



Universidade de Brasília  
Instituto de Ciências Exatas  
Departamento de Estatística

Dissertação de Mestrado

Uma adaptação do modelo de resposta ao item para  
mensuração de heterogeneidade atribuída  
à fonte desconhecida

por

Filipe Pereira da Silva

Orientador: Prof. Dr. Afrânio Márcio Corrêa Vieira

Junho de 2012

Filipe Pereira da Silva

**Uma adaptação do modelo de resposta ao item para  
mensuração de heterogeneidade atribuída  
à fonte desconhecida**

Dissertação apresentada ao Programa de Pós-Graduação em  
Estatística do Instituto de Ciências Exatas da Universidade  
de Brasília como requisito parcial à obtenção de título  
de Mestre em Estatística.

Universidade de Brasília  
Brasília, Junho de 2012

TERMO DE APROVAÇÃO

Filipe Pereira da Silva

UMA ADAPTAÇÃO DO MODELO DE RESPOSTA AO ITEM PARA  
MENSURAÇÃO DE HETEROGENEIDADE ATRIBUÍDA  
À FONTE DESCONHECIDA

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Data da defesa: 25 de junho de 2012

Orientador:

---

Prof. Dr. Afrânio Márcio Corrêa Vieira -Orientador  
Departamento de Estatística, UnB

Comissão Examinadora:

---

Prof. Dr. Donald Matthew Pianto - Membro da Banca  
Departamento de Estatística, UnB

---

Prof. Dr. Dalton Francisco de Andrade - Membro Externo da Banca  
Departamento de Informática e Estatística, Universidade Federal de Santa Catarina

Brasília, Junho de 2012

## Ficha Catalográfica

**DA SILVA, FILIPE PEREIRA**

Uma adaptação do modelo de resposta ao item para mensuração de heterogeneidade atribuída à fonte desconhecida, (UnB - IE, Mestre em Estatística, 2012).

Dissertação de Mestrado - Universidade de Brasília. Departamento de Estatística - Instituto de Ciências Exatas.

1. Teoria de Resposta ao Item 2. Modelos Lineares Generalizados Mistos 3. Prova Brasil 4. Superdispersão.

É concedida à Universidade de Brasília a permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta monografia de Projeto Final pode ser reproduzida sem a autorização por escrito do autor.

Filipe Pereira da Silva

*À Deus,*  
*Aos meus pais,*  
*À minha família*

# Agradecimentos

Agradeço a Deus por ter sido a minha força neste trabalho.

Aos meus pais, por terem me ajudado nesta caminhada com palavras de conforto, por terem acreditado em mim e investido na minha formação. Agradeço à minha namorada por ter me consolado nas horas difíceis e por se alegrar comigo nas boas notícias. À minha irmã, pelo carinho espontâneo que recebo. Agradeço a toda minha família pela ajuda e confiança depositada.

Aos meus colegas de trabalho e à minha chefia, por me passarem tranquilidade e por me apoiarem neste grande passo que estou dando.

Aos meus amigos de curso, pelos momentos felizes e pela cooperação mútua.

E ao meu orientador, por ter sempre acreditado no meu potencial. Agradeço à ele por ter sido paciente e amigo, por tentar extrair o máximo de mim e por nunca perder a fé.

# Sumário

<b>Lista de Figuras</b>	<b>3</b>
<b>Resumo</b>	<b>4</b>
<b>Abstract</b>	<b>5</b>
<b>1 Introdução</b>	<b>6</b>
<b>2 Revisão Metodológica</b>	<b>8</b>
2.1 Métodos para otimização de funções . . . . .	8
2.1.1 Método de Newton-Raphson . . . . .	8
2.1.2 Método <i>scoring</i> de Fisher . . . . .	9
2.1.3 Método da secante . . . . .	9
2.1.4 Métodos Quasi-Newton . . . . .	10
2.2 Métodos de integração numérica . . . . .	10
2.2.1 Fórmulas de Newton-Côtes . . . . .	11
2.2.2 Quadratura Gaussiana . . . . .	11
2.3 Modelos Lineares Generalizados . . . . .	11
2.3.1 Introdução . . . . .	11
2.3.2 Família Exponencial . . . . .	12
2.3.3 Família Exponencial com parâmetro de dispersão . . . . .	13
2.3.4 Modelos Lineares Generalizados . . . . .	13
2.4 Modelos Lineares Generalizados Mistos . . . . .	17
2.4.1 Introdução . . . . .	17
2.4.2 MLGM para dados longitudinais . . . . .	18
2.5 Modelos combinados para dados com superdispersão . . . . .	19

2.5.1	Modelo combinado Logit-Bernoulli-Normal-Beta . . . . .	20
2.5.2	Modelo combinado Probit-Bernoulli-Normal-Beta . . . . .	23
<b>3</b>	<b>Teoria de Resposta ao Item - TRI</b>	<b>26</b>
3.1	Definição e pressupostos . . . . .	27
3.2	Modelo de TRI com 3 parâmetros . . . . .	28
3.3	Interpretação dos parâmetro e o gráfico CCI . . . . .	29
3.4	Identificabilidade . . . . .	30
3.5	Escala de habilidade . . . . .	31
3.6	Estimação dos Parâmetros . . . . .	31
3.7	Modelo de Rasch . . . . .	32
3.8	Funcionamento Diferencial do Item - DIF . . . . .	33
<b>4</b>	<b>Modelo de Rasch Estendido para Heterogeneidade Atribuída à Fonte Desconhecida</b>	<b>34</b>
4.1	Definição do modelo . . . . .	34
4.2	Uma nova parametrização da distribuição Beta . . . . .	36
4.3	Simulação de Monte Carlo para o modelo de Rasch Estendido . . . . .	38
<b>5</b>	<b>Aplicação do Modelo de Rasch estendido aos dados da Prova Brasil 2007</b>	<b>43</b>
5.1	SAEB - Sistema de Avaliação da Educação Básica . . . . .	44
5.2	Aplicação aos dados da Prova Brasil 2007 . . . . .	46
5.3	Apontamentos da existência de uma possível heterogeneidade no banco de dados . . . . .	47
5.3.1	Comparação étnico-racial . . . . .	48
5.3.2	Comparação relacionada ao trabalho fora de casa . . . . .	49
5.3.3	Comparação relacionada a reprovação dos alunos . . . . .	50
5.3.4	Comentários . . . . .	51
5.4	Aplicação do Modelo de Rasch Estendido aos dados do SAEB . . . . .	52
5.5	Conclusões e perspectivas de trabalhos futuros . . . . .	54
	<b>Referências Bibliográficas</b>	<b>57</b>



# Lista de Figuras

3.1	Curva Característica do Item - CCI (ANDRADE et al.,2002) . . . . .	30
4.1	Gráfico da distribuição beta com $\alpha_1$ fixo em “1” e $\alpha_2$ variável . . . . .	39
5.1	CCI’s comparativas - Questão 2 . . . . .	49
5.2	CCI’s comparativas - Questão 35 . . . . .	50
5.3	CCI’s comparativas - Questão 38 . . . . .	51
5.4	CCI’s comparativas dos modelos ajustados . . . . .	54

# Resumo

A Teoria de Resposta ao Item (TRI) caracteriza-se por ser um conjunto de modelos matemáticos cujo propósito é modelar a probabilidade de uma resposta de um item  $j$  por um indivíduo  $i$  em função de parâmetros associados ao item e de traços latentes associados ao indivíduo. No entanto, outros fatores externos e desconhecidos podem influenciar esta probabilidade de acerto, causando heterogeneidade nos dados (superdispersão), muitas vezes não ajustada pelos modelos de resposta ao item.

Por outro lado, Molenberghs, Verbeke, Demétrio e Vieira (2010) descreveram alguns modelos no campo dos Modelos Lineares Generalizados Mistos (MLGM) para tratar dados binários longitudinais com superdispersão. Tais modelos que conseguem capturar este efeito acrescentando um efeito aleatório com distribuição conjugada que, desta forma, captura esta heterogeneidade atribuída à fonte desconhecida.

O modelo de resposta ao item de Rasch foi reescrito de forma que ele pudesse ser tratado como membro da família dos MLGM e um novo modelo que capture a superdispersão dos dados foi proposto e definido. Além disso, utilizando resultados analíticos apresentados por Molenberghs et al. (2010), definiu-se o modelo de Rasch estendido analítico (REMol), sendo este uma simplificação do modelo de Rasch estendido com ligação probit.

Por fim, os modelos propostos e o modelo de Rasch tradicional são aplicados a dados da Prova Brasil 2007. Os resultados demonstram que os ajustes dos modelos são similares, exceto o modelo REMol, que gerou estimativas e erros padrões distintos, indicando que capturou uma maior variabilidade dos dados, cumprindo a proposta inicial do modelo.

**Palavras Chave:** *Superdispersão, Teoria de Resposta ao Item, Modelo de Rasch, Heterogeneidade, Modelos Lineares Generalizados Mistos.*

# Abstract

The Item Response Theory (IRT) is characterized as a set of mathematical models whose purpose is fit the probability of some response of the  $i$ -th subject to respond correctly the  $j$ -th item in function of the items' parameters and ability's subject. However, others external and unknown factors may influence this probability causing heterogeneity (overdispersion) in the data, oftentimes not fitted by the item response models.

On the other hand, Molenberghs, Verbeke, Demétrio e Vieira (2010) described some models in the field Generalized Linear Mixed Models (GLMM) to treat binary longitudinal data with overdispersion. Such models capture this effects by adding an random parameters with conjugated distribution that, this way, it capture this heterogeneity from unknown source.

The Rasch model was rewritten so that it belongs to GLMM's family and a new model that capture the overdispersion from the data was proposed. Furthermore, using the analytical results shown by Molenberghs et al.(2010), it was define the analytical extended Rasch model (REMo), being this a simplification to extended Rasch model with probit link function.

Finally, the proposed models and traditional Rasch model was applied to data from Prova Brasil 2007. The results show that the fits of the models are similar, except the REMo model. This model found distinct estimates and standard errors adjusted to the greater variability from the data, fulfilling the initial proposal of the model.

**Palavras Chave:** *Overdispersion, Item Response Theory, Rasch Model, Heterogeneity, Generalized Linear Mixed Models.*

# Capítulo 1

## Introdução

Ao se realizar um teste (avaliação ou prova de conhecimentos), é muito comum calcular somente o escore total do indivíduo, o número de acertos ou até mesmo percentuais de itens corretamente respondidos. Tal abordagem tem como suporte a Teoria Clássica dos Testes (TCT), onde o foco está no desempenho do indivíduo na prova como um todo (De AYALA, 2009). No entanto, para uma comparação dos respondentes, utilizam-se os mesmos testes e os mesmos instrumentos de medida.

Na década de 1950, uma nova teoria foi proposta para revolucionar o referido cenário: a Teoria de Resposta ao Item (TRI). Esta tem como objetivo principal “representar a probabilidade de um indivíduo responder a um certo item como função dos parâmetros do item e de certos traços latentes (ou habilidades)” (ANDRADE; TAVARES; VALLE, 2000). A TRI tem como grande diferença da TCT o fato dela se basear nas estruturas dos itens, e não na prova como um todo, tornando assim provas com itens diferentes comparáveis, devido ao grau de dificuldade dos itens que foram colocados no instrumento de avaliação - cada item tem um grau de dificuldade específico e consegue discriminar os indivíduos de uma maneira diferente. Esta teoria é utilizada no Exame Nacional do Ensino Médio - ENEM, no Sistema Nacional de Avaliação da Educação Básica - SAEB, no Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo - SARESP e em vários outros aqui no Brasil.

Os modelos tradicionais de resposta ao item pressupõem, em especial, independência condicional entre as respostas de diferentes indivíduos. Tais indivíduos podem ser estudantes que pertencem a agrupamentos diversos (e.g. provenientes de escolas

públicas ou privadas, diferentes classes socio-econômicas, dentre outras), que são potenciais fontes de heterogeneidade de fonte desconhecidas (ou conhecidas mas não consideradas na modelagem dos dados) e que tendem a inflacionar a variabilidade observada dos resultados dos testes ou exames. A consequência é menor precisão na inferência sobre os parâmetros dos itens e nas habilidades latentes estimadas para os respondentes.

No entanto, Molenberghs, Verbeke, Demétrio e Vieira (2010) apresentaram resultados que estendem os modelos lineares generalizados (MLG) sobre modelos de distribuições conjugadas com um parâmetro de efeito de aleatório, onde este problema é tratado. A grande ideia presente no artigo é acrescentar parâmetros com efeitos aleatório que irá captar toda esta heterogeneidade como variação extra-Bernoulli.

O objetivo desta pesquisa é adaptar o *modelo de Resposta ao Item de 1 parâmetro* (modelo de Rasch) para a família de modelos de Molenbergh et al. (2010), criando assim um novo modelo que trata a heterogeneidade atribuída à fonte desconhecida. Após isto, utilizar um grande banco de dados para observar como esse modelo se comporta. Além disso, alguns pressupostos dos modelos de resposta ao item poderão ser testados como, por exemplo, o princípio da independência entre as respostas de diferentes indivíduos.

No primeiro capítulo tem-se a apresentação de uma revisão bibliográfica que abordará alguns métodos de otimização de curvas e integração numérica, tópicos relacionados aos Modelos Lineares Generalizados - MLG e aos Modelos Lineares Generalizados Mistos - MLGM. Finalizando a seção, o artigo de Molenberghs et al. (2010), o texto-base da pesquisa, será apresentado.

Em seguida, no segundo capítulo, o modelo de TRI será definido e alguns aspectos, como os pressupostos do modelo, a interpretação de seus parâmetros, o Funcionamento Diferencial do Item - DIF, serão apresentados. Após isto, o modelo de Rasch estendido será definido e os seus resultados analíticos serão apresentados. Neste mesmo capítulo, será apresentado um breve estudo de simulação via Monte Carlo comparando desempenho do modelo de Rasch estendido com o modelo de Rasch tradicional em cenários previamente estabelecido. Por fim, os modelos de Rasch e de Rasch estendidos demonstrados no capítulo anterior serão aplicados aos dados da Prova Brasil 2007, a fim de colaborar na reflexão sobre a importância do modelo.

# Capítulo 2

## Revisão Metodológica

Neste capítulo tem-se uma revisão metodológica de algumas técnicas e classes de modelos que são importantes para o entendimento da proposta de pesquisa. Primeiro, serão estudados alguns métodos de otimização de funções e técnicas para a integração numérica. Em seguida, serão definidos os Modelos Lineares Generalizados (MLG) e os Modelos Lineares Generalizados Mistos (MLGM). Por fim, será estudado um artigo publicado por Molenberghs, Verbeke, Demétrio e Vieira (2010), que servirá de texto-base para esse trabalho.

### 2.1 Métodos para otimização de funções

#### 2.1.1 Método de Newton-Raphson

O método de Newton-Raphson é um dos mais conhecidos para encontrar raízes de funções. Este método é baseado na aproximação de funções reais por expansões em séries infinitas de Taylor, é bastante rápido e não necessita de muitas iterações para convergir. Tal convergência é sempre garantida para um intervalo  $[a,b]$  onde as derivadas primeiras e segundas,  $f'(x)$  e  $f''(x)$ , respectivamente, sejam contínuas e  $f''(x)$  seja diferente de 0 (NOCEDAL; WRIGHT, 1999; YANG; CAO; MORRIS, 2005). Assim, basta que um valor inicial  $x_0$  esteja dentro deste intervalo para haver convergência.

A fórmula iterativa é a seguinte:

$$x^{(t+1)} = x^{(t)} - \frac{f'(x^{(t)})}{f''(x^{(t+1)})} \quad (2.1)$$

Um caso especial é quando necessitamos estimar algum parâmetro utilizando o Método da Máxima Verossimilhança. Se se quer encontrar o ponto  $\theta$  que maximiza a função de log-verossimilhança  $l(\theta)$ , tem-se a seguinte função:

$$\theta^{(t+1)} = \theta^{(t)} - \frac{l'(\theta^{(t)})}{l''(\theta^{(t)})} \quad (2.2)$$

Para uma função multivariada, o método é bem similar. Dado que  $\mathbf{X} = (X_1, \dots, X_p)^t$  é um vetor  $p$ -dimensional, a fórmula iterativa para calcular o zero de  $F(\mathbf{X})$  é:

$$\mathbf{X}^{(t+1)} = \mathbf{X}^{(t)} - \frac{f'(\mathbf{X}^{(t)})}{f''(\mathbf{X}^{(t)})} \quad (2.3)$$

No caso da maximização da função de log-verossimilhança multivariada, tem-se:

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - l'(\boldsymbol{\theta}^{(t)})l''(\boldsymbol{\theta}^{(t)})^{-1} \quad (2.4)$$

onde  $l'(\boldsymbol{\theta}^{(t)})$  é a função escore e  $l''(\boldsymbol{\theta}^{(t)})$  é a matriz Hessiana.

### 2.1.2 Método *scoring* de Fisher

Nesse caso, a matriz Hessiana é substituída pela matriz score de Fisher  $I(\boldsymbol{\theta})$ , sendo que ela é uma aproximação para a função  $-l''(\boldsymbol{\theta})$  (VIEIRA, 2008). No entanto, a matriz de informação de Fisher  $I(\boldsymbol{\theta})$  é usualmente desconhecida, assim ela é aproximada por  $-E \left[ \frac{\delta^2 l(\boldsymbol{\theta})}{\delta \boldsymbol{\theta} \delta \boldsymbol{\theta}^t} \right]$ . Logo, a função iterativa pelo método *scoring* de Fisher é:

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} + l'(\boldsymbol{\theta}^{(t)}) \left( -E \left[ \frac{\delta^2 l(\boldsymbol{\theta})}{\delta \boldsymbol{\theta} \delta \boldsymbol{\theta}^t} \right] \right)^{-1} \quad (2.5)$$

### 2.1.3 Método da secante

O método da secante é limitado pois só pode ser aplicado a funções univariadas. No entanto, o método da secante tende a convergir em casos onde o método de Newton-Rapson não converge como, por exemplo, próximo dos limites do intervalo e tem como grande vantagem o fato de não ser necessário o cálculo de  $f''(x)$ .

A função iterativa é:

$$x^{(t+1)} = x^{(t)} - f'(x^{(t)}) \frac{x^{(t)} - x^{(t-1)}}{f'(x^{(t)}) - f'(x^{(t-1)})} \quad (2.6)$$

Para iniciar este método são necessários dois valores, os valores  $x_0$  e  $x_1$ , dentro do intervalo  $[a, b]$ .

### 2.1.4 Métodos Quasi-Newton

Apesar do método de Newton-Raphson ser bastante efetivo, o cálculo da matriz jacobiana da função  $F(X)$  (ou a matrix Hessiana de log-verossimilhança) e a sua atualização a cada iteração são grandes complicadores. No entanto, segundo Yang et al. (2005), existe uma classe de métodos que propõe a substituição da matriz Hessiana por uma outra matriz de pesos  $B$  que também é atualizada a cada iteração. A sequência  $\mathbf{x}^{(k)}$  é gerada através da fórmula:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)} \quad (2.7)$$

onde  $\mathbf{s}^{(k)}$  é o resultado do sistema linear:

$$\mathbf{B}^{(k)} \mathbf{s}^{(k)} = -\mathbf{F}(\mathbf{x}^{(k)}) \quad (2.8)$$

## 2.2 Métodos de integração numérica

Em muitos casos, principalmente na estimação de parâmetros, nos deparamos com integrais analiticamente difíceis de serem resolvidas. O que pode ser feito nestes casos é tentar aproximar o integrando por uma função mais fácil de ser integrada ou aproximar a própria integral. Quando é adotada a segunda estratégia, deseja-se encontrar uma aproximação que seja da seguinte forma:

$$\int_a^b f(x) dx \approx A_0 g(x_0) + A_1 g(x_1) + \dots + A_n g(x_n) \quad (2.9)$$

$x_i \in [a, b], i = 0, 1, \dots, n$  e  $g(\cdot)$  sendo uma função a ser especificada.

Assim, abaixo serão listados alguns métodos de aproximação de integrais, mais conhecidos como métodos de integração numérica. Mais detalhes a respeito deste métodos em Vieira (2008), Yang et al. (2005) e Nocedal e Wright(1999).



### 2.2.1 Fórmulas de Newton-Côtes

As fórmulas de Newton-Côtes são métodos de interpolação de integrais definidas em intervalo  $[a, b]$  baseados em  $n$  pontos igualmente espaçados. Sendo que:

$$x_0 = a, \quad x_i = \frac{(b-a)}{n}i + a, \quad x_n = b, \quad i = 0, 1, \dots, n \quad (2.10)$$

e tem-se que o espaçamento  $h = \frac{(b-a)}{n} = x_{i+1} - x_i$ . Assim, seguindo a forma especificada mais acima, basta trocar a função  $g(\cdot)$  pela função  $f(\cdot)$  a ser integrada e a aproximação será definida.

$$\int_a^b f(x)dx \approx A_0f(x_0) + A_1f(x_1) + \dots + A_nf(x_n) \quad (2.11)$$

### 2.2.2 Quadratura Gaussiana

Em certos casos, os métodos anteriormente listados podem não ter o desempenho esperado e podem até falhar no cálculo da aproximação da integral. Uma alternativa que pode contornar esse problema é utilizar a Quadratura Gaussiana - método muito mais eficiente que os anteriores, uma vez que não há a necessidade de que os pontos que definem os subintervalos estejam igualmente espaçados e os pesos da soma em (2.11) são definidos em função da ordem da aproximação. Por definição, tem-se

$$\int_a^b f(x)dx \approx \sum_{q=1}^Q w_q f(w_q), \quad (2.12)$$

onde  $Q$  é a ordem da aproximação. Uma estratégia muito utilizada é transformar a integral  $\int_a^b f(x)dx$  para  $\int_{-1}^1 f(a + \frac{b-a}{2}(t+1))\frac{b-a}{2}dt$ , e depois calcular estes pesos. Uma particularidade deste método é que ele é exato quando para polinômios de grau menor ou igual a  $2Q + 1$ .

## 2.3 Modelos Lineares Generalizados

### 2.3.1 Introdução

Na década de 30, Kopman, Pitman e Darnois propuseram uma nova família de distribuições de probabilidade onde, dentre elas, estavam a Normal, a binomial, a beta,

a bernoulli, e muitas outras distribuições que conhecemos hoje. Esta nova classe foi denominada *Família Exponencial*, e a sua criação impulsionou grandes avanços e a descoberta de importantes resultados na Estatística devido as propriedades especiais que os seus membros possuem (DEMÉTRIO; CORDEIRO, 2007).

Já em 1972, Nelder e Wedderburn publicaram um artigo de extrema relevância que agrupou diversas metodologias e técnicas estatísticas criando uma nova classe de modelos: os *Modelos Lineares Generalizados*. No entanto, estes modelos tinham como um de seus pressupostos o fato da variável de estudo pertencer a família exponencial. Logo, primeiramente, faz-se necessário definir tal família e discorrer sobre suas propriedades mais importantes. Finalizando isto, discutir-se-á sobre os Modelos Lineares Generalizados: definição, estimativas de seus parâmetros e análise dos modelos.

### 2.3.2 Família Exponencial

Dada uma variável aleatória  $Y$  com função de densidade de probabilidade dependente unicamente do parâmetro  $\theta$ . Ela pertencerá a família exponencial de sua distribuição for da seguinte forma:

$$f(y) = h(y)\exp[\eta(\theta)t(y) - v(\theta)] \quad (2.13)$$

onde  $h(y)$ ,  $\eta(\theta)$ ,  $t(y)$  e  $v(\theta)$  são funções conhecidas.

Esta classificação não se restringe ao caso univariado e nem ao tipo de variável aleatória que está sendo estudada: contínua ou discreta. Há ainda uma outra forma de apresentar tal classe, definindo que as funções  $\eta(\theta)$  e  $t(y)$  são iguais a funções identidades e  $v(\theta)$  e  $h(y)$  serem funções conhecidas, não necessariamente iguais as funções apresentadas na fórmula anterior.

$$f(y) = h(y)\exp[\theta y - v(\theta)] \quad (2.14)$$

Nesta nova parametrização,  $\theta$  é denominado como parâmetro canônico, logo esta nova classe é definida como *família exponencial canônica*. Diante desta definição, é possível enunciar algumas propriedades importantes.

$$E(Y) = v'(\theta) \quad (2.15)$$

$$Var(Y) = v''(\theta)$$

Portanto, tem-se que a média e a variância estão conectadas ao parâmetro canônico.

### 2.3.3 Família Exponencial com parâmetro de dispersão

Agora, esta nova família é definida introduzindo um parâmetro de dispersão  $\phi^{-1} > 0$  na família exponencial canônica. Este parâmetro é um parâmetro de dispersão e foi introduzido por Nelder e Wedderburn na formulação dos MLG's. Assim, tem-se:

$$f(y) = h(y)\exp\{\phi^{-1}[\theta y - v(\theta)] + \lambda(\phi^{-1}, y)\} \quad (2.16)$$

onde as funções  $v(\cdot)$  e  $\lambda(\cdot)$  são conhecidas. Logo, as propriedades anteriormente descritas para o modelo acima são:

$$\begin{aligned} E(Y) &= v'(\theta) \\ Var(Y) &= \phi v''(\theta) \end{aligned} \quad (2.17)$$

Assim, nota-se que o parâmetro  $\phi$  está presente na fórmula da  $Var(Y)$ . Logo,  $\phi$  é um parâmetro de dispersão e o seu inverso,  $\phi^{-1}$ , é um parâmetro de precisão. A função que relaciona o parâmetro  $\theta$  com a média  $\mu$  é dado por  $q(\mu) = \theta$ , onde  $\mu = v'(\theta)$ . A função que envolve somente a média  $\mu$  na expressão da variância de  $Y$  é denominada *função de variância*. Essa função é descrita por  $V(\mu)$  e é igual a  $v''(\theta)$ . Nota-se ainda que o parâmetro canônico  $\theta$  pode ser obtido por  $\theta = \int V^{-1}(\mu)d\mu$ , pois  $\frac{d\mu}{d\theta} = v''(\theta)$ .

### 2.3.4 Modelos Lineares Generalizados

A possibilidade da criação de modelos que simulem ou, até mesmo, que modelem certos fenômenos é uma grandes contribuições da Estatística à sociedade atual. Seja na agropecuária, na indústria ou na análise risco de créditos, a modelagem é uma aliada poderosa, seja para entender o fenômeno, seja para prever comportamentos das forças que agem em um determinado evento. Assim, os modelos estatísticos vem cada vez sendo mais explorados, principalmente nesses últimos tempos, onde a informação e o conhecimento são tão valiosos.

Conforme expresso na introdução deste capítulo, Nelder e Wedderburn criaram uma nova classe de modelos que unia diversas técnicas e metodologias estatísticas em um único grupo especial de modelos lineares. Demétrio (2002) descreveu o modelo da seguinte maneira: considere-se uma amostra de  $n$  observações, sendo essas independentes e identicamente distribuídas; defina-se ainda uma matriz  $X$  com  $p + 1$  colunas

que representam  $p$  variáveis explicativas do modelo e um vetor  $y$  de observações que foram amostradas de  $Y$ . Logo, são definidos os 3 componentes de um MLG:

1. Uma variável aleatória  $Y$  membro da *família exponencial*, com uma função de probabilidade ou função densidade de probabilidade (para variáveis discretas e contínuas, respectivamente) igual a

$$f(y_i, \theta_i, \phi^{-1}) = h(y_i) \exp\{\phi^{-1}[\theta_i y_i - v(\theta_i)] + \lambda(\phi^{-1}, y_i)\} \quad (2.18)$$

e

$$\begin{aligned} E(y_i) &= v'(\theta_i) \\ \text{Var}(y_i) &= \phi v''(\theta_i) \end{aligned} \quad (2.19)$$

este componente  $Y$  é chamado de *componente sistemático* do MLG.

2. As variáveis explicativas entram no modelo como uma soma ponderada pelos fatores  $\beta$ . Assim, estas covariáveis e os fatores estão relacionados através de um *preditor linear*

$$\eta_i = \mathbf{x}_i^t \boldsymbol{\beta} \quad (2.20)$$

3. Uma função que liga o componente sistemático e o preditor linear,

$$\eta_i = g(\mu_i) \quad (2.21)$$

conhecida como *função de ligação*, sendo que esta função deve ser monótona e diferenciável.

### Estimação dos parâmetros

O objetivo agora é estimar o vetor de parâmetros  $\boldsymbol{\beta}$ , pois através deste, todas as outras estimativas podem ser realizadas utilizando as relações anteriormente descritas. No entanto, o método proposto por Nelder e Wedderburn para a estimação dos parâmetros é o método da Máxima Verossimilhança. Logo, tem-se que o logaritmo da função de verossimilhança terá a seguinte forma

$$\begin{aligned}
l(\boldsymbol{\theta}|\mathbf{y}) &= \log \left[ \prod_{i=1}^n f(y_i, \theta_i, \phi) \right] = \sum_{i=1}^n l(y_i, \theta_i) \\
&= \sum_{i=1}^n \{ \phi^{-1} [\theta_i y_i - v(\theta_i)] + \lambda(y_i, \phi_i^{-1}) \}
\end{aligned} \tag{2.22}$$

O  $j$ -ésimo componente do vetor escore é igual a

$$U_j = \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_j} = \sum_{i=1}^n \frac{dl_i}{d\theta_i} \frac{d\theta_i}{d\mu_i} \frac{d\mu_i}{d\eta_i} \frac{d\eta_i}{d\beta_i} = \sum_{i=1}^n \phi^{-1}(y_i - \mu_i) \frac{1}{V(\mu_i)} \frac{d\mu_i}{d\eta_i} x_{ij} \quad j=1, \dots, p \tag{2.23}$$

assumindo  $\phi^{-1}$  conhecido.

Assim, para a obtenção das estimativas de Máxima Verossimilhança de  $\boldsymbol{\theta}$ , deve se executar o seguinte algoritmo até que atinja-se um determinado critério de parada:

**1º Passo:**

$$\begin{aligned}
\eta_i^{(m)} &= \sum_{j=1}^p x_{ij} \beta_j^{(m)} \\
\mu_i^{(m)} &= g^{(-1)}(\eta_i^{(m)})
\end{aligned} \tag{2.24}$$

**2º Passo:**

$$\begin{aligned}
q_i^{(m)} &= \eta_i^{(m)} + (y_i - \mu_i^{(m)}) g'(\eta_i^{(m)}) \\
\mathbf{Q}^{(m)} &= [q_1^{(m)}, q_2^{(m)}, \dots, q_n^{(m)}]^T \\
W_i^{(m)} &= \frac{1}{V(\mu_i^{(m)}) [g'(\mu_i^{(m)})]^2} \\
\mathbf{W}^{(m)} &= \text{diag}\{W_i^{(m)}\}
\end{aligned} \tag{2.25}$$

**3º Passo:**

$$\boldsymbol{\beta}^{(m)} = (\mathbf{X}^T \mathbf{W}^{(m)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{(m)} \mathbf{Q}^{(m)} \tag{2.26}$$

**4º Passo:** Caso o ciclo venha a convergir,  $\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}^{(m)}$ . Se não, volta-se para o Passo 1 utilizando  $\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}^{(m)}$  no passo  $m + 1$ .

Mais detalhes sobre a estimação dos parâmetros ver em Demétrio(2000), Vieira(2008), Dobson(2002) e McCullagh(2001).

## Medidas de Discrepância

As medidas de discrepância são importantes para analisar o quão o valor predito pelo modelo estimado distante do valor real observado. Assim, Nelder e Wedderburn propuseram a seguinte medida de desvio (ou *deviance*):

$$\frac{D}{\phi} = -2 \sum_{i=1}^n l(\hat{\mu}; y_i) - l(y_i; y_i) \quad (2.27)$$

onde  $l(\hat{\mu}; y_i)$  é o valor da log-verossimilhança calculada utilizando  $y_i$  e uma estimativa de  $\mu_i$  e  $l(y_i; y_i)$  é a mesma função sendo que agora  $y_i$  é também o valor da estimativa da média. Assim, esta medida é chamada de *scaled deviance* e

$$D = \sum_{i=1}^n d_i^2, \quad (2.28)$$

denominada *deviance residual*, sendo o componente do somatório

$$d_i^2 = -2 \int_y^{\mu_i} \frac{y - u_i}{V(u_i)} du. \quad (2.29)$$

## Estimação do Parâmetro de Dispersão

Em algumas distribuições, como a Binomial e Poisson, o parâmetro de dispersão  $\phi$  é igual a 1 (DEMÉTRIO, 2002; VIEIRA, 2008). No entanto, em outros casos este parâmetro pode ser desconhecido; logo deve-se calcular uma estimativa para esse valor. Esta estimativa é necessária para a formulação dos intervalos de confiança e dos teste de hipóteses relacionados aos parâmetros  $\beta$  - ferramentas essenciais na inferência estatística. Outra utilização dessa estimativa é que ela oferece uma idéia do quão bem ajustado está um modelo. Por exemplo, se estivermos trabalhando com uma variável aleatória Bernoulli e essa estimativa  $\hat{\phi}$  for maior do que 1, isso indica que está ocorrendo o fenômeno de superdispersão, há uma variação *extra-Bernoulli* que não foi captada pelo modelo. Todavia, isso será estudado nas seções seguintes.

Demétrio (2002) define em seu trabalho alguns estimadores do parâmetro de dispersão  $\hat{\phi}$ :

- $\hat{\phi}_d = \frac{D_p}{n-p}$  - encontrado pelo método dos momentos, utilizando a função Desvio;

- $\hat{\phi}_P = \frac{1}{n-p} \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)}$  - encontrado pelo método de Pearson;
- $\hat{\phi} = \frac{2D_P}{n \left[ 1 + \left( 1 + \frac{2D_P}{3n} \right) \right]}$  - encontrado pelo método Máxima Verossimilhança.

## Análise de Resíduos

A análise de resíduos é extremamente importante para análise do ajuste do modelo. No entanto, há vários tipos de resíduos, sendo que pode-se destacar essas duas formas de cálculo:

- resíduo do componente da *deviance*:

$$r_{d_i} = \text{sign}(y_i - \hat{\mu}_i) \sqrt{d_i^2} \quad (2.30)$$

- resíduo de Person:

$$r_{P_i} = \frac{(y_i - \hat{\mu}_i)}{V(\hat{\mu}_i)} \quad (2.31)$$

## 2.4 Modelos Lineares Generalizados Mistos

### 2.4.1 Introdução

O MLGM é uma extensão do MLG, sendo que além do efeito fixo  $\beta$  no preditor linear, há um efeito aleatório, que pode ser definido como  $\mathbf{b}$ . Este modelo está ancorado na independência condicional de  $Y$  dado o efeito aleatório  $\mathbf{b}$ , na distribuição condicional  $f(y|\mathbf{b})$  e na distribuição desses efeitos aleatórios (JIANG, 2007).

Como está definido em Jiang (2007), suponha que, dado um vetor de efeitos aleatórios  $\mathbf{b}$ , as respostas  $y_1, \dots, y_n$  são condicionalmente independentes tal que a distribuição condicional de  $y_i$ , dado o vetor de efeitos aleatórios  $\mathbf{b}$  é membro da família exponencial

$$f_i(y_i|\mathbf{b}) = \exp \left\{ \phi^{-1} [y_i \theta_i - v(\theta_i)] + \lambda_i(y_i, \phi^{-1}) \right\} \quad (2.32)$$

onde  $v(\cdot)$  e  $\lambda_i(\cdot)$  são funções desconhecidas e o parâmetro de dispersão  $\phi$  pode ou não ser conhecido. O parâmetro  $\theta_i$  está associado com a média condicional  $\mu_i =$

$E(y_i|\mathbf{b})$ , associada ao preditor linear

$$\eta_i = x_i^T \boldsymbol{\beta} + z_i^T \mathbf{b} \quad (2.33)$$

onde  $x_i$  e  $z_i$  são vetores de variáveis conhecidas provenientes das matrizes  $\mathbf{X}$  e  $\mathbf{Z}$ , com dimensões  $p$  e  $q$ , respectivamente. Define-se ainda que  $\boldsymbol{\beta}$  é um vetor de parâmetros desconhecidos e fixo e a função  $g(\cdot)$ , que liga  $\mu_i$  ao preditor linear, é chamada de *função de ligação*, onde  $g(\mu_i) = \eta_i$ . E, para finalizar, é assumido que o vetor  $\mathbf{b} \sim N_q(0, \mathbf{D})$ , sendo  $\mathbf{D}$  uma matrix de variância-covariância desconhecida.

## 2.4.2 MLGM para dados longitudinais

Os MLGM's podem ser utilizados para analisar dados longitudinais, uma vez que os parâmetros de covariância dos efeitos aleatórios podem acomodar a correlação entre as medidas repetidas no tempo.

Como será mostrado mais a frente e por conveniência, será definido o MLGM para dados longitudinais, modelo este que é um caso especial do MLGM mais geral. Estes modelos são utilizados quando necessita-se modelar a ocorrência (ou não) de um evento em determinados indivíduos em diferentes tempos (MOLENBERGHS; VERBEKE, 2005). No entanto, não é necessário que os tempos sejam iguais, cada indivíduo  $i$  pode ser medido  $n_i$  vezes, em tempos totalmente diferentes, onde  $i = 1, \dots, N$ .

Suponha que  $Y_{ij}$  é a  $j$ -ésima medida de um sujeito  $i$ ,  $i = 1, \dots, N$ ,  $j = 1, \dots, n_i$  e  $\mathbf{Y}_i$  é um vetor de dimensão  $n_i$  de todas as medidas que estão disponível para o indivíduo  $i$ . Assume-se ainda que os componente do vetor  $\mathbf{b}_i$  são independentes com densidade normal multivariada, com vetor de médias  $\mathbf{0}$  e matriz de variâncias e covariâncias  $\mathbf{D}$  e as respostas  $Y_{ij}$  são independentes com as seguintes densidades:

$$f(y_{ij}|\theta_{ij}, \phi^{-1}, \mathbf{b}_i) = h(y_{ij}) \exp\{\phi^{-1}[\theta_{ij}y_{ij} - v(\theta_{ij})] + \lambda(y_{ij}, \phi^{-1})\} \quad (2.34)$$

com  $\eta(y_{ij}) = \eta(E(Y_{ij}|\mathbf{b}_i)) = \mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{b}_i$  para uma função de ligação  $\eta(\cdot)$ , com  $\mathbf{x}_{ij}^T$  e  $\mathbf{z}_{ij}^T$  vetores com dimensão  $p$  e  $q$ , respectivamente, das conhecidas covariáveis, com um vetor  $\boldsymbol{\beta}$   $p$ -dimensional dos coeficientes de regressão desconhecidos e um parâmetro de escala  $\phi$ . Para finalizar, tome  $f(\mathbf{b}_i|\mathbf{D})$  como a densidade da  $N(\mathbf{0}, \mathbf{D})$  para o vetor de efeitos aleatórios  $\mathbf{b}_i$ .



## Estimação dos parâmetros pelo Método MV

Neste modelo é necessário maximizar a função verossimilhança marginal, obtida pela integração em relação ao vetor de efeitos aleatórios  $\mathbf{b}_i$ . Assim, a contribuição do indivíduo  $i$  para a função de verossimilhança é

$$f_i(y_i|\boldsymbol{\beta}, \mathbf{D}, \phi^{-1}) = \int \prod_{j=1}^{n_i} f(y_{ij}|\theta_{ij}, \phi^{-1}, \mathbf{b}_i) f(\mathbf{b}_i|\mathbf{D}) d\mathbf{b}_i \quad (2.35)$$

para a construção da função de verossimilhança que é

$$\begin{aligned} L(\boldsymbol{\beta}, \mathbf{D}, \phi^{-1}) &= \prod_{i=1}^N f_i(y_i|\boldsymbol{\beta}, \mathbf{D}, \phi^{-1}) \\ &= \prod_{i=1}^N \int \prod_{j=1}^{n_i} f(y_{ij}|\theta_{ij}, \phi^{-1}, \mathbf{b}_i) f(\mathbf{b}_i|\mathbf{D}) d\mathbf{b}_i \end{aligned} \quad (2.36)$$

O grande problema deste tipo de estimação são as  $N$  integrais sobre os vetores de efeitos aleatórios  $q$ -dimensionais  $\mathbf{b}_i$ . O problema ainda é mais agravado se o tamanho da amostra for muito grande. Se métodos de estimação bayesiana forem utilizados, as integrais a serem resolvidas podem ser muito mais complicadas, tornando o processo computacional muito mais intensivo e complexo. Mais detalhes a respeito são descritos em Molenberghs e Verbeke (2005) e Verbeke e Molenberghs (2000).

## 2.5 Modelos combinados para dados com super-dispersão

Molenberghs, Verbeke, Demétrio e Vieira (2010) descreveram alguns modelos para tratar dados binários longitudinais com superdispersão. Este fenômeno ocorre quando há uma variabilidade maior que a predita pela relação média-variância (HINDE; DEMÉTRIO, 1998). Como foi explicado na seção relacionada aos Modelos Lineares Generalizados, quando a variável resposta tem como distribuição de probabilidade *Bernoulli*,  $\phi$  é igual a 1. No entanto, quando o valor estimado  $\hat{\phi}$  é maior do que 1, diz-se que há uma variação *extra-bernoulli*, logo, a variação existente no modelo não foi bem ajustada. Tal problema pode ser compreendido mais fácil se pensarmos no caso unidimensional, pois:

$$Var(Y) = \phi\pi(1 - \pi) = \phi\pi(1 - \pi) \quad (2.37)$$

sendo que  $\phi = 1$  no modelo Bernoulli.

Diante da questão apresentada, descreve-se uma alternativa para modelos com dados binários que estão sujeitos a superdispersão e estão sendo medidos longitudinalmente. Para este tipo de modelagem necessita-se de uma extensão particular do MLGM, classe esta mostrada em seções anteriores.

A alternativa proposta é o acréscimo de um efeito de aleatório para tratar superdispersão com distribuição *conjugada* a distribuição da variável de estudo. Tais distribuições recebem essa nomenclatura quando:

$$f(y|\gamma) = \exp \{ \phi^{-1} [yh(\gamma) - g(\gamma)] + s(y, \phi^{-1}) \} \quad (2.38)$$

$$f(\theta) = \exp \{ \omega [\psi h(\gamma) - g(\gamma)] + s(y, \psi^{-1}) \} \quad (2.39)$$

onde  $h(\gamma)$  e  $g(\gamma)$  são funções e  $\psi$ ,  $\phi$  e  $\omega$  são parâmetros, e  $s(y, \psi^{-1})$  e  $s(y, \phi^{-1})$  são constantes normalizadoras.

Definida esta questão, descreve-se uma relação de casos onde as distribuições são conjugadas, sendo que um deles é quando tem-se a distribuição da variável resposta  $Y$  com distribuição de probabilidade Bernoulli e o efeito aleatório  $\gamma$  com distribuição  $\text{Beta}(\alpha, \beta)$ , onde:

$$\begin{aligned} f(y) &= [\pi]^y [1 - \pi]^{1-y} \\ f(\gamma) &= \frac{\gamma^{\alpha_1-1} (1 - \gamma)^{\alpha_2-1}}{B(\alpha_1, \alpha_2)} \end{aligned} \quad (2.40)$$

Suponha que  $y_{ij}$  seja uma medida binária (sendo 0 caso não ocorra o evento e 1 caso contrário) do  $i$ -ésimo elemento no tempo  $j$ ;  $\gamma_{ij}$  sendo o parâmetro aleatório para a superdispersão na  $(ij)$ -ésima observação;  $\mathbf{b}_i$  o vetor aleatório de dimensão  $q$  associado ao vetor  $\mathbf{z}_{ij}$  com  $q$  fatores e/ou covariáveis;  $\boldsymbol{\beta}$  o vetor de parâmetros fixos de dimensão  $p$  associado ao vetor  $\mathbf{x}_{ij}$  com  $p$  fatores e/ou covariáveis;  $\pi_{ij}$  a probabilidade de sucesso para a  $(ij)$ -ésima medida,  $P(Y_{ij} = 1) = \pi_{ij}$ . Tendo definido esta questão, dois modelos combinados serão propostos: o modelo *logit-bernoulli-normal-beta* e o modelo *probit-bernoulli-normal-beta*.

### 2.5.1 Modelo combinado Logit-Bernoulli-Normal-Beta

Assumindo que a distribuição da variável resposta  $Y_{ij}$  é Bernoulli, que a distribuição para os efeitos aleatórios, distribuição esta que ira modelar a correlação entre

as medidas longitudinais,  $\mathbf{b}_i$  seja Normal e a distribuição para o parâmetro de superdispersão  $\gamma_{ij}$  seja beta, propôs-se um novo modelo assumindo a função de ligação logit, o modelo logit-bernoulli-normal-beta:

$$\begin{aligned} Y_{ij}|\mathbf{b}_i &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \gamma_{ij} \frac{\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)}{1 + \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)} \\ \gamma_{ij} &\sim \text{Beta}(\alpha_1, \alpha_2) \\ \mathbf{b}_i &\sim N_q(\mathbf{0}, \mathbf{D}) \end{aligned} \tag{2.41}$$

sendo que  $\alpha_1$  e  $\alpha_2$  são parâmetros da distribuição do efeito de superdispersão e a matriz  $\mathbf{D}$  é a matriz de variância-covariância referente ao efeito aleatório  $\mathbf{b}_i$ .

Como pode ser visto, o parâmetro de superdispersão foi agregado ao modelo com um efeito multiplicativo no predito linear da probabilidade de sucesso. Logo, a função de probabilidade condicional para  $Y_{ij}|\mathbf{b}_i, \gamma_{ij}$  é:

$$f(y_{ij}|\mathbf{b}_i, \gamma_{ij}) = \left[ \gamma_{ij} \frac{\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)}{1 + \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)} \right]^{y_{ij}} \left[ 1 - \gamma_{ij} \frac{\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)}{1 + \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)} \right]^{1-y_{ij}}$$

sendo que:

- o parâmetro  $\boldsymbol{\beta}$  é desconhecido e fixo;
- o parâmetro  $\mathbf{b}_i$  é aleatório, desconhecido e tem distribuição Normal com média 0 e com matriz de variância-covariância  $\mathbf{D}$ ;
- os parâmetros  $\alpha_1$  e  $\alpha_2$  são fixos e desconhecidos.

No entanto, assumiu-se a independência entre  $\gamma_{ij}$  e  $\mathbf{b}_i$ , e por isso tem-se que a contribuição da (ij)-ésima observação é:

$$f(y_{ij}, \gamma_{ij}, \mathbf{b}_i) = f(y_{ij}|\gamma_{ij}, \mathbf{b}_i) f(\mathbf{b}_i) f(\gamma_{ij}) \tag{2.42}$$

e por consequência:

$$f(\mathbf{y}_i, \boldsymbol{\gamma}_i, \mathbf{b}_i) = \prod_{j=1}^{n_i} f(y_{ij}|\gamma_{ij}, \mathbf{b}_i) f(\mathbf{b}_i) f(\gamma_{ij}). \tag{2.43}$$

Para a estimação dos parâmetros fixos  $\alpha_1$ ,  $\alpha_2$ ,  $\boldsymbol{\beta}$  e  $\mathbf{D}$  foi utilizado o método da máxima verossimilhança. Assim, para encontrar uma função de probabilidade que

dependesse somente da variável  $Y_{ij}$  foi necessário integrar a função de densidade conjunta em relação ao parâmetro de dispersão e ao vetor de parâmetros aleatórios  $\mathbf{b}_i$ . Fazendo os cálculos, sabendo que  $f(\mathbf{y}_i, \boldsymbol{\gamma}_i) = \int \prod_{j=1}^{n_i} f(y_{ij}|\gamma_{ij}, \mathbf{b}_i) f(\gamma_{ij}) f(\mathbf{b}_i) d\mathbf{b}_i$ , obteve-se a função de verossimilhança condicionada a  $\boldsymbol{\gamma}$  para os  $N$  indivíduos:

$$\begin{aligned} L(\boldsymbol{\beta}, \mathbf{D}, \alpha, \beta) &= \prod_{i=1}^N \int \prod_{j=1}^{n_i} \left[ \gamma_{ij} \frac{\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i)}{1 + \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i)} \right]^{y_{ij}} \\ &\times \left[ 1 - \gamma_{ij} \frac{\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i)}{1 + \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i)} \right]^{1-y_{ij}} \\ &\times \frac{1}{\sqrt{2\phi^{n_i}}} \frac{1}{\sqrt{|\mathbf{D}|}} \exp\left(-\frac{1}{2} \mathbf{b}_i^T \mathbf{D}^{-1} \mathbf{b}_i\right) \frac{\gamma_{ij}^{\alpha_1-1} (1-\gamma_{ij})^{\alpha_2-1}}{B(\alpha_1, \alpha_2)} d\mathbf{b}_i \end{aligned} \quad (2.44)$$

No entanto, essa função ainda está condicionada pelo efeito aleatório  $\boldsymbol{\gamma}$ . Porém, como  $Y_{ij}$  tem distribuição Bernoulli, a função  $f(y_{ij}|\mathbf{b}_i, \gamma_{ij})$  tem a seguinte relação:

$$f(y_{ij}|\mathbf{b}_i, \gamma_{ij}) = f(y_{ij} = 1|\mathbf{b}_i, \gamma_{ij})^{y_{ij}} f(y_{ij} = 0|\mathbf{b}_i, \gamma_{ij})^{1-y_{ij}} \quad (2.45)$$

Assim, definindo  $k_{ij} = [\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i) / (1 + \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i))]$ , tem-se, por definição, que  $f(y_{ij} = 1|\mathbf{b}_i) = \int_0^1 f(y_{ij} = 1|\mathbf{b}_i, \gamma_{ij}) f(\gamma_{ij}) d\gamma_{ij} = \alpha_1 k_{ij} (\alpha_1 + \alpha_2)^{-1}$  e, por outro lado,  $f(y_{ij} = 0|\mathbf{b}_i) = \int_0^1 f(y_{ij} = 1|\mathbf{b}_i, \gamma_{ij}) f(\gamma_{ij}) d\gamma_{ij} = \alpha_1 (1 - k_{ij}) (\alpha_1 + \alpha_2)^{-1}$ .

Logo, a função de verossimilhança, condicionada apenas pelo vetor de efeitos aleatórios  $\mathbf{b}_i$ , será seguinte da forma:

$$\begin{aligned} L(\boldsymbol{\beta}, \mathbf{D}, \alpha_1, \alpha_2) &= \prod_{i=1}^N \int \prod_{j=1}^{n_i} \left( \frac{\alpha_1 k_{ij}}{\alpha_1 + \alpha_2} \right)^{y_{ij}} \left[ \frac{\alpha_1 (1 + k_{ij}) + \alpha_2}{\alpha_1 + \alpha_2} \right]^{1-y_{ij}} f(\mathbf{b}_i) d\mathbf{b}_i \\ &= \prod_{i=1}^N \int \prod_{j=1}^{n_i} \frac{1}{\alpha_1 + \alpha_2} (\alpha k_{ij})^{y_{ij}} [\alpha_1 (1 + k_{ij}) + \alpha_2]^{1-y_{ij}} f(\mathbf{b}_i) d\mathbf{b}_i \end{aligned} \quad (2.46)$$

Os autores do artigo, não satisfeitos, construíram um modelo similar utilizando a função de ligação probit  $\Phi^{-1}(\cdot)$ , onde  $\Phi(x) = P(X < x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp[-\frac{t^2}{2}] dt$ .

## 2.5.2 Modelo combinado Probit-Bernoulli-Normal-Beta

O modelo probit-bernoulli-normal-beta é praticamente o mesmo, sendo que agora a função de ligação é a probit. Assim, o modelo é definido da seguinte forma:

$$\begin{aligned} Y_{ij}|\mathbf{b}_i &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \gamma_{ij}\Phi(\mathbf{x}_{ij}^T\boldsymbol{\beta} + \mathbf{z}_{ij}\mathbf{b}_i) \\ \gamma_{ij} &\sim \text{Beta}(\alpha_1, \alpha_2) \\ \mathbf{b}_i &\sim N_q(\mathbf{0}, \mathbf{D}) \end{aligned} \quad (2.47)$$

Para esse modelo, a função de probabilidade condicional  $Y_{ij}|\mathbf{b}_i, \gamma_{ij}$  é descrita como:

$$f(y_{ij}|\theta_{ij}, \mathbf{b}_i) = (\gamma_{ij}k_{ij})^{y_{ij}}(1 - \gamma_{ij}k_{ij})^{1-y_{ij}} \quad (2.48)$$

onde  $k_{ij} = \Phi(\mathbf{x}_{ij}^T\boldsymbol{\beta} + \mathbf{z}_{ij}\mathbf{b}_i)$  e  $\Phi(x) = P(X < x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}}\exp[-\frac{t^2}{2}]dt$

No entanto, se utilizar os mesmos resultados descritos anteriormente, os mesmos passos, chega-se que a função de densidade de  $Y_{ij}$  condicionada ao vetor efeitos aleatórios  $\mathbf{b}_i$  é da forma:

$$f(y_{ij}|\mathbf{b}_i) = \frac{1}{\alpha_1 + \alpha_2}(\alpha_1 k_{ij})^{y_{ij}} [\alpha_1(1 + k_{ij}) + \alpha_2]^{1-y_{ij}} \quad (2.49)$$

A função de verossimilhança condicionada ao vetor aleatório  $\mathbf{b}_i$  é igual a:

$$\begin{aligned} L(\boldsymbol{\beta}, \mathbf{D}, \alpha_1, \alpha_2) &= \prod_{i=1}^N \int \prod_{j=1}^{n_i} \left( \frac{\alpha_1 k_{ij}}{\alpha_1 + \alpha_2} \right)^{y_{ij}} \left[ \frac{\alpha_1(1 + k_{ij}) + \alpha_2}{\alpha_1 + \alpha_2} \right]^{1-y_{ij}} f(\mathbf{b}_i) d\mathbf{b}_i \\ &= \prod_{i=1}^N \int \prod_{j=1}^{n_i} \frac{1}{\alpha_1 + \alpha_2} (\alpha_1 k_{ij})^{y_{ij}} [\alpha_1(1 + k_{ij}) + \alpha_2]^{1-y_{ij}} f(\mathbf{b}_i) d\mathbf{b}_i \end{aligned} \quad (2.50)$$

Nota-se que a formula é praticamente a mesma que a apresentada para o modelo logit-binomial-normal-beta. As estimativas de  $\alpha_1$ ,  $\alpha_2$ ,  $\mathbf{D}$  e  $\boldsymbol{\beta}$ , assim como no modelo anterior podem ser encontradas através de técnicas de otimização e processos iterativos.

Todavia, o grande ponto do artigo está justamente nesse embróglio. Os autores conseguiram encontrar uma solução analíticas para as integrais envolvidas em (2.56), algo muito substancial e que trata justamente o problema das  $\sum_{i=1}^N n_i$  integrais envolvidas a cada iteração feita. Tal descoberta é ainda mais significativa quando se

está trabalhando com grande amostras. Para isto, a probabilidade conjunta de  $\mathbf{Y}_i$ ,  $\mathbf{b}_i$  e  $\gamma_i$  foi reescrita como:

$$f(\mathbf{y}_i, \gamma_i, \mathbf{b}_i) = \prod_{j=1}^{n_i} f(y_{ij} | \gamma_{ij}, \mathbf{b}_i) f(\mathbf{b}_i) \prod_{j=1}^{n_i} f(\gamma_{ij}). \quad (2.51)$$

Logo, a função de densidade marginal para  $\mathbf{Y}_i = 1$  pode ser escrita como

$$\begin{aligned} f(\mathbf{y}_i = \mathbf{1}) &= \int \cdots \int \prod_{j=1}^{n_i} f(y_{ij} = 1 | \gamma_{ij}, \mathbf{b}_i) f(\mathbf{b}_i) \prod_{j=1}^{n_i} f(\gamma_{ij}) d\mathbf{b}_i d\gamma_{ij} \\ &= \int \cdots \int \prod_{j=1}^{n_i} \gamma_{ij} \Phi(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i) f(\mathbf{b}_i) \prod_{j=1}^{n_i} f(\gamma_{ij}) d\mathbf{b}_i d\gamma_{ij} \\ &= \int \prod_{j=1}^{n_i} \Phi(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i) f(\mathbf{b}_i) d\mathbf{b}_i \int \prod_{j=1}^{n_i} \gamma_{ij} f(\gamma_{ij}) d\gamma_{ij} \\ &= \prod_{j=1}^{n_i} \int \Phi(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i) f(\mathbf{b}_i) d\mathbf{b}_i \prod_{j=1}^{n_i} \int \gamma_{ij} f(\gamma_{ij}) d\gamma_{ij} \end{aligned} \quad (2.52)$$

Na realidade, a segunda integral em (1.52) é um produtório de esperança da distribuição beta. Logo,

$$f(\mathbf{y}_i = \mathbf{1}) = \left( \frac{\alpha_1}{\alpha_1 + \alpha_2} \right)^{n_i} \prod_{j=1}^{n_i} \int \gamma_{ij} \Phi(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i) f(\mathbf{b}_i) d\mathbf{b}_i \quad (2.53)$$

Após isto, Molenberghs et al. (2010) fizeram vários cálculos analíticos e chegaram a uma solução simplificada deste modelo. Os cálculos feitos pelos autores estão no Apêndice D do artigo. A solução obtida foi que a probabilidade marginal para  $Y_i$  para o modelo probit-bernoulli-normal-beta pode ser escrita como

$$f(\mathbf{y}_i = \mathbf{1}) = \left( \frac{\alpha_1}{\alpha_1 + \alpha_2} \right)^{n_i} \Phi(\mathbf{X}_i \boldsymbol{\beta}; \mathbf{L}_i^{-1}), \quad \mathbf{L}_i = \mathbf{I} - \mathbf{Z}_i (\mathbf{D}^{-1} + \mathbf{Z}_i^T \mathbf{Z}_i)^{-1} \mathbf{Z}_i^T \quad (2.54)$$

Diante do referido resultado, as integrais que deveriam a ser calculadas desaparecem, sendo que o único termo que depende de integração é  $\Phi(\cdot)$ , porém ela é facilmente calculada em qualquer software básico de estatística. Isto torna o método muito menos intensivo e mais simples, necessitando somente encontrar os valores dos estimadores dos parâmetros que maximizam a função de verossimilhança.

É importante ainda ressaltar que, aberta essa janela para o modelo com a função de ligação probit, abre-se também uma oportunidade de aproximação ao modelo logit.

Segundo Zeger, Liang e Albert (1988):

$$\frac{e^y}{1 + e^y} \approx \Phi(\epsilon y) \quad (2.55)$$

onde  $\epsilon = (16\sqrt{3})/(15\pi)$ . E, por meio dessa aproximação, tem-se que:

$$\pi_{ij} \sim \gamma_{ij} \frac{\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)}{1 + \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)} \approx \gamma_{ij} \Phi_1(\epsilon(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij} \mathbf{b}_i)) \quad (2.56)$$

Aplicando (2.51) em (2.53) chega-se a:

$$f(\mathbf{y}_i = \mathbf{1}) \approx \left( \frac{\alpha_1}{\alpha_1 + \alpha_2} \right)^{n_i} \Phi_1(\epsilon \mathbf{X}_i \boldsymbol{\beta}; \tilde{\mathbf{L}}_i^{-1}), \quad \tilde{\mathbf{L}}_i = \mathbf{I} - \epsilon^2 \mathbf{Z}_i (\mathbf{D}^{-1} + \mathbf{Z}_i^T \mathbf{Z}_i)^{-1} \mathbf{Z}_i^T \quad (2.57)$$

Esta transformação é bastante válida, pois possibilita o cálculo de medidas como a razão de chances e a construção de seus intervalos de confiança.

## Capítulo 3

# Teoria de Resposta ao Item - TRI

Normalmente, quando se quer avaliar um grupo, indivíduo ou população, estes são submetidos há provas, testes, exames, diversos meios e métodos para avaliar o desempenho. Uma das formas de se avaliar é utilizando escore, como por exemplo, o número de acertos em um determinado teste (De AYALA, 2009). Em outros casos, podem-se ponderar tais acertos, atribuindo assim mais pontos a questões mais difíceis. Este é um cenário clássico da Teoria Clássica dos Testes (TCT), onde a modelagem é baseada no escore total que o indivíduo obtém nas provas.

A Teoria de Resposta ao Item - TRI surgiu na década de 1950 na forma de modelo ogiva-normal com um único parâmetro relacionado ao item (BAKER, 2004), e tem como principal foco não a prova em si, mas os itens, as suas dificuldades e quão bem eles discriminam os indivíduos de acordo com sua proficiência. A TRI permite com que provas, mesmos com questões diferentes e em tempos diferentes, tornem-se comparáveis mesmo apresentando itens com parâmetros de dificuldade diferentes. Além do Exame Nacional do Ensino Médio - ENEM, o Sistema Nacional de Avaliação da Educação Básica - SAEB, que foi o pioneiro do uso dessa teoria em 1995, e o Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo - SARESP são exemplos do uso da TRI nas avaliações educacionais brasileiras (ANDRADE; TAVARES; VALLE, 2000).

Diante disto, o modelo de resposta ao item será definido na seção seguinte e logo após será apresentado o modelo mais geral (modelo de 3 parâmetros). Depois, discutir-se-á a interpretação dos parâmetros através da Curva Característica do



Item - CCI, qual escala de habilidade será adotada e como é feita a estimativa dos parâmetros. Por fim, é definido o modelo de Rasch, caso particular do modelo de 3 parâmetros e que será adaptado nesta pesquisa.

### 3.1 Definição e pressupostos

Segundo Andrade et al.(2000), a Teoria de Resposta ao Item é definida por um conjunto de modelos matemáticos que tem a finalidade de modelar a probabilidade de acerto de um item  $j$  por um indivíduo  $i$  em função dos parâmetros dos item e da habilidade do indivíduo. Esta relação é baseada na idéia de quanto maior a habilidade do indivíduo, maior a probabilidade de acerto do item. De acordo com o autor, estes modelos dependem fundamentalmente de cinco fatores:

- da natureza do item: se eles são dicotômicos ou não;
- quantas populações estão sendo envolvidas, se é apenas uma ou são mais de uma população;
- a quantidade de traços latentes a serem medidos;
- do modelo ser linear ou não;
- do modelo ser cumulativo ou não-acumulativo (de desdobramento).

Pelo fato de ser somente uma demonstração do modelo e por este trabalho propor uma nova abordagem ao modelo de Rasch, serão somente apresentados modelos para itens *dicotômicos*. Nestes casos, os itens podem ser de múltipla escolha ou até mesmo com respostas abertas, sendo que exista uma definição de “certo” e “errado” na correção. Tais itens são chamados de *dicotomizados*.

Outra suposição a ser feita é que somente uma *única população* está sendo envolvida e com apenas um *único* traço latente a ser medido, sendo essa classe de modelos chamada de modelos *unidimensionais*. Em situações reais é muito difícil assumir que somente um traço latente esteja influenciando na probabilidade de acerto desse indivíduo. No entanto, supondo que haja um traço dominante, essa suposição é considerada como satisfeita. Supõe-se ainda que as respostas, condicionadas a uma

habilidade  $\theta_i$  fixa, são independentes. Esta suposição é chamada de *independência local* e é de extrema importância para que o modelo possa ser ajustado.

Em tal classe de modelos, conhecidos como modelos de 1, 2 e 3 parâmetros, que consideram, respectivamente:

- o parâmetro de dificuldade do item;
- o parâmetro de dificuldade e o parâmetro de discriminação;
- o parâmetro de dificuldade, o parâmetro de discriminação e um parâmetro de acerto ao acaso.

A fim de introduzir o modelo, será apresentado o modelo mais completo, o modelo com 3 parâmetros. Cabe ressaltar que na seção seguinte os parâmetros  $b$ 's e  $\theta$  possuem significados diferentes dos apresentados nas seções anteriores. Isto se deve ao fato de que procurou-se ser mais fiel possível a parametrização usual dos modelos de TRI apresentada pela literatura especializada, logo, esta notação foi mantida.

## 3.2 Modelo de TRI com 3 parâmetros

Nesta classe, os dois tipos de modelos mais utilizados são:

- os modelos que utilizam a função de ligação **logit**:

$$Logit(x) = \frac{\exp(x)}{1 + \exp(x)}$$

- os modelos que utilizam a função de ligação **probit**, ou **Ogiva Normal**:

$$Probit(x) = \Phi(x) = P(X < x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{t^2}{2}\right] dt$$

Normalmente, utiliza-se a ligação logit por ser mais simples que a transformação probit, que envolve uma integração não trivial. A partir daí, o modelo de TRI com 3 parâmetros e função de ligação logit é definido por Andrade et al.(2000) da seguinte maneira:

$$P(U_{ij} = 1|\theta_i) = c_j + (1 - c_j) \frac{\exp\{D[a_j(\theta_i - b_j)]\}}{1 + \exp\{D[a_j(\theta_i - b_j)]\}} \quad (3.1)$$

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

- $U_{ij}$  é uma variável dicotômica que assume valor '1' caso o indivíduo  $i$  acerte o item  $j$ , e assume o valor '0' caso contrário;
- $\theta_i$  é a habilidade (traço latente) do indivíduo  $i$ ;
- $P(U_{ij} = 1|\theta_i)$  é a probabilidade do indivíduo  $i$  responder o item  $j$ , dado que ele tenha a habilidade  $\theta_i$ . Tal função é chamada de Função de Resposta ao Item (FRI);
- $b_j$  é o parâmetro de dificuldade do item  $j$ , medido na mesma escala das habilidades dos indivíduos;
- $a_j$  é o parâmetro de discriminação do item  $j$ , com valor proporcional a inclinação da Curva de Característica do Item - CCI, no ponto  $b_j$ ;
- $c_j$  é o parâmetro que representa a probabilidade de indivíduos com habilidade muito baixa acertar o item  $j$ , comumente chamado de parâmetro de acerto ao acaso;
- $D$  é um fator constante e igual a 1. Quando deseja-se aproximar os valores fornecidos pela função ligação logit para os dados fornecidos pela probit, esse fator  $D$  é igual a 1,7, gerando estimativas dos parâmetros dos itens diferentes.

### 3.3 Interpretação dos parâmetro e o gráfico CCI

É importante notar que a função  $P(U_{ij} = 1|\theta_i)$  varia de acordo com a habilidade do indivíduo  $i$ , tendo os outros parâmetros fixos. A ideia por trás desta função é que a probabilidade de acerto do item  $j$  aumente de acordo com o grau de habilidade do indivíduo  $i$ . No entanto, a representação gráfica que mostra como é esse comportamento da função de probabilidade no item  $j$  é chamada Curva Característica do Item - CCI, curva esta que é exemplificada abaixo:

Nota-se que a curva tem um formato de sigmoidal, tendo a sua inclinação e deslocamento na escala de habilidades definidas pelos parâmetros do item. A escala do parâmetro de habilidade é *arbitrária*, pois não há importância na comparação de magnitude dessas habilidades, mas na existência de ordem, de um rank. A escala

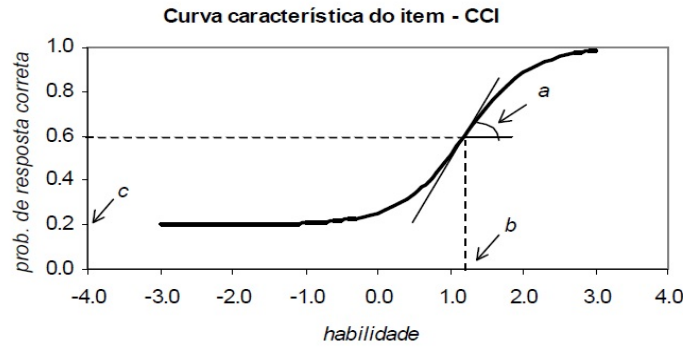


Figura 3.1: Curva Característica do Item - CCI (ANDRADE et al.,2002)

do parâmetro de dificuldade  $b$  segue a mesma escala da habilidade, enquanto que o parâmetro  $c$  varia entre 0 e 1, pois trata-se de uma probabilidade de acerto ao acaso.

Se olharmos com mais atenção para o modelo, nota-se que o parâmetro  $b$  representa a habilidade que o indivíduo  $i$  necessita para ter a probabilidade de acerto igual a  $\frac{(1+c)}{2}$ . Quando  $c$  é igual a zero, quando não considera-se o acerto ao acaso no modelo, o parâmetro  $b$  simboliza o valor da habilidade que o indivíduo necessita para ter a probabilidade de acerto igual a de erro. Caso o indivíduo tenha uma habilidade maior que  $b$ , é como se ele vencesse o item, conseguisse ter mais chance de êxito do que fracasso. Diante disso, conclui-se que quanto maior o valor de  $b$ , mais difícil é o item.

Por outro lado, o parâmetro  $a$  é responsável pelo grau de inclinação da CCI. Quanto maior o valor de  $a$ , maior a capacidade do item de separar os indivíduos com habilidade alta dos indivíduos com baixa habilidade na região em torno do parâmetros de dificuldade  $b$ . É importante notar que não há a possibilidade do  $a$  ser negativo, pois supõe-se a probabilidade de acerto do item aumente quando maior for a habilidade do indivíduo.

### 3.4 Identificabilidade

Um problema muito comum em TRI é o problema da *falta de identificabilidade*. Segunda Andrade et al. (2000), esta não-identificabilidade ocorre quando mais de um conjunto de parâmetros produz o mesmo valor no modelo de resposta ao item e, consequentemente, na verossimilhança. Para o modelo de 3 parâmetros com ligação

logit, os autores citam o exemplo onde  $\theta_i^* = \zeta\theta_i + \varrho$ ,  $b_j^* = \zeta b_j + \varrho$ ,  $a_j^* = \frac{a_j}{\zeta}$  e  $c_j^* = c_j$ , onde  $\zeta$  e  $\varrho$  são constantes reais com  $\zeta > 0$ . Logo, tem-se que:

$$\begin{aligned}
P(U_{ij} = 1 | \theta_i^*, a_j^*, b_j^*, c_j^*) &= c_j^* + (1 - c_j^*) \{1 + \exp\{-D[a_j^*(\theta_i - b_j^*)]\}\}^{-1} \\
&= c_j + (1 - c_j) \{1 + \exp\{-D\left[\frac{a_j}{\zeta}(\zeta\theta_i + \varrho - (\zeta b_j + \varrho))\right]\}\}^{-1} \\
&= c_j + (1 - c_j) \{1 + \exp\{-D[a_j(\theta_i - b_j)]\}\}^{-1} \\
&= P(U_{ij} = 1 | \theta_i, a_j, b_j, c_j)
\end{aligned} \tag{3.2}$$

Uma maneira de tratar esta identificabilidade é definindo uma métrica (unidade de medida) para as habilidades e, conseqüentemente, para os parâmetros dos itens.

### 3.5 Escala de habilidade

Na grande maioria dos testes, a análise do desempenho de um indivíduo é feita pelo escore que ele obteve em  $J$  questões, sendo as respostas classificadas como “certa” ou “errada”. No entanto, a TRI vai além e é capaz de mensurar o nível de habilidade desse indivíduo, e essa habilidade pode, teoricamente, assumir qualquer valor entre  $-\infty$  e  $+\infty$ . Assim, é necessário definir uma escala padrão que consiga representar o valor médio e a variabilidade das habilidades da população em estudo. O mais comum é tomar a média igual a “0” e o desvio igual a “1”, porém isto não é obrigatório. Em muitos casos, até para o efeito de entendimento e comparação, utilizam-se outros valores para estes parâmetros, pois talvez não seja tão trivial imaginar que um indivíduo tenha uma habilidade negativa, por exemplo. O mais interessante é que essa mudança não compromete as estimativas da probabilidade de acerto, estas permanecem iguais, independentes da escala adotada.

### 3.6 Estimação dos Parâmetros

Para ajustar o modelo é necessário que se estime os parâmetros dos modelos. Conforme Andrade et al.(2000), as estimativas para os modelos de resposta ao item podem ser feitas de três maneiras:

- estima-se os  $a_j's$ ,  $b_j's$  e  $c_j's$  fixando os valores das habilidades dos indivíduos;

- estima-se os  $\theta'_i$ s fixando os valores dos parâmetros dos modelos;
- estima-se simultaneamente todos estes parâmetros.

Caso o método de estimação seja o método da Máxima Verossimilhança (MV), este se baseará, independente da forma que for feita, na função de verossimilhança, que será da seguinte forma:

$$L(\boldsymbol{\zeta}, \boldsymbol{\theta}) = \prod_{i=1}^N \prod_{j=1}^J P(U_{ij} = u_{ij} | \theta_i, \boldsymbol{\zeta}_j) \quad (3.3)$$

onde  $\boldsymbol{\zeta}_j$  é o vetor  $(a_j, b_j, c_j)^T$ . Assim, procura-se os valores de  $\boldsymbol{\zeta}$  e  $\boldsymbol{\theta}$  que maximizem essa função. Mais detalhes sobre a estimação desses parâmetros estão em De Ayala (2009), Baker (2004) e Andrade et al. (2000). No entanto, pode-se adiantar que essas estimativas via MV são computacionalmente muito intensivas, principalmente se o tamanho da amostra for muito grande, pois além das habilidades, terá que se estimar o vetor  $\boldsymbol{\zeta}$  de parâmetros e considerando-os fixos. Considerado o traço latente como efeito aleatório, provenientes de uma função de probabilidade conhecida, simplifica a parametrização do modelo, gerando estimadores consistentes, porém demandando de outros algoritmos computacionais associados ao método de estimação por Máxima Verossimilhança Marginal (ANDRADE et al., 2000).

### 3.7 Modelo de Rasch

Este modelo foi primeiramente proposto por Rasch, utilizando a função de ligação Ogiva Normal (Probit). Posteriormente, este modelo foi definido utilizando a função de ligação Logit por ser computacionalmente mais simples que a Ogiva Normal.

$$P(U_{ij} = 1 | \theta_i) = \frac{\exp\{\theta_i - b_j\}}{1 + \exp\{\theta_i - b_j\}} \quad (3.4)$$

sendo que para o modelo de Rasch  $\alpha = 1$ .

O parâmetro  $b_j$  é o ponto de inflexão da CCI e é o valor do traço latente que o indivíduo  $i$  precisa para ter iguais as probabilidades de acerto e de erro do item.

Mesmo ele sendo muito parecido com o modelo de resposta ao item de 1 parâmetro, suas filosofias são diferentes, segundo De Ayala (2009). Segundo ele, “o modelo de 1 parâmetro ajusta os dados da melhor forma possível, dada as restrições do modelo.

Em contrapartida, o modelo de Rasch é utilizado para “construir” a variável de interesse. O autor ainda explica que o modelo de Rasch é o modelo padrão para que alguém possa criar um instrumento para medidas com traços latentes unidimensionais.

### **3.8 Funcionamento Diferencial do Item - DIF**

Segundo Lord (1980), citado por Karino et al.(2012), “um item possui DIF quando a Curva Característica do Item (CCI) difere para dois ou mais grupos, considerando o nível da variável latente”. Além de comparação das CCT's, pode-se analisar a existência de DIF por meio da comparação dos parâmetros de dificuldade dos itens. Esta técnica foi proposta por Lord e define que um item apresentará DIF se o valor do parâmetro  $b$  estimado para uma população possuir diferenças significativas com o valor estimado em outra população. (KARINO et al, 2012).

## Capítulo 4

# Modelo de Rasch Estendido para Heterogeneidade Atribuída à Fonte Desconhecida

### 4.1 Definição do modelo

Suponha um teste composto por  $J$  itens aplicado a  $N$  pessoas e que as suas respostas sejam independentes. Estas respostas, definidas como  $Y_{ij}$  sendo  $i$  referente ao  $i$ -ésimo respondente e  $j$  ao  $j$ -ésimo item, serão classificadas como erradas ou certas, “0” ou “1”, respectivamente. Assume-se o modelo probabilístico *Bernoulli* para a distribuição dessas medidas binárias, a distribuição Beta para o efeito aleatório que irá acomodar a superdispersão e a distribuição Normal para modelar a habilidade do respondente. Com base nos resultados de Molenberghs, Verbeke, Demétrio e Vieira (2010), tem-se o modelo de Rasch estendido para acomodar o efeito de heterogeneidade atribuída a fonte desconhecida (ou superdispersão).

Considerando  $Y_{ij}$  a resposta de cada respondente do teste, sendo “0” ou “1”, tem-se



que

$$\begin{aligned}
f_{Y_{ij}|\Theta_i}(y_{ij}|\theta_i, \gamma_i) &= [\pi_{ij}]^{y_{ij}} [1 - \pi_{ij}]^{1-y_{ij}} \\
\pi_{ij} &= \gamma_i \kappa(\theta_i - b_j) \\
f(\gamma_i) &\sim \text{Beta}(\alpha_1, \alpha_2) \\
\theta_i &\sim N(0, 1)
\end{aligned} \tag{4.1}$$

sendo  $\theta_i$  o nível da habilidade do  $i$ -ésimo respondente (traço latente);  $b_j$  o nível de dificuldade do  $j$ -ésimo item;  $\gamma_i$  o traço latente que irá acomodar a heterogeneidade atribuída à fonte desconhecida (ou superdispersão);  $\kappa(\cdot)$  uma função de ligação (adotada no presente trabalho como sendo a função Logit ou Probit) e  $\pi_{ij}$  a probabilidade do  $i$ -ésimo indivíduo acertar o  $j$ -ésimo item.

De acordo com a definição descrita, a densidade  $Y_{ij}$  condicionada aos efeitos aleatórios  $\theta_i$  e  $\gamma_i$  é:

$$f(y_{ij}|\theta_i, \gamma_i) = [\gamma_i c(\theta_i - b_j)]^{y_{ij}} [1 - \gamma_i \kappa(\theta_i - b_j)]^{1-y_{ij}} \tag{4.2}$$

Logo, a densidade conjunta das variáveis aleatória é definida por:

$$f_{Y_{ij}, \Theta_i, \Gamma_i}(y_{ij}, \theta_i, \gamma_i) = f_{Y_{ij}|\Theta_i, \Gamma_i}(y_{ij}|\theta_i, \gamma_i) f_{\Theta_i}(\theta_i) f_{\Gamma_i}(\gamma_i) \tag{4.3}$$

Assim, a contribuição da  $i$ -ésimo respondente, assumindo independência condicional, é:

$$f(\mathbf{y}_i, \theta_i, \gamma_i) = \prod_{j=1}^J f_{Y_{ij}|\Theta_i, \Gamma_i}(y_{ij}|\theta_i, \gamma_i) f_{\Theta_i}(\theta_i) f_{\Gamma_i}(\gamma_i) \tag{4.4}$$

Neste caso, como foi feito na formulação do modelo de Molenberghs et al.(2010), o método da estimação adotado é o da Máxima Verossimilhança. Para poder estimar o vetor de parâmetros de dificuldade dos itens  $\boldsymbol{\zeta} = (b_1, \dots, b_n)^T$ ,  $\alpha_1$  e  $\alpha_2$  é necessário definir a função de verossimilhança conjunta.

$$L(\boldsymbol{\zeta}, \alpha_1, \alpha_2 | \boldsymbol{\Theta}, \boldsymbol{\Gamma}; \mathbf{y}) = \prod_{i=1}^N \prod_{j=1}^J f_{Y_{ij}|\Theta_i, \Gamma_i}(y_{ij}|\theta_i, \gamma_i) f_{\Theta_i}(\theta_i) f_{\Gamma_i}(\gamma_i). \tag{4.5}$$

Desta forma, a função de verossimilhança marginal é definida por:

$$L(\boldsymbol{\zeta}, \alpha_1, \alpha_2; \mathbf{y}) = \prod_{i=1}^N \int \dots \int \prod_{j=1}^J f_{Y_{ij}|\Theta_i, \Gamma_i}(y_{ij}|\theta_i, \gamma_i) f_{\Theta_i}(\theta_i) f_{\Gamma_i}(\gamma_i) d\gamma_i d\theta_i \quad (4.6)$$

$$= \prod_{i=1}^N \int \dots \int \prod_{j=1}^n [\gamma_i \kappa(\theta_i - b_j)]^{y_{ij}} [1 - \gamma_i \kappa(\theta_i - b_j)]^{1-y_{ij}} \quad (4.7)$$

$$\times (\sqrt{2\pi})^{-1} e^{-\frac{\theta_i^2}{2}} \frac{\gamma_i^{\alpha_1-1} (1-\gamma_i)^{\alpha_2-1}}{B(\alpha_1, \alpha_2)} d\gamma_i d\theta_i.$$

De Molenberghs et al. (2010) tem-se como resultado a função de Máxima Verossimilhança Marginal integrada em  $\boldsymbol{\gamma}$ :

$$L(\boldsymbol{\zeta}, \alpha_1, \alpha_2; \mathbf{y}) = \prod_{i=1}^N \int \dots \int \prod_{j=1}^n \frac{1}{\alpha_1 + \alpha_2} [\alpha_1 \kappa(\theta_i - b_j)]^{y_{ij}} [(1 - \kappa(\theta_i - b_j)) \alpha_1 + \alpha_2]^{1-y_{ij}}$$

$$\times (\sqrt{2\pi})^{-1} e^{-\frac{\theta_i^2}{2}} d\theta_i$$

No entanto, quando a função  $\kappa(\cdot)$  é a probit, o modelo entra no grupo descrito em Molenberghs et al. (2010), onde:

$$f(\mathbf{Y}_i = \mathbf{1}) = \Phi(\mathbf{X}_i \boldsymbol{\beta}; L_i^{-1}), \quad L_i = \mathbf{I} - \mathbf{Z}_i (\mathbf{D}^{-1} + \mathbf{Z}_i^T \mathbf{Z}_i)^{-1} \mathbf{Z}_i^T \quad (4.8)$$

Para o modelo de Rasch estendido tem-se que  $\mathbf{X}_i = -\mathbf{I}$  e  $\boldsymbol{\beta} = \boldsymbol{\zeta}$  e, por consequência,  $\mathbf{Z}_i = \mathbf{I}$ ,  $\mathbf{D} = \mathbf{I}$  e  $L_i = \frac{1}{2} \mathbf{I}$ . Adaptando o modelo, chega-se a:

$$f(\mathbf{Y}_i = \mathbf{1}) = \left( \frac{\alpha_1}{\alpha_1 + \alpha_2} \right)^J \prod_{j=1}^J \Phi(-b_j; 2). \quad (4.9)$$

Por meio do referido resultado, os parâmetros de dificuldade  $\boldsymbol{\zeta}$  e os parâmetros  $\alpha_1$  e  $\alpha_2$  poderão ser estimadas ao maximizar essa função utilizando um dos métodos anteriormente apresentados.

## 4.2 Uma nova parametrização da distribuição Beta

No intervalo de estudos para a formulação do novo modelo foi utilizada uma nova parametrização para a distribuição Beta. Esta reparametrização foi proposta por Ferrari & Cribari-Neto (2004) e consiste na reparametrização da densidade de probabilidade Beta de forma que ela dependa de um parâmetro referente a *média* e

outro referente a *dispersão* dos dados. Dada uma variável aleatória  $X$  com distribuição de probabilidade Beta reparametrizada, a densidade de  $X$  é escrita da seguinte forma:

$$f_X(x; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} x^{\mu\phi-1} (1-x)^{(1-\mu)\phi-1}, \text{ para } 0 \leq x \leq 1$$

onde  $0 < \mu < 1$  e  $\phi > 0$ .

Para obter tal estrutura os autores adotaram a parametrização  $\mu = \frac{\alpha_1}{\alpha_1 + \alpha_2}$  e  $\phi = \alpha_1 + \alpha_2$ , isto é,  $\alpha_1 = \mu\phi$  e  $\alpha_2 = (1-\mu)\phi$ . Desta forma, as expressões da esperança e variância são:

$$\begin{aligned} E(X) &= \mu \\ Var(X) &= \frac{V(\mu)}{1 + \phi} \end{aligned}$$

onde  $V(\mu) = \mu(1-\mu)$ . Assim, nota-se que o parâmetro  $\mu$  é a média de  $X$  e  $\phi$  pode ser considerado um parâmetro de precisão pois, fixada a média, quanto maior for o seu valor, menor será a variância de  $X$ .

A partir desta nova parametrização pode-se redefinir o modelo de Rasch Estendido. Neste caso, se a distribuição de probabilidade do parâmetro aleatório  $\gamma$  for definida por essa reparametrização, a densidade de  $Y_{ij}$  condicionada a  $\theta_i$ :

$$f(y_{ij}|\theta_i) = \frac{1}{\alpha_1 + \alpha_2} [\alpha_1 \kappa(\theta_i - b_j)]^{y_{ij}} [(1 - \kappa(\theta_i - b_j)) \alpha_1 + \alpha_2]^{1-y_{ij}} \quad (4.10)$$

pode ser reparametrizada como:

$$\begin{aligned} f(y_{ij}|\theta_i) &= \frac{1}{\mu} [\phi \mu \kappa(\theta_i - b_j)]^{y_{ij}} \cdot \{[1 - \kappa(\theta_i - b_j)] \mu \phi + (1 - \mu) \phi\}^{1-y_{ij}} \\ &= \phi^{y_{ij}-1} [\mu \kappa(\theta_i - b_j)]^{y_{ij}} \cdot \{\phi [(1 - \mu) + (\mu (1 - \kappa(\theta_i - b_j)))]\}^{1-y_{ij}} \\ &= \phi^{y_{ij}-1} [\mu \kappa(\theta_i - b_j)]^{y_{ij}} \cdot \{\phi [1 - \mu + (\mu (1 - \kappa(\theta_i - b_j)))]\}^{1-y_{ij}} \\ &= \phi^{y_{ij}-1} [\mu \kappa(\theta_i - b_j)]^{y_{ij}} \cdot \{\phi [1 - \mu (\kappa(\theta_i - b_j))]\}^{1-y_{ij}} \\ &= \phi^{y_{ij}-1} \phi^{1-y_{ij}} [\mu \kappa(\theta_i - b_j)]^{y_{ij}} \cdot \{1 - \mu \kappa(\theta_i - b_j)\}^{1-y_{ij}} \\ \therefore f(y_{ij}|\theta_i) &= \mu^{y_{ij}} [\kappa(\theta_i - b_j)]^{y_{ij}} \cdot [1 - \mu \kappa(\theta_i - b_j)]^{1-y_{ij}} \end{aligned}$$

É importante notar que neste passo o parâmetro  $\phi$  do modelo desaparece da densidade condicional de  $Y_{ij}$ , simplificando o ajuste. Deste modo a função de Máxima Verossimilhança Marginal reparametrizada será:

$$\begin{aligned} L(\zeta, \mu, \phi|\mathbf{y}) &= \prod_{i=1}^N \int \cdots \int \prod_{j=1}^J \mu^{y_{ij}} [\kappa(\theta_i - b_j)]^{y_{ij}} [1 - \mu \kappa(\theta_i - b_j)]^{1-y_{ij}} \quad (4.11) \\ &\times (\sqrt{2\pi})^{-1} e^{-\frac{\theta_i^2}{2}} d\theta_i \end{aligned}$$

Do mesmo modo, se  $\kappa(\cdot)$  for a função de ligação probit, a forma analítica encontrada por Molenberghs et al. (2010) pode ser utilizada, chegando a um novo resultado para a distribuição de probabilidade marginal do indivíduo  $i$

$$f(\mathbf{Y}_i = \mathbf{1}) = (\mu)^J \prod_{j=1}^J \Phi(-b_j; 2). \quad (4.12)$$

Logo, todos estes modelos apresentados neste capítulo são, na realidade, modelos não-lineares (ou não-lineares mistos). Como todo modelo não-linear, dificuldades na obtenção de valores iniciais para iniciar o algoritmo, dificuldades na convergência de superfícies de log-verossimilhanças mal comportadas, correlação entre parâmetros estimados, etc, são parte da atividade de modelagem. Durante o trabalho de ajuste destes modelos estes problemas foram encontrados e serão melhor discutidos no Anexo C deste trabalho.

### 4.3 Simulação de Monte Carlo para o modelo de Rasch Estendido

Após a definição do modelo de Rasch estendido definido em (4.11) e do modelo de Rasch Estendido analítico definido em (4.12) foram realizadas simulações via método de Monte Carlo a fim de avaliar o ajuste do modelo a dados com efeitos aleatórios  $\gamma$  previamente estabelecidos. Além disto, desejou-se ajustar o modelo de Rasch Tradicional apresentado em (3.3) para compará-lo com os modelos propostos acima. Para isto, o modelo de Rasch foi reescrito como um MLGM e definido da seguinte maneira:

$$\begin{aligned} f_{Y_{ij}|\Theta_i}(y_{ij}|\theta_i) &= [\pi_{ij}]^{y_{ij}} [1 - \pi_{ij}]^{1-y_{ij}} \\ \pi_{ij} &= \kappa(\theta_i - b_j) \\ \theta_i &\sim N(0, 1) \end{aligned} \quad (4.13)$$

sendo que a função  $\kappa(\cdot)$  é uma função de ligação a ser definida. Para este estudo, a função  $\kappa(\cdot)$  escolhida foi a probit devido ao resultado analítico encontrado em (4.12) necessitar desta suposição.

Para a geração dos dados utilizou-se a estrutura do modelo de Rasch estendido definida em (4.1), sendo que neste caso o efeito aleatório  $\gamma$  segue a distribuição beta

reparametrizada de Ferrari & Cribari-Neto(2004). Como foi descrito anteriormente, a distribuição de probabilidade dos  $\gamma$ 's dos indivíduos depende dos parâmetros  $\mu$  e  $\phi$ . Logo, observou-se a necessidade de testar como os modelos se comportam quando a média do efeito aleatório  $\mu$  são fixos em diversos valores, criando vários cenários. Assim, fixou o valor de  $\mu$  em 0.7, 0.75, 0.8, 0.85, 0.9 e 0.95. Diante disto, era necessário fixar o valor do parâmetro de dispersão  $\phi$  da distribuição de  $\gamma$ .

A definição de um critério de escolha do valor de  $\phi$  é importante devido a sensibilidade da distribuição beta em relação a mudança deste valor e a variação da forma da sua distribuição de probabilidade, como pode ser visto abaixo.

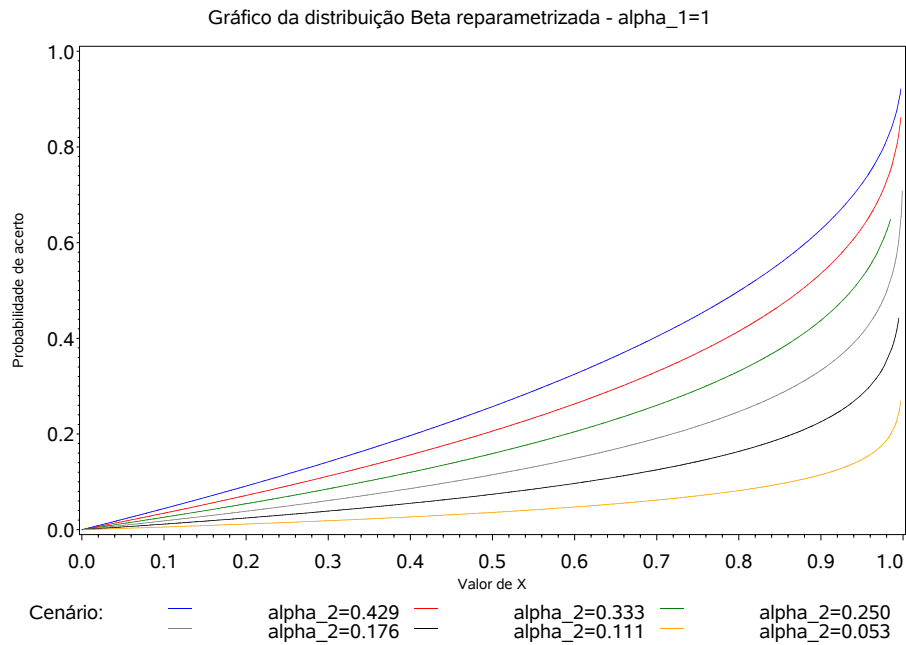


Figura 4.1: Gráfico da distribuição beta com  $\alpha_1$  fixo em “1” e  $\alpha_2$  variável

Logo, sabendo que  $\mu = \frac{\alpha_1}{\alpha_1 + \alpha_2}$  e  $\phi = \alpha_1 + \alpha_2$ , fixou-se o valor de  $\alpha_1$  em ‘1’ e, utilizando a relação  $\alpha_1 = \mu\phi$ , chegou-se aos seguinte valores de  $\phi$  para cada valor de  $\mu$  preestabelecido, apresentados na tabela 4.1.

Definidos os valores de  $\mu$  e  $\phi$  para cada um dos 6 cenários, gerou-se 100 amostras com as respostas de 20000 indivíduos para 14 itens. A tabela com os parâmetros de dificuldade dos itens segue no Anexo B.

O passo seguinte foi ajustar os modelos de Rasch estendido aos bancos de dados de cada cenário utilizando a função de Máxima Verossimilhança Marginal (MVM) descrita em (4.11) para a função probit (REMP), o modelo de Rasch estendido definido

Tabela 4.1: Cenários gerados

$\alpha_1$	$\mu$	$\alpha_2$	$\phi$
1	0,70	0,429	1,429
1	0,75	0,333	1,333
1	0,80	0,250	1,250
1	0,85	0,176	1,176
1	0,90	0,111	1,111
1	0,95	0,053	1,053

em (4.12)(REMol) e o modelo de Rasch tradicional.

Para a realizar a ajuste dos modelos foi utilizando o sistema SAS e o procedimento NLMIXED. Para a utilização deste procedimento é necessária a descrição do núcleo de sua função de máxima verossimilhança. Para o modelo REMP a função de MVM é acima apresentada e, por consequência, o núcleo também. Para o caso do modelo de Rasch tradicional o procedimento possui uma opção que se adequa as respostas *bernoulli* de cada respondente. No entanto, para o modelo REMol a função de Máxima Verossimilhança (MV) é determinada pela seguinte expressão:

$$L(\mu, \zeta | \mathbf{Y}) = \prod_{i=1}^N \prod_{j=1}^J [\mu \Phi(-b_j; 2)]^{y_{ij}} [1 - \mu \Phi(-b_j; 2)]^{1-y_{ij}} \quad (4.14)$$

Tendo definido isto, os modelos foram ajustados aos dados dos 6 cenários definidos e as médias das estimativas geradas pelo modelo seguem no Anexo B do trabalho.

Ao analisar as médias das estimativas apresentadas nos 6 cenários nota-se que o modelo REMol é o que tem as estimativas dos parâmetros de dificuldade dos itens populacionais e da médias do efeito aleatório  $\gamma$  mais consistentes. Por outro lado, as estimativas do modelo de Rasch tradicional apresentam grandes vieses em todos os cenários, tanto nas estimativas pontuais quanto nos respectivos erros-padrões. Este fato indica que o modelo de Rasch tradicional não consegue capturar e nem modelar a existência de heterogeneidade de fonte desconhecida, gerando estimativas inconsistentes e comprometendo a realização da inferência. Um ponto interessante a ser destacado é a diferença entre as estimativas do modelo REMol e as estimativas do modelo REMP. Mesmo o modelo REMol sendo resultante de uma solução analítica para o modelo REMP, as estimativas do REMP ficaram mais próximas das estimati-

vas do modelo de Rasch tradicional. Logo, este comportamento inesperado do modelo REMP sugere um estudo mais detalhado de suas propriedades.

Em relação à média dos erros-padrões estimados observa-se que para o modelo REMol estas médias são consideravelmente maiores que as dos outros modelos. Esta grandeza aumenta principalmente para os itens mais fáceis, aqueles com menores parâmetros de dificuldade populacionais. Uma possível justificativa deste comportamento das estimativas destes erros-padrões pode estar no fato de que itens muito fáceis são cada vez mais acertados, fazendo com que o vetor correspondente as respostas de um  $i$ -ésimo indivíduo aos 14 itens seja um vetor de 1's. Este fato pode estar comprometendo a estabilidade das estimativas e, quanto menor for a dificuldade do item, mais isto se torna provável de acontecer.

Tabela 4.2: Número de falhas por cenário

Cenário	Número de falhas
$\mu=0.7$	29
$\mu=0.75$	34
$\mu=0.8$	55
$\mu=0.85$	39
$\mu=0.9$	56
$\mu=0.95$	36

Uma outra vantagem do modelo REMol é que o ajuste dele é mais rápido em comparação aos demais. Em média, uma ajuste do modelo REMol demora em torno de 38s, contra 48s e 1min dos modelos de Rasch tradicional e REMP, respectivamente, utilizando com computador com procedor Intel i3 da segunda geração, com velocidade de processamento de 3.10 GHz. Entretanto, o modelo REMol apresentou falhas na convergência, como é apresentado na tabela acima.

Uma justificativa é que necessita-se estipular um valor inicial bastante razoável para a média  $\mu$  dos efeitos aleatórios  $\gamma$ . Esta escolha é determinante pois o modelo REMol é um pouco mais instável numericamente que os demais modelos. Logo, uma vez escolhido um valor inicial, ele foi empregado nos 100 ajustes do modelo aos 100 bancos de dados gerados para cada um dos cenários. Portanto, todos os ajustes tiveram os mesmos valores iniciais e uma possível alternativa seria utilizar as

estimativas do modelo de Rasch tradicional como valores iniciais.

Por fim, diante todos estes pontos, nota-se que o modelo REMol consegue ajustar muito bem os dados simulados, gerando estimativas pontuais precisas e conseguindo estimar muito bem a média do efeito aleatório que captura a superdispersão,  $\gamma$ , em todos os cenários, além de ser mais rápido que os demais. Por outro lado, o modelo de Rasch Tradicional tem um desempenho muito ruim quando a perturbação é acrescida aos dados, gerando estimativas inconsistentes.



## Capítulo 5

# Aplicação do Modelo de Rasch estendido aos dados da Prova Brasil 2007

Será apresentado neste capítulo uma aplicação do modelo de Rasch estendido a dados reais para avaliação de sua performance. Os dados escolhidos provém da Prova Brasil 2007 e a primeira seção deste capítulo é dedicada a história da avaliação em larga escala e da própria prova. Na segunda seção este banco de dados é dividido em subpopulações e as estimativas dos parâmetros de dificuldade do modelo de Rasch tradicional são comparadas, procurando evidenciar uma possível heterogeneidade existente no fenômeno. Após isto, o modelo de Rasch estendido descrito no capítulo 4 é ajustado e comparado com o ajuste do modelo de Rasch tradicional. Nesta análise o objetivo será verificar o comportamento do modelo em diversas subpopulações. Estas subpopulações são geradas separando os alunos de acordo com suas respostas a determinadas questões do Questionário do Aluno (Anexo B). Os temas (variáveis) abordados por estas questões são possíveis fontes de heterogeneidade e o objetivo é verificar a influência destas fontes no desempenho dos alunos na prova.

## 5.1 SAEB - Sistema de Avaliação da Educação Básica

O SAEB - Sistema Nacional de Avaliação de Escola Básica é uma avaliação de larga escala aplicada a cada dois anos. Esta avaliação tem por objetivo principal realizar um diagnóstico do sistema educacional brasileiro e de alguns fatores que possam interferir no desempenho do aluno, fornecendo um indicativo sobre a qualidade do ensino que é ofertado (INEP, 2012).

De acordo com Oliveira (2011), se convencionou chamar a avaliação externa de avaliação de larga escala, tendo em vista o grande número de pessoas envolvidas em cada uma de suas etapas.

A autora cita em seu trabalho que a primeira grande avaliação em larga escala do Brasil foi realizada pela Coordenação de Aperfeiçoamento de Pessoal de Nível (Capes) em 1976. No entanto, na década de 60, grandes trabalhos relacionados a levantamento de dados educacionais foram realizados. Um importante acontecimento que possivelmente pode ter influenciado para a realização destes trabalhos pioneiros é a aprovação da Lei de Diretrizes e Bases da Educação Nacional (LDB), lei nº 4024, de 1961, que no seu art. 96 apresentava a preocupação do país em coletar este tipo de dados para subsidiar os planos de aplicação de recursos, tendo em vista a melhoria da qualidade do ensino (OLIVEIRA, 2011).

Durante a década de 80 e nos princípios de 1990, o Estado brasileiro realizou uma universalização do ensino fundamental, onde a escola se viu envolvida com a responsabilidade de promover a aprendizagem de todos os alunos, independente de origens e/ou classes sociais (MARCHELLI, 2010). Os dados levantados pelo Censo Escolar anualmente, juntamente com outros estudos, apresentaram altos índices de repetência e evasão escolas. De acordo com Pestana (2007 apud OLIVEIRA, 2011), estes dados indicavam a ineficiência das escolas e necessidade de promover ações para a melhoria da qualidade do ensino.

Diante da carência de informações sobre o processo de ensino e o impacto das políticas educacionais, iniciou-se o delineamento de um sistema de avaliação do sistema educacional. Dentre as experiências que contribuíram para a implementação

deste sistema destacam-se: a desenvolvida pela Fundação Cearense de Pesquisa, no âmbito do Programa de Expansão e Melhoria do Ensino no Meio Rural do Nordeste Brasileiro (Edurural) e o Programa de Educação Básica para o Nordeste, mais conhecido como Projeto Nordeste - Segmento Educação, financiado pelo Banco Mundial (BM).

Em 1988 foi estruturado o Sistema de Avaliação do Ensino Público de 1º grau (Saep) em respostas a demanda do BM junto ao Estado de desenvolver uma sistemática do Projeto Nordeste. Em 1989 o Saep foi colocado sob os cuidados do INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira e após uma revisão dos seus instrumentos por técnicos e especialistas se transformou no SAEB inicialmente apresentado.

A primeira aplicação SAEB foi realizada em 1990 e contou com a participação de uma amostra de escolas que ofertava as 1ª, 3ª, 5ª e 7ª séries do Ensino Fundamental das escolas públicas da rede urbana. A prova avaliava o desempenho dos alunos em língua portuguesa, matemática e ciências.

Em 1995 a Teoria da Resposta ao Item -TRI foi adotada como uma nova metodologia de construção dos testes e análises de resultados abrindo a possibilidade de comparabilidade entre os resultados das avaliações no decorrer do tempo. Neste ano houve mudanças no público avaliado para as séries finais do ciclo de escolarização: 4ª e 8ª série do Ensino Fundamental (atuais 5º e 9º ano, respectivamente) e o 3º ano do Ensino Médio. O SAEB deste ano foi inovador em relação aos anteriores ao inserir uma amostra da rede privada e não realizar os testes de Ciências nas aplicações. A partir da edição de 2001, o SAEB apenas avaliou os alunos nas áreas de Matemática e Língua Portuguesa, mantendo este padrão nos anos de 2003, 2005, 2007, 2009 e 2011.

O SAEB foi reestruturado em 2005 pela Portaria Ministerial nº 931 passando a ser composto por duas avaliações: uma avaliação amostral, a Avaliação Nacional da Educação Básica (Aneb) e uma avaliação censitária, a Avaliação Nacional do Rendimento Escolar (Anresc), conhecida como Prova Brasil.

De acordo com o INEP, a Prova Brasil foi idealizada para atender a demanda dos gestores públicos, educadores, pesquisadores e da sociedade em geral por informações sobre o ensino oferecido em cada município e escola. O objetivo da avaliação é auxiliar os governantes nas decisões e no direcionamento de recursos técnicos e financeiros,

assim como a comunidade escolar, no estabelecimento de metas e na implantação de ações pedagógicas e administrativas, visando à melhoria da qualidade do ensino.

## 5.2 Aplicação aos dados da Prova Brasil 2007

Nesta seção será apresentado um estudo visando o diagnóstico de uma possível existência de heterogeneidade de fonte desconhecida no banco de dados analisado e a comparação do modelo proposto com o modelo de Rasch tradicional. Para tal, foram utilizadas respostas de alunos da 4ª série (atual 5º ano) na prova de matemática na Prova Brasil 2007. No entanto, a aplicação da Prova Brasil é realizada por meio de 20 cadernos de provas diferentes onde os itens são sorteados utilizando a técnica de Blocos Incompletos Balanceados (BIB) (MONTGOMERY; 2004). Para este estudo foi escolhido o caderno número "1" e utilizado software SAS 9.2 e o procedimento NLMIXED.

Este banco contém as respostas de 109939 alunos para os 22 itens da prova de matemática e o gabarito das mesmas. Além disto, o INEP disponibiliza dados referentes a escola do aluno e às condições socio-econômicas do estudante e de seus pais. Este banco de dados pode ser baixado gratuitamente no site do INEP ([www.inep.gov.br](http://www.inep.gov.br)).

O próprio banco disponibiliza uma transformação das respostas alunos em variáveis indicadora, comparando as suas respostas com o gabarito de cada questão. Na comparação, caso o aluno acertasse o item a variável indicadora receberia o valor "1", caso contrário ela receberia o valor "0".

Feito isto, o passo seguinte foi ajustar o modelo de Rasch tradicional a este banco de dados. Para isto, utilizou-se o modelo de Rasch reescrito como membro da família dos Modelos Lineares Generalizados Mistos descrito na seção 4.3. Os traços latentes dos indivíduos foram tratados como efeitos aleatórios e os parâmetros de dificuldades dos 22 itens como efeitos fixos. Como foi apresentado no capítulo anterior, o modelo de Rasch tradicional foi redefinido da seguinte maneira:

$$\begin{aligned} f_{Y_{ij}|\Theta_i}(y_{ij}|\theta_i) &= [\pi_{ij}]^{y_{ij}}[1 - \pi_{ij}]^{1-y_{ij}} \\ \pi_{ij} &= \kappa(\theta_i - b_j) \\ \theta_i &\sim N(0, 1) \end{aligned} \tag{5.1}$$

sendo que a função  $\kappa(\cdot)$  é uma função de ligação a ser definida. Para este estudo a função de ligação adotada foi a logit, dada por  $\log[\pi/(1 - \pi)]$ .

Assim, a função de Verossimilhança para este novo modelo é descrita como:

$$L(\mathbf{Y} | (\theta_1, \dots, \theta_{109939}), (b_1, \dots, b_{22})) = \prod_{i=1}^{109939} \prod_{j=1}^{22} [\kappa(\theta_i - b_j)]^{y_{ij}} [1 - \kappa(\theta_i - b_j)]^{1-y_{ij}} \quad (5.2)$$

Logo, para realizar o ajuste do modelo foi utilizado o código SAS descrito no Anexo A. Para a estimação foi utilizado o método de integração numérica de Gauss não adaptativa e a técnica Quasi-Newton para a maximização da verossimilhança marginal. A estimação também foi feita utilizando a maximização Newton-Raphson, sendo esta mais lenta que a anterior. A tabela com os valores estimados dos parâmetros de dificuldade para cada item segue no Anexo B.

Este ajuste será importante para futuras comparações, principalmente na análise do comportamento do ajuste do modelo de Rasch estendido.

### 5.3 Apontamentos da existência de uma possível heterogeneidade no banco de dados

Nesta seção são descritos ajustes do modelo de Rasch a grupos existentes no banco de dados que demonstram a possível existência de heterogeneidade que afeta no desempenho dos alunos nos itens da prova. Para comprovar isto, as perguntas “2”, “35” e “38” do questionário do aluno (Anexo B) foram escolhidas. Nesta seção foram feitas análises específicas para cada variável separando-se um banco para cada uma de suas respostas válidas. Após isto foram feitas CCI’s para cada uma dessas subpopulações e comparadas graficamente. Este exercício é importante pois demonstra como estas variáveis externas conhecidas influenciam no desempenho do aluno nos itens da prova.

Destaca-se ainda que para os ajustes dos modelos foi utilizado o software SAS e o procedimento NLMIXED. Toda a programação utilizada para o ajuste está disponível nos Anexo A.

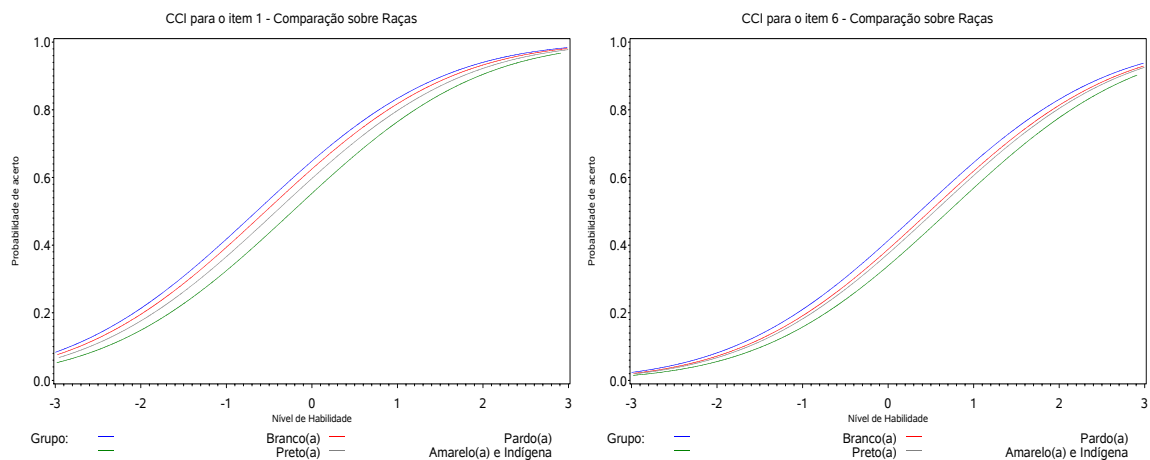
### 5.3.1 Comparação étnico-racial

A questão do questionário “2” é sobre como o aluno se considerava diante das cinco opções étnicas-raciais apresentadas: *branca*, *parda*, *negra*, *branca*, *amarela* e *indígena*. Para este estudo as respostas *amarela* e *indígena* foram agrupadas devido ao pequeno contingente de respondentes destas opções.

Após isto, estas respostas foram ajustadas ao modelo de Rasch, modelo de resposta ao item de Rasch que foi descrito na seção 3.7, e utilizando a função de ligação logit (seção 3.2), a fim de investigar se há uma grande diferença na probabilidade de acerto de uma questão dada sua habilidade entre os dois grupos. Foi suposto ainda que o parâmetro de habilidade  $\theta$  tem distribuição Normal com média e variância “0” e “1”, respectivamente.

Todavia, realizado o ajuste do modelo para os grupos em separado, chegou-se as seguintes estimativas dos parâmetros de dificuldade  $b$  para cada um dos 22 itens da prova para ambos os grupos. As tabelas com as estimativas dos parâmetros de dificuldade dos itens seguem no Anexo B.

Nota-se que os níveis de dificuldade do item ajustados para o banco referentes aos respondentes que se consideravam negros estão bem acima das estimativas para os demais grupos. Isto evidencia que os integrantes deste grupo deveriam ter um nível de habilidade maior que os demais para atingir uma probabilidade de acerto do  $p$  para um determinado item. Esta disparidade fica mais evidente se compararmos graficamente as CCI’S de cada grupo.



(a) CCI's para o item 1

(b) CCI's para o item 6

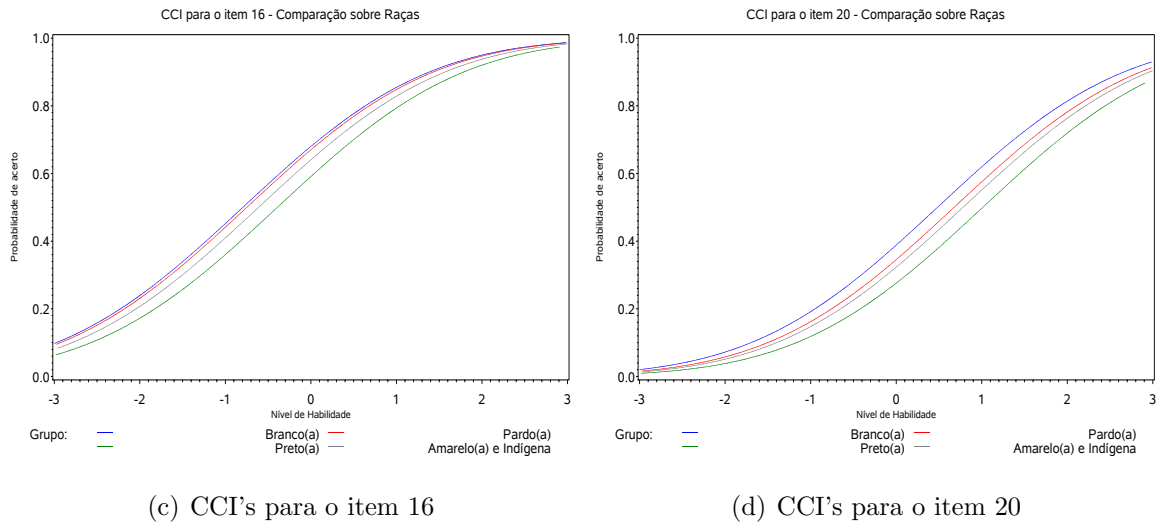


Figura 5.1: CCI's comparativas - Questão 2

Para este caso foi possível separar os indivíduos em dois grupos, no entanto nem sempre isso será possível ou até mesmo viável. Assim, a criação de um modelo que consiga ajustar essa variação proveniente de uma fonte, ainda que ela seja desconhecida, é tão importante e relevante.

### 5.3.2 Comparação relacionada ao trabalho fora de casa

A questão “35” busca saber se o aluno trabalha fora de casa ou não. O banco foi dividido em um banco com as respostas dos alunos de que trabalham fora de casa e um outro banco com as respostas dos alunos que não. Após isto, o modelo de Rasch foi ajustado gerando as estimativas que se encontram no Anexo B.

Para melhor entendimento do caso foram feitos gráficos confrontando as CCI's de cada subpopulação para cada item. Como pode ser visto nos gráficos, as curvas características do itens para os alunos que não trabalham fora estão acima das CCI's correspondentes aos alunos que trabalham. Logo, as probabalidades de acerto para estes alunos tendem a ser maiores. No entanto, este comportamento foi observado para todos os itens, exceto para o item 13, no qual a CCI dos alunos que trabalham fora está um pouco acima da curva referente ao grupo que não trabalha. Estes gráficos indicam que o fato do aluno trabalhar fora de casa pode estar influenciando de modo direto ou indireto no desempenho deste aluno na prova.

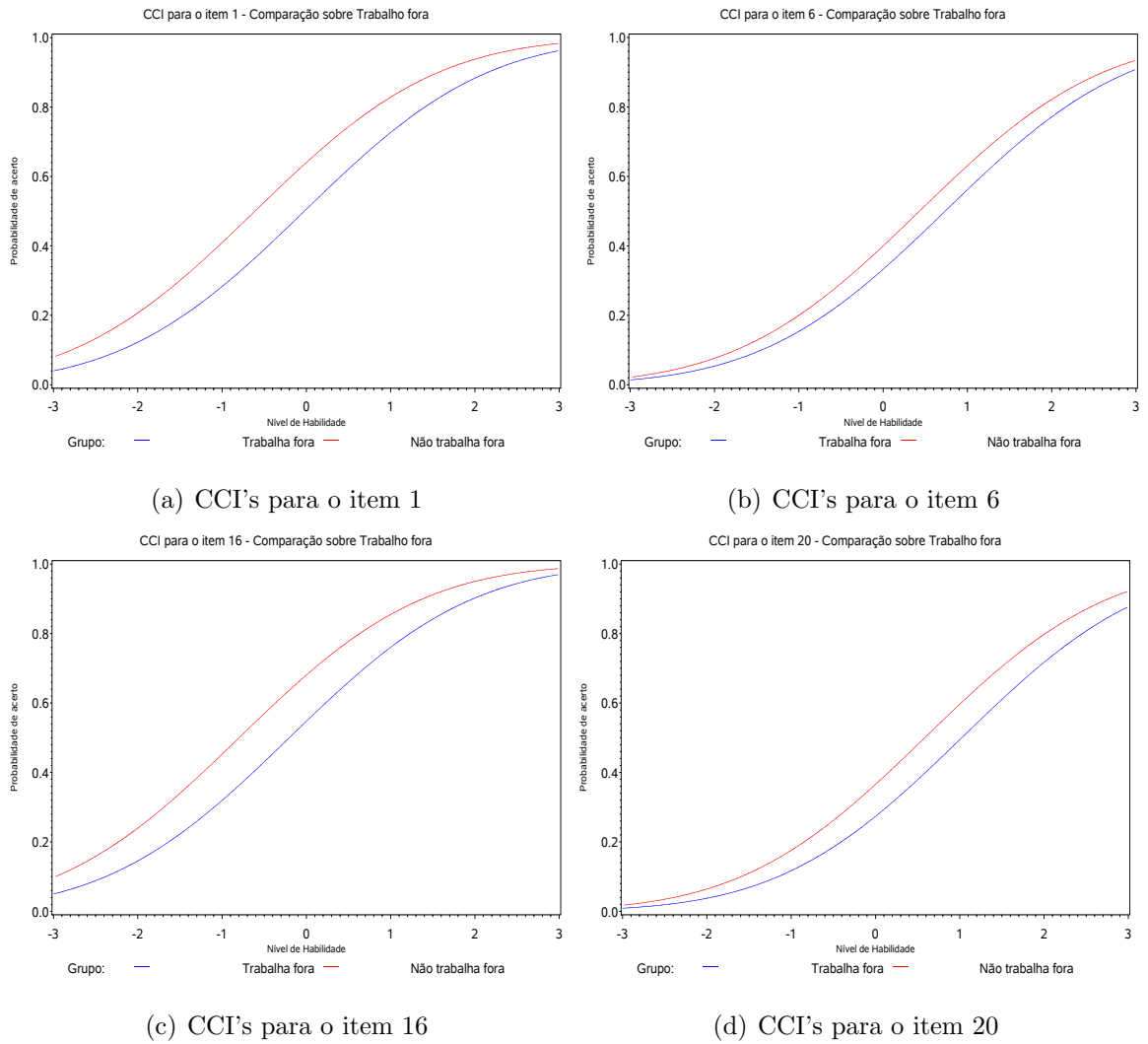


Figura 5.2: CCI's comparativas - Questão 35

### 5.3.3 Comparação relacionada a reprovação dos alunos

Nesta seção o banco de dados da Prova Brasil foi dividido em outras duas subpopulações. Utilizando a questão “38”, que pergunta ao aluno se ele já havia sido reprovado ou não, os alunos foram divididos em “alunos que nunca reprovaram” e “alunos que reprovaram pelo menos uma vez”. Esta divisão foi realizada imaginando uma possível diferença entre o desempenho dos alunos destes dois grupos.

Utilizando os mesmos passos descritos na comparação anterior, os parâmetros de dificuldade dos itens foram estimados e podem ser conferidos no Anexo B. Nota-se uma diferença entre as estimativas levando a suspeitar que os alunos que não reprovaram tiveram mais facilidade para acertar os itens que os alunos que estudavam já haviam reprovado pelo menos uma vez. Esta conclusão fica mais clara ao comparar



as CCI's das duas subpopulações.

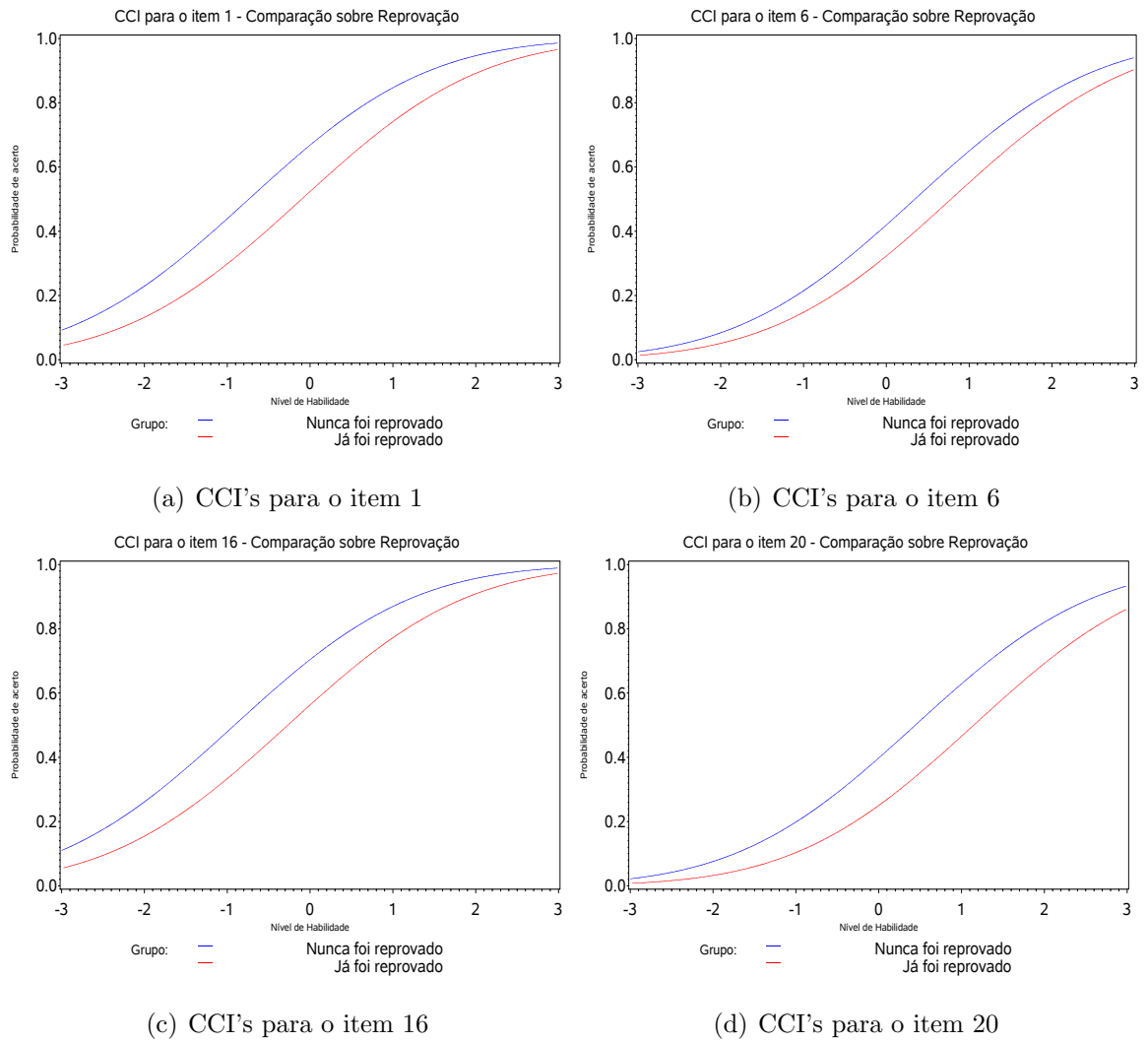


Figura 5.3: CCI's comparativas - Questão 38

### 5.3.4 Comentários

Ao realizar estes exercícios com o banco de dados é evidente que uma possível heterogeneidade pode estar influenciando no desempenho dos alunos. Logo, esta heterogeneidade apresentada acima acontece simultaneamente, demonstrando que fatores não controlados podem estar atuando sobre o modelo de Rasch. No entanto, o modelo proposto por este trabalho não define quais são estas influências e tenta capturar esta heterogeneidade desconhecida, sendo este um dos avanços apresentados por este trabalho.

## 5.4 Aplicação do Modelo de Rasch Estendido aos dados do SAEB

O passo seguinte realizado foi ajustar o modelos de Rasch estendido ao banco de dados com as respostas dos alunos da 4ª série (5º ano) na prova de matemática da Prova Brasil 2007 estudados na seção anterior. Nesta etapa ajustou-se aos dados os modelos de Rasch estendido utilizando a função de Máxima Verossimilhança Marginal (MVM) descrito em (4.11) para as funções de ligação logit (REML) e probit (REMP) e o modelo de Rasch estendido utilizando o resultado analítico de Molenberghs, Verbeke, Demétrio e Vieira (2010)(REMol) descrito em (4.12) . Foi utilizada a reparametrização da distribuição Beta (FERRARI;CRIBARI-NETO, 2004) para o ajuste do modelo estendido devido a dificuldades enfrentadas no ajuste dos modelos.

Para a realizar a ajuste de um determinado modelo utilizando o procedimento NLMIXED do software SAS é necessária a descrição do núcleo de sua função de máxima verossimilhança. Para os modelos REML e REMP a função de MVM é apresentada e, por consequência, o núcleo também. No entanto, para o modelo REMol a função de Máxima Verossimilhança (MV) é determinada por:

$$L(\mu, \mathbf{b}|\mathbf{Y}) = \prod_{i=1}^N \prod_{j=1}^J [\mu\Phi(-b_j; 2)]^{y_{ij}} [1 - \mu\Phi(-b_j; 2)]^{1-y_{ij}} \quad (5.3)$$

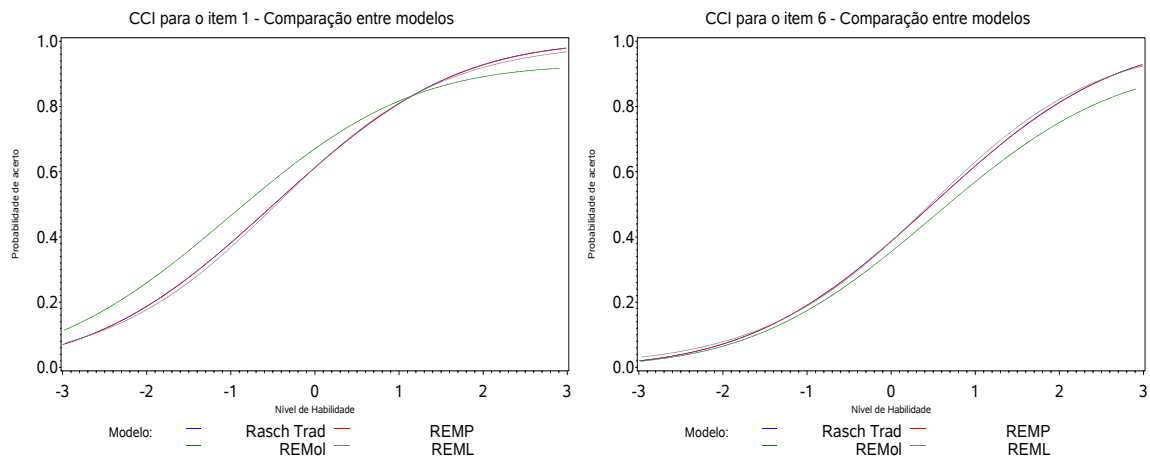
Determinadas as funções de MV's, os modelos foram ajustados e as tabelas dos parâmetros estimados e dos seus respectivos erros-padrões seguem no Anexo B. Analisando estas tabelas notou-se que as estimativas do parâmetro aleatório para acomodar o efeito de superdispersão dos modelos de Rasch estendido ficaram muito próximas de '1'. Entretanto, para o intervalo de confiança para a média do efeito  $\gamma$  não contém o valor 1, a um nível de confiança de 95%. Isto significa que não há uma evidência suficiente para dizer não há heterogeneidade atribuída à fonte desconhecida agindo sobre os dados, a um nível de significância de 5%.

Outro ponto a ser considerado é que as estimativas dos parâmetros de dificuldade são praticamente iguais, exceto as estimativas do modelo REMol. É necessário enfatizar que as estimativas do modelo REMol podem diferir das estimativas dos demais modelos, pois este não depende de integrais numéricas e os modelos REMP e REML são modelos marginais em duas dimensões ( $\theta$  e  $\gamma$ ). Para estes casos e para

o modelo de Rasch tradicional foi feita uma quadratura gaussiana com um ponto de quadratura, logo, a integral foi aproximada pelo método de Laplace (AZEVEDO-FILHO;SHACHTER,1994).

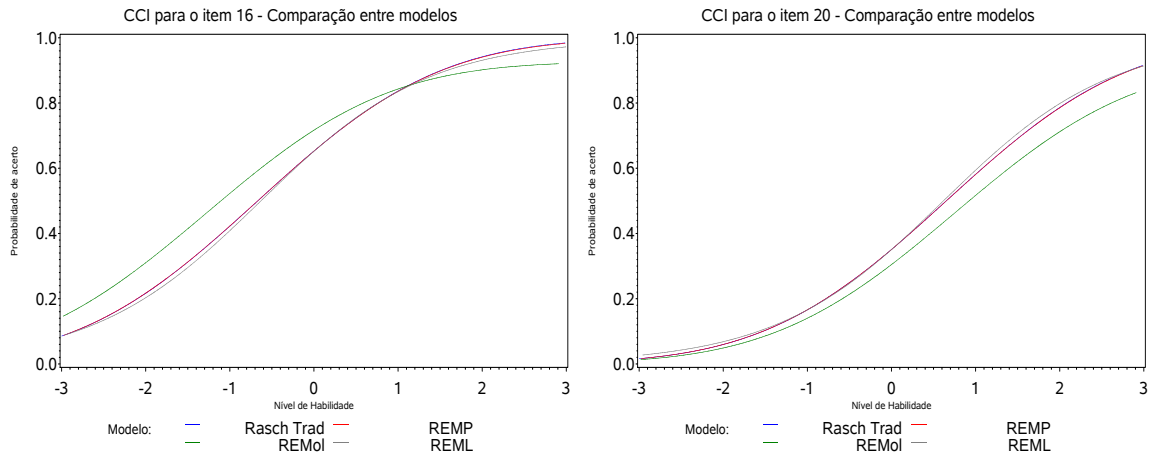
Diante disto, os erros-padrões estimados para as estimativas dos itens do modelo REMol foram maiores. Isto significa que o modelo conseguiu capturar uma maior variação existente nas respostas dos estudantes ao teste modelando com mais eficiência a heterogeneidade de fonte desconhecida que age sobre o modelo, neste caso, os fatores identificados na seção anterior. Por consequência, esta variação reflete nos intervalos de confiança, tornando-os mais amplos do que medidos pelos outros 3 modelos apresentados.

Em seu trabalho, Vieira (2008) utilizou os modelos descritos na seção 2.5 para realizar um estudo de comparação de dois medicamentos via oral para tratamento de uma patologia dermatológica. Neste trabalho, o modelo utilizando o resultado de Molenberghs et al (2010) também acomodou melhor a variabilidade existente captando assim o efeito de superdispersão. Logo, o comportamento do modelo foi similar ao comportamento do modelo REMol. Assim, modelo de Rasch estendido tem comportamento similar ao um MLG com fator de heterogeneidade constante, onde as estimativas não se alteram e sim, os erros padrões que são ajustados (HINDE; DEMÉTRIO, 1998).



(a) Comparação das CCI's para o item 1

(b) Comparação das CCI's para o item 6



(c) Comparação das CCI's para o item 16

(d) Comparação das CCI's para o item 20

Figura 5.4: CCI's comparativas dos modelos ajustados

Analisando graficamente, nota-se que para todos os itens as CCI's do modelo de Rasch tradicional e as CCI's dos modelos REMP e REML são praticamente iguais. Por outro lado, as CCI's do modelo REMol raramente coincidem com as CCI's dos outros modelos. Um comportamento visto é que as CCI's para os itens ajustados pelo modelo REMol ficam mais próximas das outras três quando o erro-padrão se aproxima do valor (0,09). Isto pode ser notado nos gráficos que seguem.

## 5.5 Conclusões e perspectivas de trabalhos futuros

Foi objetivo deste trabalho adaptar o modelo de resposta ao item de Rasch adicionando a este modelo um parâmetro multiplicativo de distribuição de probabilidade Beta a fim de ajustar uma possível heterogeneidade atribuída à fonte desconhecida. Para isto, o modelo de Rasch foi inserido na família dos Modelos Lineares Generalizados Mistos e utilizou-se o trabalho de Molenberghs et al. (2010) como referência teórica para o desenvolvimento deste novo modelo.

Após isto, uma série de simulações foram feitas, criando banco de dados com o efeito de superdispersão e o desempenho dos modelos de Rasch Tradicional, REMP e REMol foram comparados. Neste exercício, o modelo REMol conseguiu ajustar muito bem os parâmetros dos itens e a média do efeito aleatório  $\gamma$  que causava a perturbação nas respostas simuladas. Em contrapartida, os modelos de Rasch tradicional e REMP não conseguiram bons ajustes. Além disso, o modelo REMol capturou uma maior

variabilidade dos dados e ajustou os dados mais rapidamente que os demais modelos.

Para exemplificar o problema da heterogeneidade em um banco de dados real, foi realizado um estudo utilizando o banco de dados da Prova Brasil 2007 referente às respostas dos alunos da 4ª série na prova de matemática - Caderno 1. Através de estudos comparativos utilizando as CCT's de determinadas subpopulações foram notadas disparidades nas estimativas dos parâmetros de dificuldades dos itens, evidenciando uma possível existência de heterogeneidade de fonte desconhecida não capturada pelo ajuste do modelo de Rasch tradicional.

Ao aplicar os modelos de Rasch estendido ao banco de dados observou-se que as estimativas dos parâmetros de dificuldade dos itens foram praticamente iguais às estimativas do modelo de Rasch Tradicional, exceto para o modelo REMol e, além disso, os erro-padrões das estimativas dos parâmetros para este modelo foram maiores. Mesmo que a estimativa do parâmetro que ajustaria a heterogeneidade atribuída a fonte de desconhecida tenha ficado próxima de "1", ao se testar estatisticamente a possibilidade do valor  $\mu$  populacional ser igual a "1", esta hipótese foi rejeitada. Além disto, fato do erro-padrão ter aumentado também leva a crer que o modelo conseguiu captar esta heterogeneidade resultando na ampliação dos intervalos de confiança dos parâmetros estimados. Cabe ressaltar que nada aqui foi discutido sobre o impacto desta heterogeneidade atribuída à fonte desconhecida sobre a estimação das habilidades, quando esta é feita de forma simultânea. Esta é uma questão para estudos futuros.

No entanto, para a conclusão deste trabalho grandes dificuldades computacionais foram enfrentadas. Em um primeiro momento, a distribuição Beta definida para o efeito multiplicativo  $\gamma$  apresentou problemas de identificabilidade nos ajustes. Depois de inúmeras tentativas frustradas, utilizou-se a parametrização de Ferrari & Cribari-Neto (2004) limitando a estimativa do parâmetro  $\mu$  da distribuição, sendo que o parâmetro  $\phi$  não aparece na função de verossimilhança utilizada para o ajuste do modelo. Outro problema foi o tempo gasto com a estimação utilizando a integração numérica pelo método da Quadratura Gaussiana adaptativa. Diante do grande tempo demandado optou-se pela Quadratura Gaussiana não-adaptativa, sendo mais simples e mais rápida.

Porém este trabalho ainda não está completo e alguns pontos ainda deverão ser

discutidos em trabalhos futuros. Devido a problemas computacionais, que serão descritos no Anexo C do trabalho, alguns estudos de simulação do modelo, além do apresentado neste trabalho, não puderam ser feitos. Desejava-se observar ainda como o modelo de Rasch estendido se comportaria quando o número de avaliados fosse muito pequeno e muito grande, caso este muito comum em grande provas como o ENEM. Outros cenários relacionados à variação do parâmetro aleatório e às estruturas de variâncias do vetor dos parâmetros de dificuldades do itens também seriam estudados. Logo, estes estudos serão realizados em um momento posterior.

Outra possibilidade de trabalho seria colocar um preditor linear para modelar o efeito de superdispersão. Neste caso, ao invés de adotar um parâmetro aleatório com distribuição Beta para captar a heterogeneidade de fonte desconhecida, um preditor linear composto por variáveis externas ao fenômeno que poderiam estar influenciando no desempenho do indivíduo causando uma perturbação e testar esta nova abordagem em um banco de dados reais.

# Referências Bibliográficas

- [1] ANDRADE, D. F.; TAVARES, H.R.; VALLE, R.C. **Teoria de Resposta ao Item: Conceito e aplicações**. São Paulo: Associação Brasileira de Estatística, 2000.
- [2] AZEVEDO-FILHO, A.; SHACTER, R. Laplace's Method Approximations for Probabilistic Inference in Belief Networks with Continuous Variables. **UAI'94 Proceedings of the Tenth international conference on Uncertainty in artificial intelligence**. San Francisco: Morgan Kaufmann Publishers Inc. ,1994, pg 28-36.
- [3] BAKER, F. B.; KIM, S.K., **Item Theory Response - Parameter Estimation Techniques**. Nova Iorque: Dekker, 2004, 528 p.
- [4] De AYALA, **The theory and practice of item response theory**. Nova Iorque: The Guilford Press, 2009.
- [5] DEMÉTRIO, C.G.B.; CORDEIRO, G.M. **Modelos Lineares Generalizados**. In: Minicurso para o 12º SEAGRO e 52ª Reunião Anual da RBRAS, UFSM, 2007. Disponível em: < <http://www.lce.esalq.usp.br/clarice/livroSeagro.pdf> >. Acesso em: 4 mar, 2011.
- [6] DEMÉTRIO, C.G.B. **Modelos Lineares Generalizados em Experimentação Agronômica**. Piracicaba: ESALQ/USP, 2002, 113 p.
- [7] DOBSON, A. J. **An Introduction to Generalized Linear Models**. 2ed. Florida: Chapman and Hall, 2002, 218 p.
- [8] FERRARI, S. L. P.; CRIBARI-NETO, F. Beta Regression for Modelling Rates and Proportions. **Journal of Applied Statistics**, v.31, n.7, 2004, pg 799-815.

- [9] HINDE, J.; DEMÉTRIO, C.G.B. Overdispersion: Models and Estimation. **Computational Statistics and Data Analysis**, Volume 27, Abril, 1998. 151-170.
- [10] INEP. **História da Prova Brasil e do Saeb**. Disponível em: <<http://portal.inep.gov.br/web/prova-brasil-e-saeb/historico>>. Acesso em: 9 jun, 2011.
- [11] JIANG, J. **Linear and Generalizes Linear Mixed Models and Their Applications**. Nova Iorque: Springer. 2007. 253 p.
- [12] KARINO, C.A.; LAROS, J.A.; JESUS, G.R. Funcionamento Diferencial dos Itens do Teste Não-Verbal de Inteligência SON-R 2 $\frac{1}{2}$ -7[a]. **Psicologia: Teoria e Pesquisa**. Vol. 28 n. 1, pg. 15-25.
- [13] LOCCI, V; VASCONCELLOS, L.A.S. **Métodos Quase-Newton: O Método de Broyden**. São Paulo, UNESP. Disponível em: [http://wwwp.fc.unesp.br/arbalbo/Iniciacao\\_Cientifica/sistemasnaolineares/teoria/2\\_Met\\_de\\_Broyden.pdf](http://wwwp.fc.unesp.br/arbalbo/Iniciacao_Cientifica/sistemasnaolineares/teoria/2_Met_de_Broyden.pdf). Acesso em: 10 abril, 2011.
- [14] MARCHELLI, P.S. Expansão e qualidade da Educação Básica no Brasil **CADERNOS DE PESQUISA**, 2010. v. 40, n.140, pg. 561-585.
- [15] MCCULLOCH, C.E.;SEARLE, S.R. **Generalized, Linear and Mixed Models**. Nova Iorque: John Wiley and Sons, 2001, 346 p.
- [16] MOLENBERGHS, G.; VERBEKE, G. **Models For Discrete Longitudinal Data**. Nova Iorque: Springer, 2005, 683 p.
- [17] MONTGOMERY, D. G.**Design and Analysis of Experiments**.Nova Iorque: John Wiley and Sons, 2004, pg 148.
- [18] MOLENBERGHS, G.; VERBEKE, G.; DEMÉTRIO, C.G.B.; VIEIRA, A.M.C. A family of generalized linear models for repeated measures with normal e conjugate random effects. **Statistical Science**, 2010, Vol. 25, No. 3, pg 325 - 334.



- [19] NELDER, J.A.; WEDDERBURN, R. W. M. Generalized Linear Models. **Journal of the Royal Statistical Society. Series A (General)**, 1972. v. 135, n. 3, pg 370-384.
- [20] NOCEDAL, J.; WRIGHT, S. J. **Numerical Optimization**. Nova Iorque: Springer, 1999, 641 p.
- [21] OLIVEIRA, A.P. M. **A Prova Brasil como política de regulação da rede pública do Distrito Federal**. Dissertação de Mestrado. Faculdade de Educação, Universidade de Brasília, Brasília, 2011.
- [22] PESTANA, Maria Inês. **O sistema de avaliação brasileiro**. In: **Seminário Internacional de Avaliação Educacional**. Anais. Rio de Janeiro, dez. 1997. p. 15-29. *apud* OLIVEIRA, A.P. M. **A Prova Brasil como política de regulação da rede pública do Distrito Federal**. Dissertação de Mestrado. Faculdade de Educação, Universidade de Brasília, Brasília, 2011.
- [23] SAS INSTITUTE. **SAS Software Release 9.1.3**. Cary: SAS Institute, 2008.
- [24] VERBEKE, G.; MOLENBERGHS, G. **Linear Mixed Models for Longitudinal Data**. Nova Iorque: Springer, 2000, 579 p.
- [25] VIEIRA, A. M. C. **Modelagem simultânea de média e dispersão e aplicações na pesquisa agrônômica**. Piracicaba: ESALQ, Universidade de São Paulo. Tese de Doutorado. 2008, 117 p.
- [26] YANG, W. Y.; CAO, W.; MORRIS, J. **Applied numerical methods using MATLAB**. Nova Jersey: Wiley and Sons, 2005, 509 p.
- [27] ZEGER, S.L.; LIANG, K.; ALBERT, P.S. Models for longitudinal data: a generalized estimating equation approach. **Biometrics**, Washington, v. 44, n. 4, pg. 1049-1060.

# Anexo A

A1: Programação utilizada para ajustar o modelo de Rasch tradicional aos dados da Prova Brasil 2007.

```
proc nlmixed data=db.mt42007b technique=quanew maxiter=500 alpha=0.001  
qpoints=1 noad noadscale;  
PARMS b1-b22=0;  
beta=i1*b1+i2*b2+i3*b3+i4*b4+i5*b5+i6*b6+i7*b7+i8*b8+i9*b9+i10*b10+i11*b11+  
i12*b12+i13*b13+i14*b14+i15*b15+i16*b16+i17*b17+i18*b18+i19*b19+i20*b20+i21*b21+i22*b22;  
eta =theta - beta;  
pb=probnorm(eta*(16*sqrt(3))/(15*3.141593));  
model resp ~ binary(pb);  
random theta ~ normal(0,1) subject=id_ aluno;  
ods output ParameterEstimates=pe;  
run;
```

A2: Programação utilizada para ajustar o modelo de Rasch estendido utilizando a máxima verossimilhança marginal e a função de ligação logit aos dados da Prova Brasil 2007.

```
proc nlmixed data=db.mt42007b method=gauss technique=quanew maxiter=500  
qpoints=1 noad noadscale empirical alpha=0.001;  
parms mu=0.85 / data=pe;  
bounds 0<mu<1;  
beta=i1*b1+i2*b2+i3*b3+i4*b4+i5*b5+i6*b6+i7*b7+i8*b8+i9*b9+ i10*b10+i11*b11+i12*b12+  
i13*b13+i14*b14+i15*b15+i16*b16+i17*b17+i18*b18+i19*b19+ i20*b20+i21*b21+i22*b22; eta =theta - beta;  
pb=1/(1+exp(-theta+beta));;
```

```

ll = resp*log(mu) + resp*log(pb) + (1-resp)*log(1-mu*pb);

model resp ~ general(ll);

random theta ~ normal(0,1) subject=id_ aluno;

run;

```

A3: Programação utilizada para ajustar o modelo de Rasch estendido utilizando a máxima verossimilhança marginal e a função de ligação probit aos dados da Prova Brasil 2007.

```

proc nlmixed data=db.mt42007b method=gauss technique=quanew maxiter=500

qpoints=1 noad noadscale empirical alpha=0.001;

parms mu=0.85 / data=pe;

bounds 0<mu<1;

beta=i1*b1+i2*b2+i3*b3+i4*b4+i5*b5+i6*b6+i7*b7+i8*b8+i9*b9+ i10*b10+i11*b11+i12*b12+
i13*b13+i14*b14+i15*b15+i16*b16+i17*b17+i18*b18+i19*b19+ i20*b20+i21*b21+i22*b22;

eta =theta - beta;

pb=probnorm(eta*(16*sqrt(3))/(15*3.141593));

ll = resp*log(mu) + resp*log(pb) + (1-resp)*log(1-mu*pb);

model resp ~ general(ll);

random theta ~ normal(0,1) subject=ID_ aluno;

run;

```

A4: Programação utilizada para ajustar o modelo de Rasch estendido utilizando o resultado de Molenberghs et al. (2010) e a ligação probit aos dados da Prova Brasil 2007

```

proc nlmixed data=db.mt42007b technique=quanew maxiter=500 alpha=0.001;

parms mu=0.85 / data=pe;

bounds 0< mu < 1;

beta=i1*b1+i2*b2+i3*b3+i4*b4+i5*b5+i6*b6+i7*b7+i8*b8+i9*b9+ i10*b10+i11*b11+i12*b12+
i13*b13+i14*b14+i15*b15+i16*b16+i17*b17+i18*b18+i19*b19+ i20*b20+i21*b21+i22*b22;

eta = -beta*(16*sqrt(3))/(15*3.141593);

pb=CDF('NORMAL',eta,0,sqrt(2));

ll = resp*log(mu) + resp*log(pb) + (1-resp)*log(1-mu*pb);

```

```
model resp = general(11);  
  
run;
```

## Anexo B

Tabela B1: Parâmetros de dificuldade para a simulação

Parâmetro	Valor original
$b_1$	-1
$b_2$	-0,5
$b_3$	1
$b_4$	0,8
$b_5$	-0,75
$b_6$	1,25
$b_7$	1,35
$b_8$	0,05
$b_9$	-0,25
$b_{10}$	-0,03
$b_{11}$	0,57
$b_{12}$	1,45
$b_{13}$	-1,96
$b_{14}$	1,75

Tabela B2: Médias das estimativas para os dados simulados - Média 0.7

Parâmetros	Valor original	Analítico		MVM		Rasch Tradicional	
		Estimativa	Erro-Padrão	Estimativa	Erro-Padrão	Estimativa	Erro-Padrão
$b_1$	-1,000	-1,044	1,261	-0,294	0,017	-0,128	0,012
$b_2$	-0,500	-0,531	0,864	0,040	0,017	0,170	0,012
$b_3$	1,000	0,988	0,386	1,282	0,018	1,320	0,014
$b_4$	0,800	0,787	0,421	1,103	0,017	1,152	0,014
$b_5$	-0,750	-0,788	1,037	-0,134	0,017	0,014	0,012
$b_6$	1,250	-1,314	1,581	-0,442	0,018	-0,255	0,012
$b_7$	1,350	1,336	0,337	1,599	0,019	1,619	0,016
$b_8$	0,050	0,028	0,614	0,463	0,016	0,556	0,013
$b_9$	-0,250	-0,277	0,735	0,226	0,016	0,339	0,012
$b_{10}$	-0,030	-0,055	0,644	0,396	0,016	0,495	0,013
$b_{11}$	0,570	0,553	0,469	0,901	0,017	0,963	0,013
$b_{12}$	1,450	1,437	0,324	1,690	0,019	1,706	0,016
$b_{13}$	-1,960	-2,080	3,417	-0,761	0,019	-0,520	0,013
$b_{14}$	1,750	1,742	0,291	1,972	0,021	1,973	0,018
$\mu$	0,7	0,693	0,244	0,922	0,003	-	-

Tabela B3: Médias das estimativas para os dados simulados - Média 0,75

Parâmetros	Valor original	Analítico		MVM		Rasch Tradicional	
		Estimativa	Erro-Padrão	Estimativa	Erro-Padrão	Estimativa	Erro-Padrão
$b_1$	-1,000	-1,043	0,873	-0,399	0,017	-0,252	0,012
$b_2$	-0,500	-0,529	0,599	-0,045	0,017	0,068	0,012
$b_3$	1,000	0,987	0,268	1,237	0,017	1,266	0,014
$b_4$	0,800	0,786	0,292	1,054	0,017	1,093	0,014
$b_5$	-0,750	-0,784	0,716	-0,227	0,017	-0,098	0,012
$b_6$	1,250	-1,300	1,084	-0,553	0,018	-0,387	0,013
$b_7$	1,350	1,342	0,233	1,562	0,018	1,575	0,015
$b_8$	0,050	0,028	0,426	0,394	0,016	0,474	0,013
$b_9$	-0,250	-0,273	0,508	0,149	0,016	0,246	0,012
$b_{10}$	-0,030	-0,048	0,445	0,330	0,016	0,414	0,013
$b_{11}$	0,570	0,554	0,325	0,847	0,017	0,898	0,013
$b_{12}$	1,450	1,443	0,225	1,656	0,019	1,665	0,016
$b_{13}$	-1,960	-2,073	2,338	-0,905	0,019	-0,684	0,013
$b_{14}$	1,750	1,738	0,203	1,934	0,020	1,931	0,017
$\mu$	0,75	0,741	0,181	0,937	0,003	-	-

Tabela B4: Médias das estimativas para os dados simulados - Média 0,8

Parâmetros	Valor original	Analítico		MVM		Rasch Tradicional	
		Estimativa	Erro-Padrão	Estimativa	Erro-Padrão	Estimativa	Erro-Padrão
$b_1$	-1,000	-1,031	2,502	-0,514	0,017	-0,389	0,012
$b_2$	-0,500	-0,523	1,722	-0,135	0,016	-0,042	0,012
$b_3$	1,000	0,990	0,774	1,192	0,017	1,211	0,014
$b_4$	0,800	0,787	0,845	1,002	0,017	1,029	0,014
$b_5$	-0,750	-0,780	2,061	-0,333	0,017	-0,224	0,012
$b_6$	1,250	-1,294	3,115	-0,684	0,018	-0,541	0,013
$b_7$	1,350	1,338	0,675	1,521	0,018	1,526	0,015
$b_8$	0,050	0,031	1,230	0,323	0,016	0,386	0,013
$b_9$	-0,250	-0,269	1,462	0,071	0,016	0,150	0,012
$b_{10}$	-0,030	-0,053	1,290	0,252	0,016	0,320	0,012
$b_{11}$	0,570	0,558	0,937	0,790	0,016	0,827	0,013
$b_{12}$	1,450	1,438	0,650	1,615	0,018	1,617	0,016
$b_{13}$	-1,960	-2,048	6,572	-1,069	0,020	-0,870	0,013
$b_{14}$	1,750	1,741	0,583	1,906	0,020	1,897	0,017
$\mu$	0,8	0,792	0,559	0,951	0,003	-	-



Tabela B5:Médias das estimativas para os dados simulados - Média 0,85

Parâmetros	Valor original	Analítico		MVM		Rasch Tradicional	
		Estimativa	Erro-Padrão	Estimativa	Erro-Padrão	Estimativa	Erro-Padrão
$b_1$	-1,000	-1,050	2,774	-0,634	0,017	-0,527	0,013
$b_2$	-0,500	-0,534	1,902	-0,231	0,016	-0,153	0,012
$b_3$	1,000	0,983	0,849	1,145	0,017	1,160	0,014
$b_4$	0,800	0,785	0,926	0,955	0,017	0,976	0,013
$b_5$	-0,750	-0,792	2,279	-0,438	0,017	-0,346	0,012
$b_6$	1,250	-1,312	3,458	-0,819	0,018	-0,696	0,013
$b_7$	1,350	1,339	0,737	1,484	0,018	1,487	0,015
$b_8$	0,050	0,029	1,348	0,257	0,016	0,309	0,012
$b_9$	-0,250	-0,279	1,614	-0,013	0,016	0,053	0,012
$b_{10}$	-0,030	-0,055	1,411	0,183	0,016	0,238	0,012
$b_{11}$	0,570	0,551	1,030	0,736	0,016	0,766	0,013
$b_{12}$	1,450	1,435	0,712	1,579	0,018	1,579	0,015
$b_{13}$	-1,960	-2,095	7,540	-1,255	0,020	-1,078	0,013
$b_{14}$	1,750	1,739	0,640	1,874	0,020	1,866	0,017
$\mu$	0,85	0,838	0,650	0,964	0,002	-	-

Tabela B6:Médias das estimativas para os dados simulados - Média 0,9

Parâmetros	Valor original	Analítico		MVM		Rasch Tradicional	
		Estimativa	Erro-Padrão	Estimativa	Erro-Padrão	Estimativa	Erro-Padrão
$b_1$	-1,000	-1,066	2,146	-0,764	0,017	-0,677	0,013
$b_2$	-0,500	-0,547	1,466	-0,328	0,016	-0,265	0,012
$b_3$	1,000	0,983	0,650	1,101	0,017	1,110	0,014
$b_4$	0,800	0,777	0,710	0,902	0,016	0,918	0,013
$b_5$	-0,750	-0,803	1,760	-0,552	0,017	-0,477	0,013
$b_6$	1,250	-1,331	2,690	-0,969	0,018	-0,867	0,013
$b_7$	1,350	1,332	0,565	1,442	0,018	1,443	0,015
$b_8$	0,050	0,020	1,034	0,182	0,016	0,222	0,012
$b_9$	-0,250	-0,289	1,244	-0,099	0,016	-0,047	0,012
$b_{10}$	-0,030	-0,063	1,085	0,104	0,016	0,147	0,012
$b_{11}$	0,570	0,545	0,789	0,680	0,016	0,702	0,013
$b_{12}$	1,450	1,435	0,544	1,542	0,018	1,540	0,015
$b_{13}$	-1,960	-2,127	6,071	-1,474	0,021	-1,317	0,014
$b_{14}$	1,750	1,741	0,489	1,840	0,019	1,833	0,017
$\mu$	0,9	0,883	0,521	0,974	0,002	-	-

Tabela B7:Médias das estimativas para os dados simulados - Média 0,95

Parâmetros	Valor original	Analítico		MVM		Rasch Tradicional	
		Estimativa	Erro-Padrão	Estimativa	Erro-Padrão	Estimativa	Erro-Padrão
$b_1$	-1,000	-1,062	1,940	-0,902	0,017	-0,837	0,013
$b_2$	-0,500	-0,542	1,317	-0,427	0,016	0,061	0,012
$b_3$	1,000	0,980	0,585	1,052	0,017	0,637	0,013
$b_4$	0,800	0,779	0,638	0,852	0,016	1,502	0,015
$b_5$	-0,750	-0,801	1,584	-0,667	0,017	-1,601	0,015
$b_6$	1,250	-1,322	2,412	-1,129	0,018	1,800	0,016
$b_7$	1,350	1,334	0,509	1,402	0,018	-0,382	0,013
$b_8$	0,050	0,021	0,932	0,111	0,016	1,058	0,014
$b_9$	-0,250	-0,286	1,114	-0,185	0,016	0,862	0,013
$b_{10}$	-0,030	-0,059	0,977	0,031	0,016	-0,613	0,013
$b_{11}$	0,570	0,546	0,709	0,622	0,016	-1,052	0,014
$b_{12}$	1,450	1,433	0,490	1,504	0,018	1,402	0,015
$b_{13}$	-1,960	-2,121	5,435	-1,728	0,022	0,138	0,012
$b_{14}$	1,750	1,739	0,439	1,806	0,019	-0,148	0,012
$\mu$	0,95	0,933	0,496	0,983	0,001	-	-

Tabela B8: Modelo de Rasch Tradicional

Parâmetro	Estimativa	Erro padrão	Parâmetro	Estimativa	Erro padrão
$b_1$	-0,4876	0,006524	$b_{12}$	0,8013	0,006694
$b_2$	1,2424	0,007092	$b_{13}$	0,07681	0,00643
$b_3$	0,2682	0,006457	$b_{14}$	1,3998	0,007285
$b_4$	-0,5123	0,006535	$b_{15}$	1,1474	0,006989
$b_5$	-0,7496	0,00666	$b_{16}$	-0,6674	0,006611
$b_6$	0,4905	0,006525	$b_{17}$	0,01572	0,006428
$b_7$	0,9261	0,006786	$b_{18}$	0,9975	0,006846
$b_8$	1,2977	0,007156	$b_{19}$	-0,01262	0,006428
$b_9$	1,2862	0,007142	$b_{20}$	0,6518	0,006602
$b_{10}$	0,6538	0,006603	$b_{21}$	0,4251	0,006501
$b_{11}$	-0,5297	0,006542	$b_{22}$	1,2598	0,007112

Tabela B9: Modelo de Rasch - Trabalha

Parâmetro	Estimativa	Parâmetro	Estimativa
$b_1$	-0,02373	$b_{12}$	0,7517
$b_2$	1,3694	$b_{13}$	0,3351
$b_3$	0,4882	$b_{14}$	1,5109
$b_4$	-0,1415	$b_{15}$	0,9533
$b_5$	-0,2914	$b_{16}$	-0,2019
$b_6$	0,7357	$b_{17}$	0,3703
$b_7$	1,0617	$b_{18}$	1,0732
$b_8$	1,4891	$b_{19}$	0,3845
$b_9$	1,3056	$b_{20}$	1,0245
$b_{10}$	0,8452	$b_{21}$	0,6351
$b_{11}$	-0,1654	$b_{22}$	1,2722

Tabela B10: Modelo de Rasch - Não trabalha

Parâmetro	Estimativa	Parâmetro	Estimativa
$b_1$	-0,6096	$b_{12}$	0,7963
$b_2$	1,2107	$b_{13}$	0,002603
$b_3$	0,2119	$b_{14}$	1,3723
$b_4$	-0,6119	$b_{15}$	1,1736
$b_5$	-0,8679	$b_{16}$	-0,7975
$b_6$	0,4315	$b_{17}$	-0,0624
$b_7$	0,9035	$b_{18}$	0,9709
$b_8$	1,2692	$b_{19}$	-0,09993
$b_9$	1,292	$b_{20}$	0,585
$b_{10}$	0,6292	$b_{21}$	0,3883
$b_{11}$	-0,6041	$b_{22}$	1,2939

Tabela B11: Modelo de Rasch - Não reprovou

Parâmetro	Estimativa	Parâmetro	Estimativa
$b_1$	-0,7389	$b_{12}$	0,7418
$b_2$	1,1661	$b_{13}$	-0,1004
$b_3$	0,1466	$b_{14}$	1,3188
$b_4$	-0,7338	$b_{15}$	1,2407
$b_5$	1,0097	$b_{16}$	-0,9125
$b_6$	0,345	$b_{17}$	-0,1798
$b_7$	0,8345	$b_{18}$	0,9107
$b_8$	1,2075	$b_{19}$	-0,2059
$b_9$	1,2424	$b_{20}$	0,4417
$b_{10}$	0,5701	$b_{21}$	0,3013
$b_{11}$	-0,707	$b_{22}$	1,2485

Tabela B12: Modelo de Rasch - Já reprovou pelo menos uma vez

Parâmetro	Estimativa	Parâmetro	Estimativa
$b_1$	-0,1005	$b_{12}$	0,8845
$b_2$	1,3892	$b_{13}$	0,3706
$b_3$	0,4778	$b_{14}$	1,5621
$b_4$	-0,1564	$b_{15}$	0,9379
$b_5$	-0,3211	$b_{16}$	-0,2685
$b_6$	0,7812	$b_{17}$	0,4123
$b_7$	1,1307	$b_{18}$	1,1593
$b_8$	1,5265	$b_{19}$	0,3754
$b_9$	1,4194	$b_{20}$	1,1473
$b_{10}$	0,8654	$b_{21}$	0,7017
$b_{11}$	-0,1659	$b_{22}$	1,3762

Tabela B13: Modelo de Rasch - Branco(a)

Parâmetro	Estimativa	Parâmetro	Estimativa
$b_1$	-0,6466	$b_{12}$	0,6763
$b_2$	1,1492	$b_{13}$	-0,07438
$b_3$	0,1729	$b_{14}$	1,3063
$b_4$	-0,6426	$b_{15}$	1,1816
$b_5$	-0,8937	$b_{16}$	-0,7951
$b_6$	0,3726	$b_{17}$	-0,1112
$b_7$	0,8323	$b_{18}$	0,8907
$b_8$	1,1923	$b_{19}$	-0,1352
$b_9$	1,2094	$b_{20}$	0,4827
$b_{10}$	0,6057	$b_{21}$	0,3192
$b_{11}$	-0,6568	$b_{22}$	1,225

Tabela B14: Modelo de Rasch - Pardo(a)

Parâmetro	Estimativa	Parâmetro	Estimativa
$b_1$	-0,5409	$b_{12}$	0,8207
$b_2$	1,2508	$b_{13}$	0,04821
$b_3$	0,2427	$b_{14}$	1,4263
$b_4$	-0,5656	$b_{15}$	1,1614
$b_5$	-0,8113	$b_{16}$	-0,7482
$b_6$	0,4862	$b_{17}$	-0,00302
$b_7$	0,9411	$b_{18}$	1,0118
$b_8$	1,3226	$b_{19}$	-0,03946
$b_9$	1,3088	$b_{20}$	0,6768
$b_{10}$	0,6572	$b_{21}$	0,4387
$b_{11}$	-0,5409	$b_{22}$	1,3132

Tabela B15: Modelo de Rasch - Preto(a)

Parâmetro	Estimativa	Parâmetro	Estimativa
$b_1$	-0,2255	$b_{12}$	0,9403
$b_2$	1,3976	$b_{13}$	0,2863
$b_3$	0,41825	$b_{14}$	1,5294
$b_4$	-0,2919	$b_{15}$	0,9925
$b_5$	-0,4657	$b_{16}$	-0,3915
$b_6$	0,7068	$b_{17}$	0,3172
$b_7$	1,1011	$b_{18}$	1,1894
$b_8$	1,5033	$b_{19}$	0,2837
$b_9$	1,4566	$b_{20}$	1,0125
$b_{10}$	0,8142	$b_{21}$	0,6587
$b_{11}$	-0,2573	$b_{22}$	1,3534

Tabela B16: Modelo de Rasch - Amarelo(a) e Indígena

Parâmetro	Estimativa	Parâmetro	Estimativa
$b_1$	-0,4167	$b_{12}$	0,8143
$b_2$	1,2716	$b_{13}$	0,1389
$b_3$	0,3116	$b_{14}$	1,3905
$b_4$	-0,4667	$b_{15}$	1,1375
$b_5$	-0,682	$b_{16}$	-0,6113
$b_6$	0,5448	$b_{17}$	0,1026
$b_7$	0,9772	$b_{18}$	0,9821
$b_8$	1,4082	$b_{19}$	0,07241
$b_9$	1,3737	$b_{20}$	0,7813
$b_{10}$	0,7023	$b_{21}$	0,4905
$b_{11}$	-0,4105	$b_{22}$	1,3127



Tabela B17: Modelo de Rasch estendido - Ligação Probit utilizando a MVM

Parâmetro	Estimativa	Erro padrão	Parâmetro	Estimativa	Erro padrão
$b_1$	-0,4915	0,008441	$b_{13}$	0,07352	0,008788
$b_2$	1,2405	0,007783	$b_{14}$	1,3982	0,007957
$b_3$	0,2652	0,008552	$b_{15}$	1,1454	0,008089
$b_4$	-0,5163	0,008151	$b_{16}$	-0,6714	0,00712
$b_5$	-0,7537	0,005642	$b_{17}$	0,01236	0,008657
$b_6$	0,4877	0,008102	$b_{18}$	0,9954	0,007891
$b_7$	0,9239	0,007897	$b_{19}$	-0,01601	0,008602
$b_8$	1,2959	0,007843	$b_{20}$	0,6492	0,007853
$b_9$	1,2844	0,007883	$b_{21}$	0,4222	0,008433
$b_{10}$	0,6512	0,008073	$b_{22}$	0,258	0,00809
$b_{11}$	-0,5337	0,008496	$\mu$	0,9985	0,002651
$b_{12}$	0,7989	0,008018			

Tabela B18: Modelo de Rasch estendido - Ligação Logit utilizando a MVM

Parâmetro	Estimativa	Erro padrão	Parâmetro	Estimativa	Erro padrão
$b_1$	-0,462	0,007494	$b_{13}$	0,07002	0,006921
$b_2$	1,1928	0,007028	$b_{14}$	1,3529	0,008154
$b_3$	0,2501	0,006877	$b_{15}$	1,0976	0,007563
$b_4$	-0,4856	0,007251	$b_{16}$	-0,6338	0,007365
$b_5$	-0,7129	0,005714	$b_{17}$	0,01262	0,006633
$b_6$	0,4603	0,005556	$b_{18}$	0,9493	0,007797
$b_7$	0,8794	0,007755	$b_{19}$	-0,014	0,006273
$b_8$	1,2487	0,007017	$b_{20}$	0,6141	0,005867
$b_9$	1,237	0,007036	$b_{21}$	0,3982	0,005796
$b_{10}$	0,616	0,006193	$b_{22}$	1,2104	0,007033
$b_{11}$	-0,5022	0,00758	$\mu$	0,9989	0,001948
$b_{12}$	0,758	0,007424			

Tabela B19: Modelo de Rasch estendido - Ligação Probit Analítico

Parâmetro	Estimativa	Erro padrão	Parâmetro	Estimativa	Erro padrão
$b_1$	-1,008	0,02499	$b_{13}$	-0,1381	0,08621
$b_2$	1,6143	0,1184	$b_{14}$	1,8275	0,04988
$b_3$	0,1615	0,05217	$b_{15}$	1,4821	0,05375
$b_4$	-1,0451	0,07792	$b_{16}$	-1,2739	0,1308
$b_5$	-1,3927	0,1201	$b_{17}$	-0,2337	0,08913
$b_6$	0,5073	0,1367	$b_{18}$	1,2684	0,5657
$b_7$	1,1645	0,06984	$b_{19}$	-0,278	0,09053
$b_8$	1,6901	0,05807	$b_{20}$	0,7546	0,06491
$b_9$	1,6744	0,0515	$b_{21}$	0,4059	0,07205
$b_{10}$	0,7576	0,06485	$b_{22}$	1,6384	0,0519
$b_{11}$	-1,0711	0,1213	$\mu$	0,9269	0,02499
$b_{12}$	0,98	0,06096			

B20: Questionário de aluno - Parte 1

 Ministério da Educação  GOVERNO FEDERAL	<h1>CADERNO</h1>
<b>SAEB E PROVA BRASIL – 2007</b>	
PARA USO DO APLICADOR <input type="checkbox"/> ← Aluno ausente <input type="checkbox"/> ← Aluno presente e NÃO respondeu a prova <input type="checkbox"/> ← Aluno respondeu no Caderno de Prova, mas NÃO preencheu a Folha de Respostas <input type="checkbox"/> ← Aluno transcreveu corretamente as marcações do Caderno de Prova para a Folha de Respostas <input type="checkbox"/> ← ANE	6552494005 CÓDIGO DA TURMA
ESCREVA O SEU NOME COMPLETO	

## Folha de Respostas

4.ª Série (5.º Ano) do Ensino Fundamental

BLOCO 1					BLOCO 2					BLOCO 3					BLOCO 4				
1	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	12	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	1	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	12	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
2	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	13	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	2	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	13	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
3	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	14	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	3	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	14	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
4	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	15	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	4	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	15	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
5	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	16	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	5	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	16	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
6	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	17	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	6	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	17	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
7	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	18	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	7	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	18	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
8	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	19	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	8	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	19	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
9	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	20	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	9	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	20	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
10	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	21	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	10	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	21	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
11	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	22	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	11	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	22	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D

### QUESTIONÁRIO DO ALUNO (a ser respondido após a prova)

<b>1. Sexo:</b> <input type="checkbox"/> A Masculino. <input type="checkbox"/> B Feminino.	<b>7. Na sua casa tem videocassete ou DVD?</b> <input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.
<b>2. Como você se considera?</b> <input type="checkbox"/> A Branco(a). <input type="checkbox"/> C Preto(a). <input type="checkbox"/> E Indígena. <input type="checkbox"/> B Pardo(a). <input type="checkbox"/> D Amarelo(a).	<b>8. Na sua casa tem geladeira?</b> <input type="checkbox"/> A Sim, uma. <input type="checkbox"/> B Duas ou mais. <input type="checkbox"/> C Não tem.
<b>3. Qual é o mês do seu aniversário?</b> <input type="checkbox"/> A Janeiro. <input type="checkbox"/> E Maio. <input type="checkbox"/> I Setembro. <input type="checkbox"/> B Fevereiro. <input type="checkbox"/> F Junho. <input type="checkbox"/> J Outubro. <input type="checkbox"/> C Março. <input type="checkbox"/> G Julho. <input type="checkbox"/> K Novembro. <input type="checkbox"/> D Abril. <input type="checkbox"/> H Agosto. <input type="checkbox"/> L Dezembro.	<b>9. Na sua casa tem freezer separado da geladeira?</b> <input type="checkbox"/> A Sim. <input type="checkbox"/> B Não. <input type="checkbox"/> C Não sei.
<b>4. Qual é a sua idade?</b> <input type="checkbox"/> A 8 anos ou menos. <input type="checkbox"/> D 11 anos. <input type="checkbox"/> G 14 anos. <input type="checkbox"/> B 9 anos. <input type="checkbox"/> E 12 anos. <input type="checkbox"/> H 15 anos ou mais. <input type="checkbox"/> C 10 anos. <input type="checkbox"/> F 13 anos.	<b>10. Na sua casa tem uma máquina de lavar roupa? (Não é tanquinho)</b> <input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.
<b>5. Na sua casa tem televisão em cores?</b> <input type="checkbox"/> A Sim, uma. <input type="checkbox"/> B Sim, duas. <input type="checkbox"/> C Sim, três ou mais. <input type="checkbox"/> D Não tem.	<b>11. Na sua casa tem aspirador de pó?</b> <input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.
<b>6. Na sua casa tem rádio?</b> <input type="checkbox"/> A Sim, um. <input type="checkbox"/> B Sim, dois. <input type="checkbox"/> C Sim, três ou mais. <input type="checkbox"/> D Não tem.	<b>12. Na sua casa tem carro?</b> <input type="checkbox"/> A Sim, um. <input type="checkbox"/> B Sim, dois. <input type="checkbox"/> C Sim, três ou mais. <input type="checkbox"/> D Não.
	<b>13. Na sua casa tem um computador?</b> <input type="checkbox"/> A Sim, com internet. <input type="checkbox"/> B Sim, sem internet. <input type="checkbox"/> C Não.
	<b>14. Na sua casa tem banheiro?</b> <input type="checkbox"/> A Sim, um. <input type="checkbox"/> C Sim, três. <input type="checkbox"/> E Não. <input type="checkbox"/> B Sim, dois. <input type="checkbox"/> D Sim, mais de três.

## B21: Questionário de aluno - Parte 2

<p><b>15. Na sua casa trabalha alguma empregada doméstica?</b></p> <p><input type="checkbox"/> A Sim, uma diarista, uma ou duas vezes por semana.</p> <p><input type="checkbox"/> B Sim, uma, todos os dias úteis.</p> <p><input type="checkbox"/> C Sim, duas ou mais, todos os dias úteis.</p> <p><input type="checkbox"/> D Não.</p>	<p><b>28. Seus pais ou responsáveis incentivam você a fazer o dever de casa e os trabalhos da escola?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>
<p><b>16. Na sua casa tem quartos para dormir?</b></p> <p><input type="checkbox"/> A Sim, um. <input type="checkbox"/> C Sim, três. <input type="checkbox"/> E Não.</p> <p><input type="checkbox"/> B Sim, dois. <input type="checkbox"/> D Sim, quatro ou mais.</p>	<p><b>29. Seus pais ou responsáveis incentivam você a ler?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>
<p><b>17. Quantas pessoas moram com você?</b></p> <p><input type="checkbox"/> A Moro sozinho(a) ou com mais 1 pessoa.</p> <p><input type="checkbox"/> B Moro com mais 2 pessoas.</p> <p><input type="checkbox"/> C Moro com mais 3 pessoas.</p> <p><input type="checkbox"/> D Moro com mais 4 ou 5 pessoas.</p> <p><input type="checkbox"/> E Moro com mais 6 a 8 pessoas.</p> <p><input type="checkbox"/> F Moro com mais do que 8 pessoas.</p>	<p><b>30. Seus pais ou responsáveis incentivam você a ir à escola e não faltar às aulas?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>
<p><b>18. Você mora com sua mãe?</b></p> <p><input type="checkbox"/> A Sim.</p> <p><input type="checkbox"/> B Não.</p> <p><input type="checkbox"/> C Não. Moro com outra mulher responsável por mim.</p>	<p><b>31. Seus pais ou responsáveis conversam com você sobre o que acontece na escola?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>
<p><b>19. Até que série sua mãe ou a mulher responsável por você estudou?</b></p> <p><input type="checkbox"/> A Nunca estudou ou não completou a 4.ª série (antigo primário).</p> <p><input type="checkbox"/> B Completou a 4.ª série, mas não completou a 8.ª série (antigo ginásio).</p> <p><input type="checkbox"/> C Completou a 8.ª série, mas não completou o Ensino Médio (antigo 2.º grau).</p> <p><input type="checkbox"/> D Completou o Ensino Médio, mas não completou a Faculdade.</p> <p><input type="checkbox"/> E Completou a Faculdade.</p> <p><input type="checkbox"/> F Não sei.</p>	<p><b>32. Além dos livros escolares, quantos livros têm em sua casa?</b></p> <p><input type="checkbox"/> A O bastante para encher uma prateleira (1 a 20 livros).</p> <p><input type="checkbox"/> B O bastante para encher uma estante (21 a 100 livros).</p> <p><input type="checkbox"/> C O bastante para encher várias estantes (mais de 100 livros).</p> <p><input type="checkbox"/> D Nenhum.</p>
<p><b>20. Sua mãe ou a mulher responsável por você sabe ler e escrever?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>	<p><b>33. Em dia de aula, quanto tempo você gasta assistindo TV?</b></p> <p><input type="checkbox"/> A 1 hora ou menos. <input type="checkbox"/> C 3 horas. <input type="checkbox"/> E Não assisto TV.</p> <p><input type="checkbox"/> B 2 horas. <input type="checkbox"/> D 4 horas ou mais.</p>
<p><b>21. Você vê sua mãe ou a mulher responsável por você lendo?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>	<p><b>34. Em dia de aula, quanto tempo você gasta fazendo trabalhos domésticos em casa?</b></p> <p><input type="checkbox"/> A 1 hora ou menos. <input type="checkbox"/> D 4 horas ou mais.</p> <p><input type="checkbox"/> B 2 horas. <input type="checkbox"/> E Não faço trabalhos domésticos.</p> <p><input type="checkbox"/> C 3 horas.</p>
<p><b>22. Você mora com seu pai?</b></p> <p><input type="checkbox"/> A Sim.</p> <p><input type="checkbox"/> B Não.</p> <p><input type="checkbox"/> C Não. Moro com outro homem responsável por mim.</p>	<p><b>35. Você trabalha fora de casa?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>
<p><b>23. Até que série seu pai ou o homem responsável por você estudou?</b></p> <p><input type="checkbox"/> A Nunca estudou ou não completou a 4.ª série (antigo primário).</p> <p><input type="checkbox"/> B Completou a 4.ª série, mas não completou a 8.ª série (antigo ginásio).</p> <p><input type="checkbox"/> C Completou a 8.ª série, mas não completou o Ensino Médio (antigo 2.º grau).</p> <p><input type="checkbox"/> D Completou o Ensino Médio, mas não completou a Faculdade.</p> <p><input type="checkbox"/> E Completou a Faculdade.</p> <p><input type="checkbox"/> F Não sei.</p>	<p><b>36. Quando você entrou na escola?</b></p> <p><input type="checkbox"/> A No maternal (jardim de infância). <input type="checkbox"/> C Na primeira série.</p> <p><input type="checkbox"/> B Na pré-escola. <input type="checkbox"/> D Depois da primeira série.</p>
<p><b>24. Seu pai ou o homem responsável por você sabe ler e escrever?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>	<p><b>37. Desde a primeira série você estudou sempre nesta mesma escola?</b></p> <p><input type="checkbox"/> A Sim.</p> <p><input type="checkbox"/> B Não, mas só estudei em escola pública.</p> <p><input type="checkbox"/> C Não, mas já estudei em escola particular.</p>
<p><b>25. Você vê seu pai ou o homem responsável por você lendo?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>	<p><b>38. Você já foi reprovado?</b></p> <p><input type="checkbox"/> A Não. <input type="checkbox"/> B Sim, uma vez. <input type="checkbox"/> C Sim, duas vezes ou mais.</p>
<p><b>26. Com que frequência seus pais ou responsáveis vão à reunião de pais?</b></p> <p><input type="checkbox"/> A Sempre ou quase sempre. <input type="checkbox"/> C Nunca ou quase nunca.</p> <p><input type="checkbox"/> B De vez em quando.</p>	<p><b>39. Você já abandonou a escola durante o período de aulas e ficou fora da escola o resto do ano?</b></p> <p><input type="checkbox"/> A Não. <input type="checkbox"/> B Sim, uma vez. <input type="checkbox"/> C Sim, duas vezes ou mais.</p>
<p><b>27. Seus pais ou responsáveis incentivam você a estudar?</b></p> <p><input type="checkbox"/> A Sim. <input type="checkbox"/> B Não.</p>	<p><b>40. Você faz o dever de casa de língua portuguesa?</b></p> <p><input type="checkbox"/> A Sempre ou quase sempre. <input type="checkbox"/> C Nunca ou quase nunca.</p> <p><input type="checkbox"/> B De vez em quando. <input type="checkbox"/> D O professor não passa dever de casa.</p>
	<p><b>41. O professor corrige o dever de casa de língua portuguesa?</b></p> <p><input type="checkbox"/> A Sempre ou quase sempre. <input type="checkbox"/> C Nunca ou quase nunca.</p> <p><input type="checkbox"/> B De vez em quando.</p>
	<p><b>42. Você faz o dever de casa de matemática?</b></p> <p><input type="checkbox"/> A Sempre ou quase sempre. <input type="checkbox"/> C Nunca ou quase nunca.</p> <p><input type="checkbox"/> B De vez em quando. <input type="checkbox"/> D O professor não passa dever de casa.</p>
	<p><b>43. O professor corrige o dever de casa de matemática?</b></p> <p><input type="checkbox"/> A Sempre ou quase sempre. <input type="checkbox"/> C Nunca ou quase nunca.</p> <p><input type="checkbox"/> B De vez em quando.</p>
	<p><b>44. Seus professores elogiam ou dão parabéns quando você tira boas notas?</b></p> <p><input type="checkbox"/> A Sempre ou quase sempre. <input type="checkbox"/> C Nunca ou quase nunca.</p> <p><input type="checkbox"/> B De vez em quando.</p>

4134493693

# Anexo C - Problemas computacionais e códigos utilizados

Nesta seção serão listados alguns problemas computacionais enfrentados na aplicação do modelo de Rasch estendido e as alternativas que foram adotadas para poder realizar os ajustes. Destaca-se entre todos estes problemas o problema da estimação dos parâmetros  $\alpha_1$  e  $\alpha_2$  do modelo. Uma série de tentativas de estimação destes parâmetros foram realizadas e frustadas, atrapalhando substancialmente o desenvolvimento do trabalho. Por fim, optou-se por utilizar outra parametrização, a parametrização de Ferrari & Cribari-Neto (2004). Abaixo seguem os tópicos críticos e suas soluções encontradas.

## Estrutura de variância dos parâmetros de dificuldade dos itens

De acordo com a TRI, estes parâmetros tem distribuição Normal com média “0” e variância igual a “1”. Logo, nos primeiros programas testados tentou-se não fixar a variância destes parâmetros, mas estimá-la, sentindo a necessidade de verificar se esta suposição do modelo estava sendo respeitada. No entanto, após inúmeras tentativas frustradas de estimação, decidiu-se simplificar o modelo supondo que esta suposição estava sendo atendida, deixando o ajuste mais rápido.

## Parametrização da distribuição Beta

Para os modelos de Rasch estendidos apresentados no capítulo 4, o parâmetro aleatório  $\gamma$ , que captaria o efeito de superdispersão, inicialmente teria distribuição Beta ( $\alpha_1, \alpha_2$ ). No entanto, para realizar estes ajustes era necessário que  $\alpha$  e  $\beta$  fossem relacionados de alguma forma. Vieira(2008) utiliza em seu trabalho a relações  $\alpha_1 +$

$\alpha_2 = c$  e  $\frac{\alpha_1}{\alpha_2} = c$  para estimar estes parâmetros. Primeiramente, utilizou-se a relação  $\alpha_1 + \alpha_2 = c$  para o ajuste do modelo. No entanto, ao agregar a constante  $c$  que linkaria os valores de  $\alpha_1$  e  $\alpha_2$  mais uma incerteza foi inserida no modelo. Diante deste problema resolveu-se testar diversos valores de  $c$  e uma série de valores iniciais de  $\alpha_1$ , sendo o parâmetro  $\alpha_2$  do modelo substituído por  $c - \alpha_1$ .

Logo, diversas combinações de valores para  $c$  e  $\alpha_1$  foram testadas para ajustar estes modelos. Entretanto, o ajuste destes bancos de dados levavam muito tempo e não eram satisfatórios. Neste momento as integrações numéricas eram realizadas pelo método da Quadratura Gaussiana (QG), adaptativa e utilizando 30 pontos. Diante desta grande demora diversos bancos de dados menores foram amostrados do banco original. Estes bancos menores dariam mais dinamismo ao trabalho e mais possibilidades de modelagem em menos tempo. Caso um resultado fosse encontrado neste pequenos bancos, este resultado seria posteriormente testado no banco original.

Depois de todos estes cuidados serem tomados e todas estas combinações de modelos serem modeladas os resultados encontrados não foram satisfatórios. Na grande maioria das vezes as estimativas dos parâmetros de dificuldade eram justamente os limites superiores estabelecidos na programação, -5 e 5. Em outros casos os valores estimados para  $\alpha_1$  eram muito grandes e consequentemente as estimativas dos  $b$ 's também eram absurdas.

No entanto, depois de muitas tentativas frustradas, o modelo foi reescrito utilizando a parametrização de Ferrari & Cribari-Neto. Ao fazer as contas observou-se que a função de máxima verossimilhança só dependia do parâmetro  $\mu$ , não precisando assim de estimar o parâmetro  $\phi$ .

## Integração Numérica

O método escolhido para realizar as integrações numéricas existentes nos modelos de Rasch estendido utilizando função de Máxima Verossimilhança Marginal foi a Quadratura Gaussiana (QG). Inicialmente optou-se pela QG adaptativa com 30 pontos, no entanto esta combinação faz com que o ajuste do modelo seja muito demorado. Após isto, optou-se utilizar a QG não-adaptativa para o ajuste do modelo de Rasch estendido. Diante da demora também apresentada, a QG não-adaptativa

foi executada com um ponto apenas, ou seja, utilizando o método de aproximação de Laplace.

## Programação utilizada

C1: Programação escrita para manipular os dados da Prova Brasil

```
data ts_ resposta;  
  
infile "C: Users EduardaFilipe Desktop Prova Brasil 2007 DADOS TS_ resposta.TXT"  
  
missover lrecl=246347;  
  
input  
  
@1 ID_ ALUNO 8. /* Código do Aluno na Prova Brasil */  
  
@9 ID_ SERIE 1. /* Código da Série */  
  
@10 DS_ DISCIPLINA . /* Código da Disciplina */  
  
@11 NU_ COD_ BNI 8. /* Código do item no BNI, fornecido pelo INEP */  
  
@19 NU_ ITEM_ BLOCO 2. /* Número do item dentro do bloco */  
  
@21 ID_ CADERNO 8. /* Código do Caderno */  
  
@29 NU_ BLOCO 1. /* Numero do Bloco */  
  
@30 RESPOSTA_ ALUNO /* Resposta do Aluno */  
  
@31 GABARITO_ ITEM /* Gabarito */  
  
@32 COD_ CORRECAO /* Correção de determinado item */  
  
;run;  
  
quit;  
  
  
data db.Ts_ resposta;  
  
set ts_ resposta;  
  
run;  
  
  
data db.Ts_ resposta4;  
  
set db.Ts_ resposta;  
  
if ID_ SERIE=8 then delete;  
  
run;
```

```

data db.Ts_ resposta4Mspec;

set db.Ts_ resposta4;

if DS_ Disciplina='L' then delete;

keep Id_ aluno Id_ serie Ds_ disciplina ID_ CADERNO NU_ ITEM_ BLOCO;

run;

/*Removendo duplicações*/

proc sort data=db.Ts_ resposta4Mspec nodupkey;

by Id_ aluno;

run;

/*Colocando as respostas dos questionários em ordem crescente*/

proc sort data=db.TS_ quest_ aluno4;

by Id_ aluno;

run;

/*Juntando com o arquivo de respostas do questionário do aluno*/

data db.Ts_ Qalun4M;

merge db.Ts_ resposta4Mspec db.TS_ quest_ aluno4;

by Id_ aluno;

run;

/*Tirando as observações do alunos que não fizeram a prova do caderno 1*/

data db.Ts_ Qalun4M;

set db.Ts_ Qalun4M;

if ID_ CADERNO  $\neq$  20071 then delete;

run;

/*Para as análises serão usadas as respostas dos alunos que receberam o caderno 1 (ID_C ADERNO=20071)*/

data db.Ts_ resposta4cad1;

set db.Ts_ resposta4;

```



```

if DS_ Disciplina='L' or ID_ CADERNO = 20071 then delete;

run;

proc sort data=db.Ts_ resposta4cad1;

by Id_ aluno;

run;

data db.mt42007;

merge db.Ts_ resposta4cad1 db.Ts_ Qalun4M;

by Id_ aluno;

run;

data db.mt42007;

set db.mt42007;

if NU_ bloco=1 then Questao=NU_ item_ bloco;

else Questao=NU_ item_ bloco+11;

run;

data db.mt42007;

set db.mt42007;

if Cod_ Correcao='0' then resp=0;

else resp=1;

run;

proc sort data=db.mt42007;

by Id_ aluno NU_ bloco NU_ item_ bloco;

run;

/* Dados só com as respostas e as suas respectivas questões*/

data db.mt42007a;

set db.mt42007;

per=_ N_;

keep Id_ aluno QUESTAO COD_ CORRECAO per resp;run;

```

```

data db.mt42007b;

set db.mt42007a;

arrat items22 i1-i22;

do j=1 TO 22;

items{ j }=0;

end;

items{QUESTAO}=1;

drop i j score1;

run;

```

C2: Programação básica utilizada para realizar os ajustes adotando a parametrização tradicional da beta para o parâmetro aleatório  $\gamma$ :

```

proc nlmixed data=db.mt42007b method=gauss technique=newrap noad qpoints=20;

parms b1-b39=1 alpha=3.6;

bounds alpha_1<10 -5 < b1-b22 <5; const=10;

beta=i1*b1+i2*b2+i3*b3+i4*b4+i5*b5+i6*b6+i7*b7+i8*b8+i9*b9+

i10*b10+i11*b11+i12*b12+i13*b13+i14*b14+i15*b15+i16*b16+i17*b17+i18*b18+i19*b19+

i20*b20+i21*b21+i22*b22;

ex=exp(theta-beta);

p=ex/(1+ex);

ll=-log(const)+resp*log(alpha_1*p)+(1-resp)*log((1-p)*const);

model resp ~ general(ll);

random theta ~ normal(0,1) subject=Id_ aluno ;

contrast 'alpha_2 ' const-alpha_1;

run;

quit;

```

C3: Macro utilizada para gerar os dados da simulação:

```
data=nome de banco

n=número de observações

t=número de itens

a=valor de  $\mu$ 

b=valor de  $\phi$ 


% macro tovo(data,n,t,a,b);

data un;

array dife(&t.) d1-d&t.;

%do i=1 %to &t.;

dife[&i]=rand('Normal',0,1)

; %End;

%do i=1 %to &n.;

per=&i.;

output;

%end;

run;


data &data.;

set un;

run;

data &data.;

set &data.;

theta=rand('Normal',0,1);

gamma=rand('beta',&a.*&b.,(1-&a.)&b.);

run;
```

```

data &data.;

set &data.;

array item(&t.) a1-a&t.;

%do i=1 %to &t.;

item[&i]=rand('bernoulli',gamma*probnorm(theta-d&i.));

%End;

run;

```

```

data &data.;

set &data.;

%do i=1 %to &t.;

i&i.=a&i.;

%End;

drop a1-a&t.;

run;

%mend tovo;

```

C4: Exemplo da macro utilizada para ajustar os modelos na simulação:

```

%macro tovo1(data,obs,in);

%do s=1 %to &obs.;

title ' MODELO PROBIT COM FUNCAO MVM ANALITICA - valor real 0.75';

proc nlmixed data=&data.&s. technique=quanew maxiter=500 alpha=0.01;

parms mu=&in. b1-b14=0;

bounds 0< mu < 1;

beta=i1*b1+i2*b2+i3*b3+i4*b4+i5*b5+i6*b6+i7*b7+i8*b8+i9*b9+

i10*b10+i11*b11+i12*b12+i13*b13+i14*b14; /*+i15*b15+i16*b16+i17*b17+i18*b18+i19*b19+

i20*b20+i21*b21+i22*b22;

eta = -beta;

pb=cdf('NORMAL',eta,0,sqrt(2));

ll = resp*log(mu) + resp*log(pb) + (1-resp)*log(1-mu*pb);

model resp general(ll);

```

```
ods output ParameterEstimates=pe&s.;  
  
run;  
  
data pe&s.;  
  
set pe&s.;  
  
amostra=&s.;  
  
run;  
  
%end;
```