,32	0,96	0,96	0,96	0,96
	b_5	0,07	0,13	0,19	0,21	0,95	0,94	0,95	0,94
	b_6	0,10	0,29	0,34	0,34	0,41	0,52	0,64	0,65
	θ	0,90	0,92	0,94	0,95	0,91	0,92	0,94	0,95
1000	a	0,90	0,93	0,89	0,90	0,99	0,99	0,99	0,99
	b_1	0,92	0,92	0,91	0,90	0,99	0,99	0,99	0,99
	b_2	0,86	0,86	0,86	0,86	0,99	0,99	0,99	0,99
	b_3	0,61	0,63	0,64	0,63	0,99	0,99	0,99	0,99
	b_4	0,24	0,28	0,32	0,33	0,97	0,96	0,97	0,97
	b_5	0,11	0,17	0,22	0,23	0,95	0,95	0,96	0,95
	b_6	0,15	0,33	0,34	0,33	0,40	0,57	0,67	0,69
	θ	0,90	0,92	0,93	0,93	0,91	0,92	0,94	0,95

Tabela 15 – REQM médio das estimativas dos parâmetros dos itens (a , b_1 , b_2 , b_3 , b_4 , b_5 , b_6) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação fraca entre os traços latentes (ρ_{fr}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	0,66	1,03	0,64	0,75	0,86	0,46	0,46	0,63	0,95	0,42	0,52	0,39
	b_1	0,52	0,73	1,30	0,65	0,83	0,28	0,36	1,07	0,45	0,42	0,44	0,38
	b_2	0,30	0,43	0,59	0,48	0,42	0,20	0,35	0,53	0,53	0,35	0,53	0,34
	b_3	0,24	0,45	0,30	0,40	0,35	0,24	0,36	0,39	0,63	0,34	0,60	0,35
	b_4	0,20	0,81	0,35	0,60	0,42	0,29	0,55	0,36	0,74	0,47	0,72	0,41
	b_5	0,23	1,12	0,57	0,90	0,50	0,33	0,75	0,44	0,81	0,63	0,74	0,41
	b_6	0,48	1,65	1,05	1,18	1,04	0,68	0,89	1,02	0,97	0,78	0,82	0,48
	θ	0,57	0,91	0,98	0,65	1,13	0,76	0,42	0,73	0,81	0,38	0,58	0,50
150	a	0,50	1,15	0,77	0,62	0,99	0,45	0,30	0,69	1,00	0,27	0,34	0,28
	b_1	0,31	0,76	1,61	0,52	1,04	0,20	0,22	1,15	0,28	0,26	0,24	0,25
	b_2	0,18	0,26	0,79	0,33	0,50	0,16	0,25	0,57	0,37	0,22	0,33	0,20
	b_3	0,15	0,35	0,53	0,31	0,36	0,18	0,28	0,37	0,43	0,23	0,40	0,22
	b_4	0,18	0,77	0,57	0,50	0,56	0,22	0,42	0,46	0,54	0,31	0,53	0,27
	b_5	0,23	1,00	0,89	0,75	0,83	0,26	0,60	0,65	0,62	0,45	0,59	0,31
	b_6	0,36	1,88	1,66	1,03	1,59	0,45	0,72	1,24	0,79	0,58	0,70	0,46
	θ	0,50	0,95	1,00	0,59	1,13	0,77	0,38	0,65	0,82	0,34	0,50	0,43
300	a	0,29	0,79	0,66	0,43	0,56	0,23	0,19	0,46	0,76	0,19	0,24	0,18
	b_1	0,18	0,31	1,31	0,23	0,33	0,11	0,16	0,54	0,17	0,18	0,17	0,15
	b_2	0,12	0,16	0,65	0,21	0,19	0,11	0,17	0,28	0,26	0,17	0,23	0,14
	b_3	0,10	0,22	0,47	0,24	0,16	0,14	0,18	0,25	0,32	0,19	0,29	0,16
	b_4	0,10	0,34	0,61	0,33	0,22	0,17	0,28	0,34	0,43	0,28	0,36	0,20
	b_5	0,09	0,49	0,83	0,45	0,31	0,21	0,40	0,45	0,50	0,35	0,44	0,25
	b_6	0,18	0,85	1,60	0,53	0,57	0,31	0,49	0,82	0,66	0,46	0,53	0,36
	θ	0,46	0,73	0,87	0,52	0,89	0,67	0,36	0,54	0,76	0,33	0,44	0,42
500	a	0,25	0,85	0,73	0,41	0,56	0,21	0,17	0,42	0,67	0,14	0,16	0,14
	b_1	0,14	0,29	1,43	0,18	0,29	0,09	0,11	0,52	0,16	0,13	0,13	0,12
	b_2	0,09	0,13	0,70	0,14	0,15	0,10	0,13	0,26	0,27	0,13	0,17	0,10
	b_3	0,07	0,19	0,58	0,15	0,17	0,11	0,15	0,24	0,38	0,14	0,20	0,11
	b_4	0,07	0,36	0,80	0,24	0,23	0,14	0,20	0,33	0,52	0,20	0,25	0,15
	b_5	0,07	0,49	1,05	0,32	0,31	0,15	0,26	0,43	0,63	0,26	0,29	0,19
	b_6	0,13	0,91	1,97	0,46	0,58	0,22	0,38	0,75	0,81	0,33	0,37	0,28
	θ	0,44	0,72	0,89	0,50	0,89	0,67	0,35	0,53	0,75	0,33	0,44	0,42
750	a	0,18	1,35	0,93	0,30	1,19	0,15	0,15	0,64	1,28	0,12	0,14	0,13
	b_1	0,09	1,03	2,70	0,30	1,60	0,08	0,08	1,57	0,23	0,10	0,09	0,09
	b_2	0,05	0,36	1,36	0,21	0,72	0,07	0,09	0,81	0,47	0,09	0,14	0,08
	b_3	0,04	0,69	1,13	0,20	0,76	0,08	0,11	0,73	0,70	0,12	0,17	0,09
	b_4	0,05	1,41	1,54	0,27	1,30	0,11	0,15	1,01	0,97	0,15	0,22	0,12
	b_5	0,07	1,86	1,98	0,44	1,71	0,13	0,25	1,30	1,14	0,23	0,27	0,16
	b_6	0,13	3,30	3,59	0,62	3,09	0,21	0,30	2,23	1,46	0,31	0,36	0,25
	θ	0,45	0,96	0,98	0,49	1,12	0,78	0,36	0,57	0,83	0,33	0,45	0,42

Tabela 16 – Viés médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação fraca entre os traços latentes (ρ_{fr}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	-0,21	-0,74	-0,61	-0,40	-0,62	0,08	0,08	-0,44	0,41	-0,04	0,24	0,13
	b_1	-0,40	-0,65	-1,12	-0,51	-0,71	-0,19	-0,25	-0,77	-0,35	-0,32	-0,39	-0,32
	b_2	-0,24	-0,32	-0,46	-0,38	-0,27	-0,13	-0,28	-0,38	-0,47	-0,27	-0,47	-0,30
	b_3	-0,19	-0,12	-0,12	-0,27	-0,08	-0,08	-0,26	-0,20	-0,52	-0,23	-0,50	-0,29
	b_4	-0,12	-0,03	0,25	-0,25	0,14	-0,03	-0,30	0,05	-0,57	-0,23	-0,58	-0,30
	b_5	-0,05	0,04	0,52	-0,29	0,29	0,06	-0,35	0,20	-0,57	-0,23	-0,54	-0,21
	b_6	0,18	0,55	0,93	0,04	0,84	0,37	-0,27	0,71	-0,48	-0,05	-0,41	-0,07
θ	-0,13	-0,04	-0,01	-0,11	-0,15	-0,05	-0,17	-0,28	-0,12	-0,16	-0,32	-0,25	
150	a	0,09	-0,92	-0,75	-0,24	-0,79	0,10	0,04	-0,48	0,32	-0,02	0,13	0,09
	b_1	-0,06	-0,62	-1,40	-0,38	-0,87	-0,13	-0,14	-0,73	-0,20	-0,15	-0,18	-0,14
	b_2	-0,04	-0,09	-0,49	-0,25	-0,29	-0,08	-0,18	-0,29	-0,26	-0,13	-0,26	-0,15
	b_3	-0,06	0,14	-0,08	-0,18	0,01	-0,04	-0,18	-0,09	-0,23	-0,12	-0,29	-0,14
	b_4	-0,07	0,39	0,48	-0,17	0,43	-0,04	-0,25	0,21	-0,25	-0,13	-0,36	-0,13
	b_5	-0,08	0,57	0,82	-0,22	0,70	0,04	-0,32	0,41	-0,20	-0,14	-0,37	-0,11
	b_6	-0,08	1,20	1,62	0,00	1,44	0,21	-0,33	0,90	-0,09	-0,07	-0,25	-0,05
θ	-0,05	-0,01	-0,01	-0,04	-0,08	-0,03	-0,13	-0,15	-0,07	-0,06	-0,16	-0,10	
300	a	-0,18	-0,62	-0,65	-0,28	-0,47	0,02	0,00	-0,33	0,14	-0,01	0,10	0,03
	b_1	-0,10	-0,25	-1,10	-0,17	-0,26	-0,05	-0,10	-0,36	-0,12	-0,13	-0,12	-0,09
	b_2	-0,08	-0,06	-0,34	-0,16	-0,09	-0,04	-0,13	-0,14	-0,10	-0,10	-0,18	-0,10
	b_3	-0,06	0,01	0,05	-0,16	0,00	-0,03	-0,12	-0,01	-0,06	-0,11	-0,21	-0,10
	b_4	-0,05	0,10	0,48	-0,17	0,13	-0,03	-0,16	0,14	-0,06	-0,13	-0,25	-0,11
	b_5	-0,03	0,16	0,73	-0,20	0,22	-0,01	-0,21	0,24	-0,01	-0,14	-0,26	-0,09
	b_6	0,00	0,40	1,53	-0,13	0,46	0,04	-0,19	0,57	0,12	-0,11	-0,25	-0,04
θ	-0,04	-0,03	-0,03	-0,06	-0,04	-0,03	-0,08	-0,09	-0,06	-0,08	-0,10	-0,08	
500	a	-0,15	-0,69	-0,72	-0,29	-0,49	0,02	0,00	-0,33	0,04	0,02	0,03	0,00
	b_1	-0,08	-0,24	-1,23	-0,12	-0,24	-0,05	-0,06	-0,33	-0,05	-0,09	-0,09	-0,06
	b_2	-0,05	-0,01	-0,32	-0,09	-0,03	-0,06	-0,07	-0,09	0,04	-0,09	-0,12	-0,04
	b_3	-0,04	0,12	0,14	-0,06	0,07	-0,05	-0,08	0,01	0,11	-0,09	-0,14	-0,04
	b_4	-0,02	0,25	0,66	-0,09	0,18	-0,05	-0,09	0,16	0,19	-0,11	-0,16	-0,02
	b_5	-0,01	0,32	0,94	-0,11	0,27	-0,04	-0,11	0,24	0,25	-0,12	-0,19	-0,01
	b_6	0,04	0,68	1,92	-0,03	0,53	0,02	-0,12	0,51	0,37	-0,10	-0,14	0,01
θ	-0,03	-0,02	-0,02	-0,01	-0,01	-0,03	-0,03	-0,06	-0,03	-0,06	-0,08	-0,04	
750	a	-0,01	-1,17	-0,92	-0,17	-1,11	0,00	0,06	-0,50	0,20	0,02	0,04	0,02
	b_1	0,01	-0,81	-2,23	-0,21	-1,36	-0,03	-0,01	-0,87	-0,04	-0,05	-0,05	-0,03
	b_2	-0,02	0,24	-0,55	-0,14	-0,13	-0,02	-0,03	-0,21	0,25	-0,05	-0,09	-0,04
	b_3	-0,02	0,67	0,26	-0,09	0,41	-0,02	-0,05	0,12	0,43	-0,05	-0,10	-0,04
	b_4	-0,04	1,37	1,28	-0,05	1,18	0,01	-0,08	0,55	0,61	-0,07	-0,13	-0,03
	b_5	-0,06	1,82	1,80	-0,05	1,61	0,01	-0,11	0,78	0,76	-0,09	-0,13	-0,02
	b_6	-0,07	3,10	3,55	0,09	3,03	0,05	-0,14	1,46	0,97	-0,07	-0,12	0,02
θ	-0,02	-0,01	0,00	-0,02	-0,01	-0,01	-0,01	-0,03	0,00	-0,03	-0,04	-0,03	

Tabela 17 – Correlação média das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação fraca entre os traços latentes (ρ_{fr}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	0,04	0,02	-0,53	0,38	0,15	0,71	0,70	0,71	0,45	0,76	0,74	0,66
	b_1	0,27	0,85	0,94	0,86	0,86	0,85	0,91	0,96	0,86	0,93	0,91	0,88
	b_2	0,65	0,89	0,96	0,94	0,94	0,80	0,95	0,95	0,78	0,96	0,88	0,94
	b_3	0,80	0,88	0,97	0,97	0,91	0,74	0,95	0,94	0,66	0,95	0,87	0,90
	b_4	0,78	0,82	0,95	0,94	0,82	0,68	0,94	0,88	0,54	0,94	0,84	0,81
	b_5	0,77	0,85	0,92	0,93	0,80	0,61	0,95	0,84	0,56	0,93	0,83	0,71
	b_6	0,45	0,44	0,80	0,74	0,58	0,39	0,86	0,44	0,52	0,83	0,71	0,78
	θ	0,82	0,42	0,30	0,74	0,29	0,85	0,92	0,73	0,59	0,94	0,88	0,89
150	a	0,36	0,00	-0,68	0,51	0,22	0,75	0,86	0,71	0,49	0,89	0,84	0,79
	b_1	0,11	0,92	0,95	0,82	0,85	0,92	0,96	0,96	0,93	0,95	0,96	0,91
	b_2	0,65	0,93	0,98	0,95	0,93	0,82	0,97	0,98	0,78	0,97	0,92	0,96
	b_3	0,84	0,91	0,98	0,98	0,95	0,79	0,98	0,96	0,70	0,97	0,91	0,93
	b_4	0,76	0,89	0,97	0,96	0,87	0,72	0,97	0,87	0,60	0,97	0,87	0,87
	b_5	0,72	0,93	0,96	0,94	0,80	0,68	0,98	0,80	0,60	0,97	0,87	0,78
	b_6	0,44	0,50	0,82	0,76	0,47	0,54	0,94	0,47	0,63	0,92	0,75	0,81
	θ	0,87	0,36	0,19	0,77	0,27	0,84	0,93	0,76	0,60	0,94	0,88	0,91
300	a	0,57	0,82	-0,33	0,84	0,66	0,90	0,95	0,85	0,51	0,94	0,92	0,88
	b_1	0,28	0,96	0,98	0,96	0,95	0,96	0,98	0,98	0,98	0,98	0,98	0,97
	b_2	0,86	0,97	0,99	0,99	0,98	0,92	0,99	0,98	0,85	0,99	0,96	0,98
	b_3	0,94	0,97	0,99	0,99	0,98	0,89	0,99	0,97	0,81	0,99	0,96	0,96
	b_4	0,92	0,97	0,98	0,99	0,96	0,83	0,99	0,92	0,74	0,98	0,94	0,93
	b_5	0,94	0,98	0,98	0,99	0,95	0,76	0,99	0,89	0,77	0,98	0,92	0,87
	b_6	0,83	0,92	0,92	0,97	0,85	0,62	0,96	0,76	0,79	0,95	0,87	0,90
	θ	0,89	0,70	0,50	0,86	0,53	0,83	0,93	0,84	0,66	0,94	0,90	0,91
500	a	0,74	0,86	-0,45	0,89	0,81	0,91	0,95	0,89	0,51	0,97	0,96	0,90
	b_1	0,24	0,98	0,98	0,98	0,97	0,97	0,99	0,98	0,98	0,99	0,98	0,97
	b_2	0,92	0,98	0,99	0,99	0,99	0,94	0,99	0,98	0,84	0,99	0,98	0,98
	b_3	0,97	0,98	1,00	0,99	0,98	0,94	0,99	0,98	0,79	0,99	0,98	0,98
	b_4	0,96	0,98	0,99	0,99	0,97	0,89	0,99	0,93	0,73	0,99	0,97	0,95
	b_5	0,98	0,98	0,98	0,99	0,97	0,86	0,99	0,90	0,75	0,99	0,97	0,91
	b_6	0,90	0,95	0,93	0,98	0,93	0,69	0,97	0,79	0,77	0,97	0,94	0,92
	θ	0,89	0,68	0,47	0,86	0,54	0,84	0,94	0,85	0,67	0,94	0,90	0,91
750	a	0,72	-0,38	-0,81	0,83	0,17	0,95	0,97	0,84	0,47	0,98	0,97	0,92
	b_1	0,51	0,99	0,99	0,87	0,94	0,97	0,99	0,96	0,99	0,99	0,99	0,98
	b_2	0,96	0,97	0,99	0,97	0,98	0,96	0,99	0,97	0,69	1,00	0,98	0,99
	b_3	0,98	0,97	1,00	0,98	0,96	0,96	0,99	0,93	0,65	0,99	0,97	0,98
	b_4	0,98	0,97	0,99	0,98	0,91	0,94	0,99	0,74	0,57	0,99	0,97	0,97
	b_5	0,98	0,98	0,99	0,97	0,88	0,89	0,99	0,65	0,63	0,99	0,97	0,93
	b_6	0,91	0,79	0,94	0,91	0,66	0,78	0,98	0,38	0,65	0,98	0,94	0,95
	θ	0,89	0,29	0,18	0,86	0,27	0,86	0,93	0,82	0,58	0,94	0,89	0,91

Tabela 18 – REQM médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação média entre os traços latentes (ρ_{me}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	0,53	0,72	0,58	0,59	0,54	0,61	0,36	0,47	0,79	0,40	0,48	0,35
	b_1	0,36	0,41	1,10	0,57	0,59	0,30	0,44	0,70	0,47	0,51	0,49	0,39
	b_2	0,21	0,39	0,59	0,41	0,34	0,26	0,38	0,48	0,60	0,46	0,59	0,42
	b_3	0,21	0,46	0,37	0,41	0,26	0,27	0,40	0,39	0,66	0,48	0,65	0,43
	b_4	0,20	0,75	0,27	0,59	0,29	0,31	0,55	0,33	0,83	0,56	0,78	0,46
	b_5	0,22	1,07	0,39	0,85	0,34	0,33	0,73	0,35	0,86	0,71	0,80	0,46
	b_6	0,48	1,28	0,99	1,06	0,69	0,54	0,89	0,65	0,99	0,86	0,87	0,52
	θ	0,49	0,70	0,88	0,54	0,90	0,67	0,44	0,65	0,76	0,45	0,58	0,51
150	a	0,35	0,74	0,63	0,44	0,56	0,39	0,34	0,40	0,88	0,27	0,40	0,27
	b_1	0,22	0,35	1,17	0,34	0,48	0,16	0,25	0,61	0,34	0,28	0,27	0,22
	b_2	0,13	0,25	0,53	0,27	0,28	0,14	0,27	0,36	0,40	0,26	0,38	0,23
	b_3	0,10	0,30	0,38	0,28	0,24	0,16	0,28	0,27	0,43	0,29	0,45	0,26
	b_4	0,13	0,55	0,44	0,47	0,23	0,22	0,40	0,26	0,57	0,36	0,58	0,31
	b_5	0,18	0,66	0,61	0,59	0,32	0,23	0,57	0,33	0,64	0,50	0,66	0,37
	b_6	0,26	1,13	1,21	0,82	0,63	0,44	0,73	0,60	0,78	0,60	0,74	0,45
	θ	0,47	0,73	0,90	0,51	0,89	0,66	0,39	0,57	0,76	0,35	0,48	0,43
300	a	0,30	1,26	0,84	0,29	1,13	0,29	0,20	0,69	1,13	0,20	0,26	0,18
	b_1	0,14	0,83	2,01	0,24	1,24	0,09	0,12	1,28	0,20	0,16	0,14	0,13
	b_2	0,09	0,22	1,03	0,22	0,55	0,10	0,17	0,66	0,29	0,15	0,22	0,13
	b_3	0,10	0,40	0,76	0,25	0,50	0,12	0,20	0,51	0,39	0,17	0,28	0,13
	b_4	0,12	0,87	1,00	0,35	0,89	0,16	0,29	0,68	0,56	0,23	0,38	0,18
	b_5	0,13	1,23	1,31	0,55	1,22	0,19	0,44	0,93	0,65	0,34	0,45	0,21
	b_6	0,19	2,16	2,43	0,65	2,21	0,32	0,52	1,68	0,84	0,40	0,57	0,33
	θ	0,47	0,96	1,00	0,50	1,13	0,79	0,37	0,62	0,82	0,35	0,47	0,42
500	a	0,23	1,32	0,90	0,24	1,22	0,21	0,18	0,67	1,23	0,15	0,18	0,15
	b_1	0,10	0,98	2,45	0,18	1,49	0,08	0,10	1,52	0,20	0,12	0,13	0,11
	b_2	0,06	0,27	1,23	0,17	0,66	0,07	0,14	0,74	0,36	0,11	0,17	0,11
	b_3	0,06	0,55	0,91	0,19	0,67	0,09	0,16	0,64	0,52	0,12	0,21	0,12
	b_4	0,08	1,19	1,26	0,27	1,12	0,14	0,22	0,87	0,72	0,19	0,26	0,16
	b_5	0,08	1,59	1,65	0,44	1,50	0,17	0,29	1,12	0,87	0,27	0,31	0,19
	b_6	0,15	2,97	3,10	0,54	2,78	0,27	0,34	2,00	1,14	0,35	0,40	0,30
	θ	0,46	0,95	0,98	0,46	1,12	0,79	0,36	0,59	0,83	0,33	0,46	0,41
750	a	0,26	0,82	0,73	0,43	0,57	0,16	0,14	0,41	0,74	0,12	0,13	0,13
	b_1	0,13	0,27	1,42	0,20	0,28	0,06	0,10	0,52	0,17	0,09	0,09	0,09
	b_2	0,07	0,12	0,71	0,13	0,13	0,07	0,09	0,27	0,34	0,10	0,14	0,08
	b_3	0,06	0,17	0,61	0,12	0,14	0,09	0,12	0,29	0,49	0,12	0,17	0,10
	b_4	0,05	0,35	0,87	0,16	0,22	0,13	0,15	0,43	0,67	0,17	0,22	0,13
	b_5	0,06	0,46	1,13	0,24	0,29	0,15	0,22	0,56	0,80	0,21	0,27	0,16
	b_6	0,13	0,84	1,98	0,33	0,56	0,22	0,31	0,93	1,04	0,31	0,36	0,25
	θ	1,14	0,79	0,72	1,11	0,96	0,85	1,22	0,60	1,01	1,23	0,50	1,12

Tabela 19 – Viés médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação média entre os traços latentes (ρ_{me}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	-0,06	-0,45	-0,54	-0,30	-0,41	0,18	0,03	-0,18	0,35	0,01	0,23	0,12
	b_1	-0,21	-0,34	-0,96	-0,47	-0,49	-0,23	-0,36	-0,56	-0,39	-0,43	-0,44	-0,32
	b_2	-0,15	-0,29	-0,46	-0,35	-0,26	-0,18	-0,31	-0,40	-0,55	-0,40	-0,53	-0,37
	b_3	-0,13	-0,22	-0,18	-0,32	-0,15	-0,15	-0,31	-0,29	-0,58	-0,40	-0,57	-0,38
	b_4	-0,08	-0,23	0,07	-0,35	-0,05	-0,14	-0,36	-0,16	-0,69	-0,39	-0,65	-0,36
	b_5	-0,07	-0,27	0,28	-0,38	0,05	-0,06	-0,41	-0,05	-0,66	-0,40	-0,62	-0,31
	b_6	0,07	0,01	0,85	-0,21	0,40	0,16	-0,33	0,26	-0,61	-0,29	-0,51	-0,16
	θ	-0,12	-0,09	-0,06	-0,16	-0,18	-0,08	-0,25	-0,31	-0,20	-0,29	-0,36	-0,30
150	a	-0,10	-0,52	-0,61	-0,21	-0,46	0,10	0,06	-0,24	0,31	0,04	0,17	0,10
	b_1	-0,10	-0,25	-0,96	-0,24	-0,37	-0,09	-0,17	-0,44	-0,28	-0,19	-0,21	-0,14
	b_2	-0,06	-0,12	-0,31	-0,21	-0,21	-0,09	-0,20	-0,26	-0,32	-0,20	-0,32	-0,18
	b_3	-0,05	-0,06	-0,04	-0,20	-0,10	-0,06	-0,20	-0,17	-0,32	-0,20	-0,37	-0,19
	b_4	-0,04	-0,06	0,33	-0,27	0,06	-0,05	-0,25	-0,05	-0,39	-0,20	-0,43	-0,21
	b_5	-0,07	-0,02	0,55	-0,30	0,19	0,01	-0,32	0,07	-0,40	-0,24	-0,45	-0,19
	b_6	-0,02	0,22	1,17	-0,23	0,48	0,16	-0,27	0,33	-0,28	-0,23	-0,33	-0,15
	θ	-0,07	-0,03	-0,01	-0,09	-0,10	-0,02	-0,14	-0,17	-0,14	-0,13	-0,19	-0,13
300	a	-0,05	-1,05	-0,83	-0,07	-1,05	0,06	0,07	-0,55	0,27	0,03	0,10	0,03
	b_1	-0,02	-0,68	-1,72	-0,15	-1,06	-0,05	-0,06	-0,76	-0,11	-0,07	-0,09	-0,06
	b_2	-0,05	0,03	-0,51	-0,14	-0,18	-0,06	-0,10	-0,24	-0,05	-0,09	-0,16	-0,08
	b_3	-0,06	0,31	0,13	-0,13	0,23	-0,06	-0,12	0,04	0,01	-0,08	-0,19	-0,07
	b_4	-0,07	0,72	0,86	-0,16	0,81	-0,06	-0,17	0,41	0,03	-0,11	-0,23	-0,06
	b_5	-0,08	0,99	1,22	-0,23	1,14	-0,04	-0,24	0,62	0,05	-0,14	-0,24	-0,05
	b_6	-0,11	1,75	2,40	-0,20	2,15	0,01	-0,26	1,25	0,22	-0,13	-0,25	-0,03
	θ	-0,06	-0,02	-0,01	-0,04	-0,03	-0,02	-0,06	-0,08	-0,04	-0,04	-0,08	-0,06
500	a	0,01	-1,13	-0,88	-0,03	-1,13	0,01	0,08	-0,53	0,23	0,03	0,05	0,03
	b_1	0,03	-0,79	-2,07	-0,10	-1,26	-0,02	-0,03	-0,87	-0,05	-0,04	-0,08	-0,06
	b_2	-0,02	0,11	-0,56	-0,12	-0,14	0,00	-0,07	-0,21	0,13	-0,04	-0,12	-0,06
	b_3	-0,04	0,52	0,16	-0,12	0,35	0,00	-0,08	0,08	0,25	-0,04	-0,15	-0,06
	b_4	-0,05	1,13	1,09	-0,16	1,03	0,02	-0,12	0,48	0,34	-0,05	-0,17	-0,05
	b_5	-0,06	1,49	1,54	-0,23	1,43	0,05	-0,16	0,69	0,45	-0,07	-0,18	-0,05
	b_6	-0,09	2,67	3,06	-0,19	2,71	0,10	-0,19	1,35	0,61	-0,05	-0,17	-0,01
	θ	-0,04	-0,01	-0,01	-0,05	-0,01	0,01	-0,03	-0,05	0,00	-0,01	-0,07	-0,05
750	a	-0,21	-0,66	-0,72	-0,33	-0,49	0,03	-0,01	-0,31	0,02	0,02	0,04	0,03
	b_1	-0,07	-0,21	-1,19	-0,15	-0,24	-0,04	-0,04	-0,32	-0,03	-0,04	-0,06	-0,04
	b_2	-0,03	0,03	-0,29	-0,09	-0,04	-0,03	-0,05	-0,06	0,16	-0,05	-0,09	-0,05
	b_3	-0,02	0,11	0,14	-0,07	0,05	-0,03	-0,06	0,05	0,25	-0,05	-0,11	-0,06
	b_4	-0,01	0,27	0,72	-0,05	0,17	-0,03	-0,07	0,21	0,37	-0,07	-0,13	-0,06
	b_5	0,01	0,35	1,01	-0,07	0,23	-0,04	-0,09	0,31	0,47	-0,08	-0,13	-0,06
	b_6	0,09	0,68	1,93	-0,01	0,49	0,00	-0,09	0,59	0,64	-0,11	-0,10	-0,05
	θ	-0,02	-0,01	0,00	-0,02	-0,01	-0,02	-0,02	-0,04	-0,01	-0,04	-0,05	-0,03

Tabela 20 – Correlação média das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação média entre os traços latentes (ρ_{me}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	0,15	0,50	-0,53	0,57	0,58	0,65	0,77	0,75	0,47	0,76	0,74	0,68
	b_1	0,10	0,87	0,91	0,90	0,93	0,86	0,91	0,95	0,88	0,93	0,91	0,88
	b_2	0,79	0,93	0,95	0,97	0,96	0,75	0,96	0,94	0,82	0,95	0,89	0,92
	b_3	0,84	0,90	0,96	0,97	0,95	0,74	0,96	0,94	0,77	0,95	0,89	0,88
	b_4	0,76	0,88	0,95	0,97	0,89	0,70	0,95	0,91	0,59	0,94	0,86	0,81
	b_5	0,79	0,86	0,95	0,96	0,88	0,66	0,95	0,89	0,59	0,93	0,84	0,69
	b_6	0,49	0,70	0,80	0,85	0,69	0,32	0,85	0,65	0,56	0,81	0,71	0,78
θ	0,88	0,74	0,48	0,84	0,54	0,83	0,93	0,82	0,68	0,94	0,90	0,91	
150	a	0,45	0,52	-0,51	0,74	0,58	0,79	0,82	0,85	0,54	0,89	0,80	0,75
	b_1	0,42	0,92	0,94	0,94	0,90	0,92	0,95	0,97	0,95	0,96	0,96	0,94
	b_2	0,83	0,94	0,97	0,98	0,96	0,89	0,98	0,97	0,83	0,98	0,93	0,96
	b_3	0,95	0,94	0,97	0,98	0,95	0,87	0,98	0,97	0,82	0,98	0,93	0,93
	b_4	0,89	0,95	0,96	0,98	0,94	0,76	0,97	0,93	0,71	0,96	0,89	0,86
	b_5	0,87	0,97	0,96	0,98	0,93	0,75	0,97	0,91	0,71	0,97	0,87	0,74
	b_6	0,73	0,78	0,90	0,93	0,78	0,51	0,90	0,75	0,68	0,92	0,75	0,83
θ	0,88	0,70	0,49	0,85	0,55	0,82	0,93	0,84	0,67	0,94	0,89	0,91	
300	a	0,49	-0,35	-0,72	0,85	0,18	0,85	0,94	0,76	0,49	0,94	0,90	0,87
	b_1	0,31	0,97	0,98	0,92	0,94	0,97	0,98	0,97	0,97	0,97	0,98	0,97
	b_2	0,91	0,95	0,99	0,98	0,97	0,94	0,99	0,98	0,79	0,99	0,96	0,98
	b_3	0,94	0,94	0,99	0,98	0,96	0,92	0,99	0,96	0,72	0,98	0,95	0,97
	b_4	0,91	0,94	0,98	0,98	0,91	0,87	0,99	0,83	0,62	0,98	0,93	0,94
	b_5	0,93	0,95	0,98	0,97	0,88	0,82	0,99	0,75	0,65	0,98	0,91	0,90
	b_6	0,81	0,74	0,96	0,93	0,67	0,59	0,97	0,46	0,70	0,97	0,85	0,90
θ	0,88	0,33	0,20	0,86	0,27	0,86	0,93	0,78	0,60	0,94	0,89	0,91	
500	a	0,65	-0,30	-0,80	0,91	0,22	0,92	0,96	0,79	0,47	0,97	0,95	0,91
	b_1	0,55	0,98	0,98	0,95	0,97	0,98	0,99	0,96	0,98	0,98	0,98	0,98
	b_2	0,95	0,96	0,99	0,99	0,98	0,95	0,99	0,97	0,74	0,99	0,97	0,98
	b_3	0,97	0,96	0,99	0,99	0,97	0,94	0,99	0,94	0,69	0,99	0,97	0,97
	b_4	0,97	0,96	0,99	0,99	0,93	0,88	0,99	0,78	0,61	0,99	0,96	0,95
	b_5	0,98	0,97	0,99	0,99	0,93	0,85	1,00	0,70	0,65	0,98	0,96	0,92
	b_6	0,90	0,71	0,96	0,96	0,75	0,68	0,99	0,45	0,66	0,97	0,93	0,92
θ	0,89	0,29	0,18	0,89	0,28	0,85	0,94	0,80	0,59	0,94	0,89	0,91	
750	a	0,76	0,89	-0,37	0,92	0,87	0,95	0,97	0,91	0,49	0,98	0,97	0,93
	b_1	0,47	0,98	0,97	0,98	0,98	0,98	0,99	0,98	0,99	0,99	0,99	0,98
	b_2	0,95	0,98	0,99	0,99	0,99	0,95	1,00	0,99	0,80	0,99	0,98	0,99
	b_3	0,97	0,99	0,99	1,00	0,99	0,94	1,00	0,97	0,77	0,99	0,98	0,99
	b_4	0,98	0,99	0,99	0,99	0,98	0,91	0,99	0,91	0,71	0,99	0,97	0,97
	b_5	0,98	0,99	0,99	0,99	0,98	0,89	0,99	0,87	0,73	0,99	0,96	0,94
	b_6	0,92	0,97	0,95	0,98	0,94	0,75	0,99	0,74	0,77	0,98	0,94	0,94
θ	0,28	0,60	0,83	0,31	0,45	0,52	0,21	0,80	0,35	0,20	0,87	0,31	

Tabela 21 – REQM médio das estimativas dos parâmetros dos itens (a , b_1 , b_2 , b_3 , b_4 , b_5 , b_6) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação forte entre os traços latentes (ρ_{f_0}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	0,53	0,64	0,29	0,56	0,59	0,39	0,41	0,49	0,42	0,42	0,44	0,33
	b_1	0,46	0,45	0,50	0,55	0,70	0,33	0,47	0,70	0,47	0,62	0,51	0,50
	b_2	0,27	0,41	0,33	0,45	0,42	0,27	0,49	0,45	0,60	0,54	0,62	0,50
	b_3	0,19	0,52	0,30	0,44	0,31	0,25	0,49	0,39	0,66	0,54	0,66	0,49
	b_4	0,19	0,79	0,31	0,60	0,28	0,30	0,62	0,35	0,81	0,62	0,77	0,54
	b_5	0,19	1,11	0,33	0,87	0,37	0,34	0,84	0,35	0,91	0,77	0,81	0,55
	b_6	0,34	1,28	0,53	1,09	0,70	0,64	0,93	0,57	0,98	0,89	0,87	0,56
	θ	0,43	0,64	0,47	0,48	0,70	0,44	0,45	0,60	0,52	0,47	0,58	0,51
150	a	0,39	0,63	0,23	0,52	0,56	0,49	0,32	0,47	0,25	0,30	0,36	0,26
	b_1	0,27	0,33	0,33	0,38	0,59	0,20	0,35	0,52	0,29	0,37	0,29	0,30
	b_2	0,17	0,26	0,24	0,31	0,28	0,18	0,33	0,33	0,37	0,32	0,40	0,26
	b_3	0,15	0,30	0,24	0,31	0,22	0,17	0,35	0,26	0,41	0,32	0,47	0,27
	b_4	0,15	0,49	0,25	0,42	0,22	0,23	0,44	0,28	0,53	0,43	0,59	0,31
	b_5	0,16	0,68	0,26	0,63	0,31	0,30	0,65	0,35	0,61	0,57	0,64	0,36
	b_6	0,27	1,04	0,47	0,80	0,64	0,55	0,74	0,58	0,78	0,68	0,74	0,48
	θ	0,43	0,65	0,48	0,44	0,69	0,40	0,39	0,54	0,46	0,35	0,46	0,39
300	a	0,30	0,61	0,17	0,41	0,57	0,23	0,22	0,42	0,14	0,19	0,21	0,18
	b_1	0,20	0,22	0,24	0,29	0,41	0,14	0,19	0,46	0,15	0,18	0,17	0,17
	b_2	0,11	0,16	0,16	0,18	0,22	0,11	0,17	0,24	0,21	0,19	0,23	0,16
	b_3	0,09	0,20	0,12	0,18	0,17	0,13	0,19	0,22	0,24	0,20	0,28	0,17
	b_4	0,09	0,28	0,15	0,24	0,26	0,19	0,27	0,24	0,35	0,29	0,37	0,21
	b_5	0,11	0,44	0,19	0,39	0,34	0,24	0,38	0,34	0,39	0,39	0,42	0,23
	b_6	0,18	0,68	0,25	0,50	0,64	0,34	0,45	0,67	0,48	0,45	0,53	0,34
	θ	0,42	0,63	0,47	0,45	0,69	0,40	0,35	0,51	0,42	0,32	0,42	0,35
500	a	0,25	0,67	0,15	0,33	0,62	0,18	0,17	0,40	0,13	0,15	0,18	0,13
	b_1	0,15	0,22	0,20	0,22	0,36	0,10	0,14	0,43	0,11	0,13	0,13	0,13
	b_2	0,10	0,13	0,12	0,15	0,18	0,08	0,13	0,23	0,15	0,14	0,18	0,12
	b_3	0,07	0,16	0,11	0,13	0,20	0,10	0,13	0,21	0,20	0,15	0,23	0,12
	b_4	0,07	0,24	0,12	0,18	0,31	0,15	0,16	0,31	0,26	0,19	0,29	0,15
	b_5	0,07	0,35	0,16	0,26	0,43	0,19	0,24	0,40	0,32	0,24	0,33	0,18
	b_6	0,15	0,64	0,31	0,40	0,81	0,30	0,33	0,68	0,44	0,33	0,41	0,25
	θ	0,42	0,63	0,46	0,43	0,69	0,39	0,34	0,50	0,42	0,31	0,41	0,34
750	a	0,24	0,67	0,12	0,38	0,60	0,16	0,17	0,37	0,12	0,12	0,12	0,13
	b_1	0,10	0,19	0,14	0,23	0,36	0,08	0,12	0,39	0,09	0,12	0,09	0,11
	b_2	0,08	0,10	0,09	0,13	0,18	0,08	0,09	0,20	0,12	0,11	0,14	0,09
	b_3	0,05	0,14	0,09	0,10	0,20	0,10	0,09	0,21	0,16	0,12	0,17	0,10
	b_4	0,05	0,25	0,09	0,13	0,29	0,15	0,13	0,31	0,23	0,17	0,21	0,13
	b_5	0,06	0,35	0,12	0,21	0,36	0,19	0,18	0,39	0,28	0,23	0,24	0,17
	b_6	0,11	0,64	0,21	0,33	0,68	0,30	0,27	0,66	0,37	0,30	0,31	0,26
	θ	0,84	0,75	0,79	0,92	0,67	0,82	0,78	0,71	0,81	0,73	0,72	0,88

Tabela 22 – Viés médio das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação forte entre os traços latentes (ρ_{f_0}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	-0,26	-0,35	-0,04	-0,26	-0,45	0,00	-0,02	-0,28	0,22	-0,01	0,15	0,11
	b_1	-0,39	-0,41	-0,41	-0,48	-0,61	-0,27	-0,42	-0,56	-0,41	-0,53	-0,47	-0,44
	b_2	-0,22	-0,28	-0,28	-0,40	-0,32	-0,21	-0,43	-0,38	-0,55	-0,49	-0,57	-0,47
	b_3	-0,15	-0,28	-0,21	-0,34	-0,17	-0,17	-0,41	-0,28	-0,59	-0,47	-0,59	-0,45
	b_4	-0,10	-0,33	-0,12	-0,38	-0,02	-0,10	-0,44	-0,18	-0,70	-0,49	-0,65	-0,47
	b_5	-0,05	-0,39	-0,02	-0,45	0,10	-0,02	-0,50	-0,12	-0,73	-0,51	-0,63	-0,43
	b_6	0,14	-0,21	0,03	-0,20	0,48	0,26	-0,44	0,11	-0,61	-0,36	-0,47	-0,27
θ	-0,13	-0,09	-0,12	-0,17	-0,19	-0,18	-0,28	-0,31	-0,30	-0,35	-0,39	-0,39	
150	a	-0,21	-0,43	-0,09	-0,32	-0,49	-0,04	-0,08	-0,32	0,09	0,00	0,10	0,06
	b_1	-0,21	-0,27	-0,23	-0,32	-0,50	-0,15	-0,30	-0,40	-0,25	-0,30	-0,24	-0,24
	b_2	-0,11	-0,18	-0,17	-0,27	-0,20	-0,13	-0,29	-0,23	-0,31	-0,27	-0,35	-0,23
	b_3	-0,08	-0,13	-0,12	-0,23	-0,10	-0,07	-0,28	-0,15	-0,34	-0,26	-0,38	-0,21
	b_4	-0,04	-0,15	-0,09	-0,23	0,08	-0,02	-0,30	-0,04	-0,41	-0,29	-0,47	-0,22
	b_5	-0,02	-0,20	-0,05	-0,28	0,20	0,05	-0,36	0,08	-0,43	-0,33	-0,46	-0,17
	b_6	0,06	0,05	0,02	-0,09	0,54	0,18	-0,31	0,29	-0,39	-0,29	-0,39	-0,08
θ	-0,05	-0,04	-0,04	-0,08	-0,10	-0,09	-0,18	-0,17	-0,18	-0,17	-0,20	-0,20	
300	a	-0,22	-0,46	-0,04	-0,30	-0,51	-0,03	-0,05	-0,33	0,01	0,01	0,10	0,02
	b_1	-0,13	-0,17	-0,16	-0,23	-0,32	-0,09	-0,16	-0,33	-0,11	-0,14	-0,12	-0,12
	b_2	-0,08	-0,07	-0,12	-0,15	-0,10	-0,01	-0,13	-0,12	-0,14	-0,14	-0,18	-0,12
	b_3	-0,06	-0,03	-0,07	-0,12	0,02	0,02	-0,12	-0,03	-0,14	-0,14	-0,20	-0,11
	b_4	-0,04	0,00	-0,05	-0,09	0,19	0,08	-0,14	0,11	-0,19	-0,18	-0,25	-0,12
	b_5	-0,01	0,01	-0,03	-0,12	0,26	0,10	-0,17	0,21	-0,17	-0,21	-0,26	-0,09
	b_6	0,05	0,13	0,02	-0,03	0,58	0,20	-0,15	0,51	-0,11	-0,19	-0,24	-0,05
θ	-0,05	-0,03	-0,05	-0,05	-0,06	-0,05	-0,09	-0,09	-0,09	-0,10	-0,11	-0,11	
500	a	-0,19	-0,53	-0,14	-0,24	-0,56	-0,05	-0,06	-0,34	-0,01	0,03	0,06	0,00
	b_1	-0,13	-0,18	-0,17	-0,19	-0,30	-0,07	-0,11	-0,27	-0,05	-0,07	-0,09	-0,10
	b_2	-0,07	-0,02	-0,07	-0,13	-0,02	-0,01	-0,08	-0,06	-0,07	-0,09	-0,15	-0,08
	b_3	-0,04	0,05	-0,01	-0,08	0,11	0,03	-0,07	0,04	-0,06	-0,09	-0,17	-0,07
	b_4	-0,03	0,14	0,07	-0,06	0,27	0,07	-0,06	0,19	-0,07	-0,11	-0,21	-0,07
	b_5	-0,01	0,18	0,09	-0,07	0,38	0,10	-0,08	0,27	-0,05	-0,12	-0,22	-0,06
	b_6	0,07	0,43	0,20	0,03	0,75	0,20	-0,03	0,54	-0,03	-0,13	-0,20	-0,04
θ	-0,04	-0,02	-0,03	-0,03	-0,01	-0,03	-0,05	-0,05	-0,05	-0,06	-0,08	-0,08	
750	a	-0,18	-0,53	-0,09	-0,31	-0,54	-0,05	-0,09	-0,33	-0,04	0,02	0,04	0,02
	b_1	-0,06	-0,13	-0,09	-0,19	-0,31	-0,05	-0,08	-0,24	-0,03	-0,07	-0,07	-0,06
	b_2	-0,05	0,01	-0,04	-0,09	-0,04	0,02	-0,05	-0,03	0,01	-0,07	-0,09	-0,06
	b_3	-0,03	0,08	-0,02	-0,04	0,08	0,05	-0,03	0,06	0,02	-0,07	-0,10	-0,05
	b_4	0,00	0,16	0,01	0,02	0,24	0,09	-0,02	0,19	0,06	-0,10	-0,13	-0,06
	b_5	0,01	0,18	0,05	0,04	0,32	0,13	-0,03	0,27	0,08	-0,12	-0,12	-0,05
	b_6	0,05	0,46	0,12	0,15	0,65	0,21	0,02	0,50	0,12	-0,11	-0,12	-0,03
θ	-0,02	-0,01	-0,01	-0,03	-0,02	-0,03	-0,03	-0,04	-0,03	-0,06	-0,05	-0,05	

Tabela 23 – Correlação média das estimativas dos parâmetros dos itens ($a, b_1, b_2, b_3, b_4, b_5, b_6$) e do traço latente (θ) em cada subteste (S1, S2 e S3) do modelo de resposta gradual multiunidimensional estimado via amostrador de Gibbs, considerando associação forte entre os traços latentes (ρ_{f_0}) e diferentes tamanhos amostrais e tamanhos do teste

n	Par	8			12			16			22		
		S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
75	a	0,29	0,51	0,17	0,71	0,53	0,81	0,75	0,77	0,50	0,76	0,79	0,74
	b_1	0,36	0,89	0,91	0,95	0,93	0,89	0,93	0,95	0,89	0,89	0,91	0,87
	b_2	0,79	0,91	0,97	0,97	0,95	0,81	0,95	0,96	0,81	0,95	0,88	0,94
	b_3	0,91	0,89	0,97	0,97	0,95	0,77	0,96	0,94	0,77	0,95	0,88	0,90
	b_4	0,83	0,88	0,94	0,97	0,91	0,68	0,95	0,89	0,69	0,96	0,84	0,80
	b_5	0,86	0,90	0,93	0,97	0,87	0,63	0,95	0,89	0,64	0,95	0,82	0,67
	b_6	0,61	0,75	0,83	0,85	0,72	0,39	0,89	0,65	0,65	0,83	0,72	0,77
	θ	0,90	0,78	0,88	0,89	0,76	0,92	0,93	0,85	0,90	0,95	0,90	0,94
150	a	0,46	0,70	-0,02	0,84	0,69	0,70	0,86	0,84	0,66	0,88	0,86	0,76
	b_1	0,27	0,91	0,96	0,96	0,92	0,92	0,96	0,97	0,95	0,95	0,95	0,94
	b_2	0,80	0,96	0,98	0,99	0,96	0,85	0,98	0,97	0,87	0,98	0,92	0,97
	b_3	0,89	0,96	0,97	0,99	0,97	0,86	0,98	0,96	0,86	0,98	0,91	0,94
	b_4	0,85	0,95	0,96	0,98	0,95	0,75	0,98	0,93	0,80	0,97	0,88	0,89
	b_5	0,88	0,97	0,96	0,98	0,94	0,71	0,98	0,89	0,79	0,97	0,88	0,78
	b_6	0,69	0,86	0,85	0,94	0,80	0,41	0,93	0,75	0,75	0,91	0,80	0,80
	θ	0,90	0,77	0,88	0,90	0,75	0,92	0,94	0,86	0,90	0,95	0,91	0,94
300	a	0,70	0,90	0,29	0,92	0,82	0,92	0,92	0,88	0,83	0,94	0,94	0,87
	b_1	0,25	0,96	0,96	0,98	0,96	0,96	0,98	0,97	0,98	0,98	0,97	0,96
	b_2	0,89	0,97	0,99	0,99	0,98	0,92	0,99	0,98	0,93	0,99	0,96	0,98
	b_3	0,97	0,97	0,99	0,99	0,98	0,90	0,98	0,98	0,91	0,99	0,96	0,96
	b_4	0,95	0,98	0,98	0,99	0,97	0,87	0,98	0,95	0,85	0,98	0,94	0,94
	b_5	0,94	0,98	0,98	0,99	0,96	0,81	0,98	0,94	0,86	0,98	0,93	0,89
	b_6	0,81	0,95	0,95	0,97	0,92	0,69	0,97	0,84	0,86	0,96	0,88	0,91
	θ	0,91	0,79	0,88	0,90	0,75	0,92	0,94	0,86	0,91	0,95	0,91	0,94
500	a	0,69	0,92	0,67	0,96	0,88	0,94	0,96	0,92	0,86	0,97	0,95	0,92
	b_1	0,37	0,97	0,98	0,99	0,98	0,98	0,99	0,99	0,98	0,98	0,98	0,98
	b_2	0,92	0,98	0,99	0,99	0,99	0,94	0,99	0,99	0,95	0,99	0,98	0,99
	b_3	0,97	0,98	0,99	1,00	0,99	0,95	0,99	0,98	0,92	0,99	0,97	0,98
	b_4	0,97	0,99	0,99	0,99	0,98	0,90	0,99	0,95	0,89	0,99	0,97	0,96
	b_5	0,97	0,99	0,99	0,99	0,98	0,85	0,99	0,93	0,89	0,99	0,97	0,93
	b_6	0,87	0,97	0,96	0,98	0,95	0,60	0,98	0,85	0,88	0,98	0,94	0,95
	θ	0,91	0,77	0,88	0,90	0,74	0,92	0,94	0,86	0,91	0,95	0,91	0,94
750	a	0,73	0,94	0,56	0,98	0,91	0,96	0,97	0,94	0,90	0,98	0,98	0,93
	b_1	0,42	0,99	0,98	0,99	0,99	0,98	0,99	0,99	0,99	0,99	0,99	0,98
	b_2	0,96	0,99	1,00	1,00	0,99	0,95	1,00	0,99	0,96	0,99	0,98	0,99
	b_3	0,98	0,99	1,00	1,00	0,99	0,95	1,00	0,99	0,94	1,00	0,98	0,98
	b_4	0,98	0,99	0,99	0,99	0,99	0,92	0,99	0,96	0,93	0,99	0,98	0,97
	b_5	0,98	0,99	0,99	0,99	0,98	0,90	0,99	0,95	0,92	0,99	0,97	0,93
	b_6	0,93	0,97	0,98	0,99	0,96	0,77	0,98	0,87	0,93	0,98	0,96	0,95
	θ	0,61	0,65	0,63	0,57	0,76	0,64	0,68	0,72	0,64	0,72	0,71	0,59

Tabela 24 – Correlação média entre os traços latentes estimados do modelo de resposta gradual multiunidimensional nos diferentes tamanhos amostrais (75, 150, 300, 500 e 750), tamanhos de teste (8, 12, 16 e 22) e níveis de associação (fraca, média e forte).

<i>n</i>	8			12			16			22		
	ρ_{12}	ρ_{23}	ρ_{13}	ρ_{12}	ρ_{23}	ρ_{13}	ρ_{12}	ρ_{23}	ρ_{13}	ρ_{12}	ρ_{23}	ρ_{13}
Fraca												
75	0,99	0,97	0,96	0,23	0,99	0,34	0,31	0,53	0,20	0,25	0,36	0,12
150	0,99	0,97	0,96	0,23	0,99	0,35	0,28	0,49	0,14	0,23	0,34	0,11
300	0,99	0,97	0,96	0,20	0,99	0,32	0,27	0,48	0,12	0,24	0,34	0,12
500	0,99	0,97	0,97	0,20	0,99	0,33	0,27	0,47	0,12	0,24	0,36	0,12
750	0,99	0,97	0,97	0,22	0,99	0,34	0,27	0,44	0,11	0,23	0,36	0,11
Média												
75	1,00	1,00	1,00	0,65	0,97	0,81	0,73	0,69	0,49	0,69	0,59	0,45
300	1,00	1,00	1,00	0,64	0,97	0,80	0,73	0,68	0,50	0,68	0,58	0,44
150	1,00	1,00	1,00	0,63	0,97	0,80	0,73	0,69	0,51	0,67	0,57	0,46
500	1,00	1,00	1,00	0,64	0,97	0,81	0,72	0,68	0,51	0,69	0,59	0,46
750	1,00	1,00	1,00	0,64	0,97	0,80	0,72	0,68	0,50	0,68	0,58	0,45
Forte												
75	1,00	1,00	1,00	0,90	0,95	0,99	0,86	0,93	0,98	0,79	0,88	0,95
150	1,00	1,00	1,00	0,90	0,95	0,99	0,86	0,93	0,98	0,77	0,87	0,95
300	1,00	1,00	1,00	0,90	0,96	0,99	0,86	0,93	0,98	0,78	0,88	0,95
500	1,00	1,00	1,00	0,90	0,96	0,99	0,86	0,93	0,98	0,79	0,88	0,95
750	1,00	1,00	1,00	0,90	0,95	0,99	0,86	0,93	0,98	0,79	0,88	0,95