



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

Um método bayesiano para verificação de ajuste do modelo logístico de três parâmetros em teoria da resposta ao item

por

Rodrigo Marques dos Santos

Brasília, 07 de Fevereiro de 2023

Um método bayesiano para verificação de ajuste do modelo logístico de três parâmetros em teoria da resposta ao item

por

Rodrigo Marques dos Santos

Dissertação apresentada ao Departamento de
Estatística da Universidade de Brasília, como
requisito parcial para obtenção do título de
Mestre em Estatística.

Orientador: Prof. Dr. Antonio Eduardo Go-
mes

Brasília, 06 de Abril de 2022

Dissertação submetida ao Programa de Pós-Graduação em Estatística do Departamento de Estatística da Universidade de Brasília como parte dos requisitos para a obtenção do título de Mestre em Estatística.

Texto aprovado por:

Prof. Antônio Eduardo Gomes
Orientador, EST/UnB

Prof. Raul Yukihiro Matsushita
EST/UnB

Prof. André Luiz Fernandes Cançado
EST/UnB

Prof. Dalton Francisco de Andrade
INE-CTC/UFSC

A randomização é importante demais para ser deixada ao acaso.

(Joseph D. Petrucci)

Para todas as pessoas que me permitem ser quem eu sou de verdade.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

A Teoria de Resposta ao Item tem sido cada vez mais utilizada em estudos que tem como objetivo estimar o traço latente e, dentre os modelos existentes, tem-se que os logísticos são uns dos mais utilizados. Porém, cada vez mais estudos mostram que o pressuposto de que as Curvas Características dos Itens (CCI's) seguem a forma Logística não são válidos, tornando a checagem desse pressuposto cada vez mais importante. Por esse motivo, estimar a CCI de formas alternativas e não paramétricas pode ser uma poderosa ferramenta para comparar com a CCI gerada pelo modelo logístico e, assim, permitir inferência sobre a veracidade desse pressuposto. Esse estudo propõe um teste não paramétrico que se utiliza de inferência Bayesiana, mais especificamente o método de Posterior Predictive Model Checking (PPMC) para testar essa hipótese. Para comparar com a CCI calculada pelo Modelo Logístico, foram utilizadas as regressões Isotônica e de Nadaraya-Watson para criar 6 estatísticas do teste. Foram feitas duas análises, uma utilizando uma simulação e outra aplicando esse teste a dados reais de uma aplicação do SARESP. Os resultados da simulação foram satisfatórios, com o teste indicando diferenças significativas em pouquíssimos itens que de fato seguiam o Modelo Logístico de 3 parâmetros, e conseguindo reconhecer bem os itens que tinham CCI's não monotônicos. Apesar disso, o teste reconheceu apenas um item que era uma mistura de distribuições. Para os dados reais, os estimadores de Regressão Isotônica indicaram valores diferentes dos que foram indicados pela Regressão de Nadaraya-Watson, em sua maioria.

Palavras-chave:

Teoria de Resposta ao Item; Teste de Adequabilidade; Regressão Isotônica; Inferência Bayesiana; Markov Chain Monte Carlo (MCMC); Posterior Predictive Model Checking (PPMC); Regressão de Nadaraya-Watson;

Abstract

The Item Response Theory has been increasingly used in studies that aim to estimate the latent trait and, among the existing models, the logistic ones are the most used. However, more and more studies show that the assumption that Item Characteristic Curves (ICC's) follow the Logistic form are not valid, making it increasingly important to check this assumption. Therefore, estimating the ICC in alternative, nonparametric ways can be a powerful tool to compare with the ICC generated by the logistic model and thus allow inference about the veracity of this assumption. This study proposes a nonparametric test that uses Bayesian inference, more specifically the Posterior Predictive Model Checking (PPMC) method to test this hypothesis. To compare with the ICC calculated by the Logistic Model, Isotonic and Nadaraya-Watson regressions were used to create 6 test statistics. Two analyses were done, one using a simulated data set and the other applying this test to real data from a SARESP application. The simulation results were satisfactory, with the test indicating significant differences in very few items that actually followed the 3-parameter Logistic Model, and managing to recognize well those items that had a non-monotonic ICC. Despite this, the test recognized only one item that were mixtures of distributions. For the real data, the Isotonic Regression estimators indicated different values than those indicated by the Nadaraya-Watson Regression, for the most part of items.

Key words:

Item Response Theory; Test of Adequacy; Isotonic Regression; Bayesian Inference;
Markov Chain Monte Carlo (MCMC); Posterior Predictive Model Checking (PPMC);
Nadaraya-Watson Regression;

Sumário

1	Introdução	1
2	Teoria da Resposta ao Item	3
2.1	Modelos Logísticos	3
2.1.1	Curva Característica do Item	4
2.1.2	Função de Informação do Item	6
2.2	Estimação não Paramétrica da CCI	8
2.2.1	Regressão Isotônica	8
2.2.2	Regressão Isotônica Suavizada	11
2.3	Estatísticas para Adequabilidade de Ajuste	12
3	Estimação Bayesiana	15
3.1	Teoria Bayesiana	15
3.2	Método de Monte Carlo via Cadeia de Markov (MCMC)	17
3.3	Algoritmo Metropolis-Hastings	18
4	Posterior Predictive Model Checking (PPMC)	19
5	Simulação e Aplicação a Dados Reais	22
5.1	Aplicação a dados simulados	23
5.2	Aplicação a dados reais	40

6	Conclusão e Trabalhos Futuros	46
6.1	Conclusão	46
6.2	Trabalhos Futuros	47
7	Referências Bibliográficas	48
A	Código em R	51

Lista de Tabelas

5.1	Parâmetros reais para cada item ML3.	24
5.2	Parâmetros dos 4 itens de mistura.	24
5.3	Parâmetros dos 4 itens não monotônicos.	25
5.4	Parâmetros estimados para os 56 itens da prova simulada.	35
5.5	Valores do PPP para os 56 itens em cada uma das 6 estatísticas do teste na prova simulada.	40
5.6	Valores do PPP para os 56 itens em cada uma das 6 estatísticas do Teste na prova do SARESP.	42
5.7	Parâmetros estimados dos 30 itens da Prova SARESP.	45

Lista de Figuras

2.1	Exemplo de uma Curva Característica do Item geral explicando os valores dos parâmetros no modelo (Andrade et al., 2000).	5
2.2	Exemplos de curvas características e de informação para quatro itens diferentes (Andrade et al., 2000).	7
2.3	Exemplo de um DSA com os minorantes calculados.	10
5.1	CCI's do Item 49.	25
5.2	CCI's do Item 50.	26
5.3	CCI's do Item 51.	27
5.4	CCI's do Item 52.	28
5.5	CCI's do Item 53.	29
5.6	CCI's do Item 54.	30
5.7	CCI's do Item 55.	31
5.8	CCI's do Item 56.	32
5.9	Item 6.	41
5.10	CCI's do Item 5.	43
5.11	CCI's do Item 15.	44

Abreviações e Siglas

AAS	Amostra Aleatória Simples
AASc	Amostra Aleatória Simples com reposição
AASs	Amostra Aleatória Simples sem reposição
AE	Amostra Estratificada
AEpr	Amostra Estratificada proporcional
AEot	Amostra Estratificada ótima
TSL	Linearização por Série de <i>Taylor</i>
BRR	<i>Balanced Repeated Replication</i>
JRR	<i>Jackknife Repeated Replication</i>
RDD	<i>Randon Digit Dialing</i>
H_0	Hipótese nula
H_1	Hipótese alternativa
Var	Variância
ns	Nível de Significância
cv	Coefficiente de Variação
IBGE	Instituto Brasileiro de Geografia e Estatística
PNADc	Pesquisa Nacional de Amostra por Domicílio contínua

Capítulo 1

Introdução

A Teoria da Resposta ao Item (TRI) compreende um conjunto de técnicas estatísticas que é uma das mais utilizadas para estimar traços latentes a partir de instrumentos de coletas de dados, principalmente os questionários. Essa capacidade é de extrema importância, visto que os traços latentes são características que não podem ser medidas de nenhuma forma direta, sendo, assim, possível apenas obter uma estimativa dela.

Existem algumas formas de se estimar esse traço latente, como Análise Fatorial Confirmatória, ou a Teoria Clássica dos Testes (TCT), por exemplo. Porém, diferentemente da TCT, que tem foco na prova inteira e no resultado obtido nela, a TRI foca no item, gerando um modelo e, assim, uma probabilidade de acerto para cada item, tornando possíveis ajustes individuais em itens da prova e uma maior facilidade em equilibrar uma prova em itens mais fáceis e itens mais difíceis, tornando o teste o mais discriminante possível.

Por esse motivo, a TRI hoje é a principal ferramenta para auxiliar o ensino brasileiro, estando presente na elaboração das mais importantes provas brasileiras, como o Exame Nacional do Ensino Médio (ENEM) e o Sistema de Avaliação de Educação Básica (Saeb), por exemplo. A TRI traz ferramentas que são capazes de modelar a probabilidade de acerto de um item usando parâmetros como discriminação, proficiência, probabilidade de

acerto ao acaso e dificuldade do item.

Porém, assim como a grande maioria das técnicas estatísticas, essas modelagens tem alguns pressupostos. Dois deles são a distribuição da curva característica do item ser logística e, portanto, não decrescente, além de a proficiência do indivíduo seguir uma distribuição normal, entre outras suposições, como independência local e unidimensionalidade do instrumento.

Portanto, é importante que seja possível checar se o ajuste da modelagem está satisfatória, um excelente indicador de que os pressupostos estão sendo atendidos e que a probabilidade de acerto obtida com o modelo é uma boa aproximação do valor real.

Assim, esse trabalho tem como objetivo principal desenvolver um teste de adequabilidade do modelo de três parâmetros que a TRI ajusta em um conjunto de dados e, para isso, tem como objetivos secundários a simulação dos parâmetros de forma bayesiana, um processo de reamostragem a partir desses valores gerados e, por fim, o cálculo de uma estatística do Teste.

Capítulo 2

Teoria da Resposta ao Item

2.1 Modelos Logísticos

A Teoria da Resposta ao Item, ou TRI, complementa a Teoria Clássica dos Testes (TCT) desenvolvida em meados dos anos 50. A TCT tem como foco analisar os escores totais dos respondentes, ou seja, o somatório dos acertos que um respondente obteve em um teste. Esse tipo de abordagem traz consigo alguns problemas, como, por exemplo, a incapacidade de comparação do traço latente entre uma pessoa que fez uma prova e outra que fez outra prova, visto que, ao analisar o escore total de todos os itens, as provas se tornam independentes uma da outra. (Andrade et al., 2000).

Pensando em corrigir esses problemas, a TRI analisa cada item separadamente, criando modelos que permitem que o item estime o traço latente de forma individual. Essa construção permite que o teste seja construído como um conjunto de itens independentes, equilibrando a exigência de diferentes respondentes e possibilitando, assim, a comparação entre testes.

Para isso, a TRI conta com 1 modelo principal com dois casos particulares, o modelo logístico de 3 parâmetros. Esse modelo tem três parâmetros: o de discriminação, chamado de parâmetro a_i , o de dificuldade, chamado de parâmetro b_i e o parâmetro de probabilidade

de acerto ao acaso, chamado de parâmetro c_i . Este modelo, que será utilizado neste trabalho, foi proposto por Birnbaum (1968) e, quando os itens são dicotômicos ou foram dicotomizados a probabilidade de acerto do item i pelo indivíduo j é dada por

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + \exp(-Da_i(\theta_j - b_i))} \quad (2.1)$$

sendo $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, n$, I a quantidade de itens e n o tamanho da amostra, ou seja, o número de respondentes. Além disso, D é o parâmetro de escala, geralmente constante com valor $D = 1$ caso se queira aproximar a distribuição obtida para uma função ogiva normal, usa-se $D = 1,702$, e θ_j é o traço latente associado à j -ésima pessoa (Baker, 1985).

Esse modelo se transforma no modelo logístico de 2 parâmetros quando $c_i = 0$ e, mantido isso constante, se transforma no modelo de um parâmetro quando $a_i = 1$.

Também é importante citar que a_i tem o valor proporcional à inclinação da Curva Característica do Item (CCI).

2.1.1 Curva Característica do Item

A Curva Característica do Item, ou CCI, é uma função que representa a relação da probabilidade de acerto, obtida por um dos 3 modelos já discutidos anteriormente, e o valor do traço latente. Um exemplo geral da CCI pode ser visto na Figura 2.1.

Um ponto importante sobre a CCI é que ela é sempre monotonicamente crescente, visto que é sempre esperado que a correlação entre a probabilidade de acerto e o traço latente seja positiva e não linear, ou seja, quanto maior o θ , maior a probabilidade do indivíduo acertar o item (Andrade et al., 2000). Além disso, também é possível deduzir os valores estimados dos parâmetros a_i , b_i e c_i apenas olhando para a curva, como é mostrado na figura.

Com relação aos limites dos valores que os parâmetros a_i , b_i e c_i podem assumir, tem-

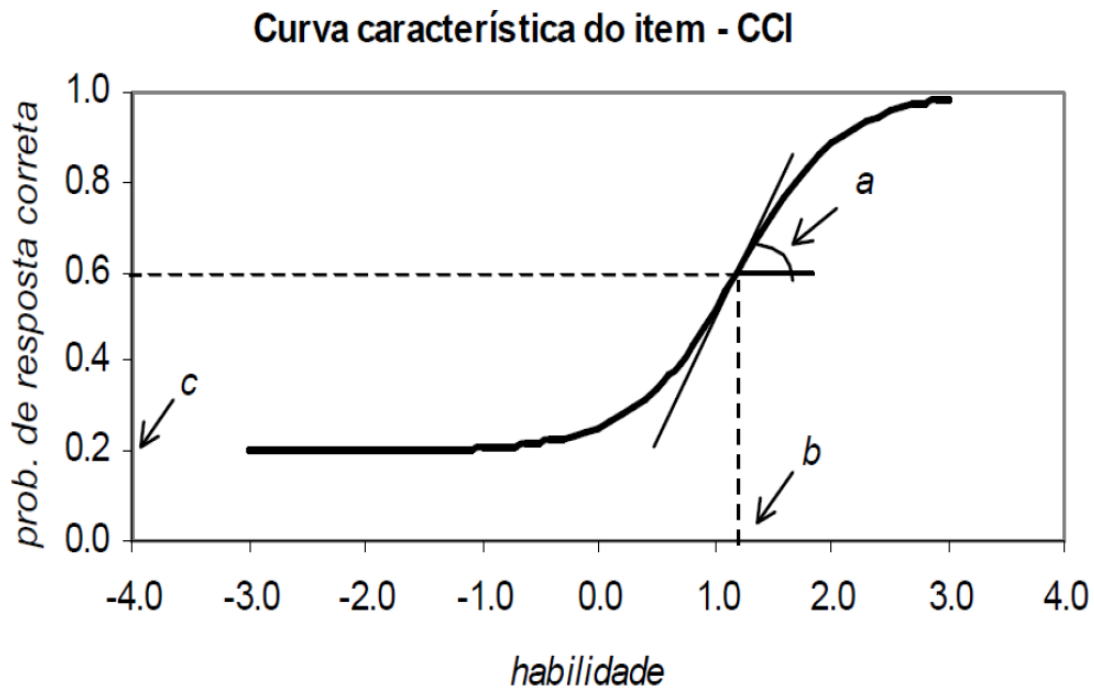


Figura 2.1: Exemplo de uma Curva Característica do Item geral explicando os valores dos parâmetros no modelo (Andrade et al., 2000).

se que o parâmetro c varia de 0 a 1, por se tratar de uma probabilidade, e o parâmetro a_i tem que ser positivo, pois, como visto na figura 1 e já comentado antes, ele é proporcional à inclinação da CCI no ponto $\theta = b_i$

Também é importante entender que a escala de θ é arbitrária, visto que um traço latente, por definição, não pode ser medido, ou seja, não possui um valor real. Isso faz com que os valores possíveis para o b_i também tenham uma escala arbitrária, que posteriormente será a mesma escala do θ . Apesar disso, a maioria dos itens tem uma CCI na qual o θ varia em torno do zero por facilidade computacional.

2.1.2 Função de Informação do Item

Outra medida importante para o estudo exploratório de um item é a informação dele. Essa medida tem como objetivo medir quanto de informação este item é capaz de captar para cada valor de θ . Essa conta para um item dicotômico é dada pela expressão

$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad (2.2)$$

sendo $P'_{i,x}(\theta)$ a derivada da probabilidade de acerto dado θ , ou seja, a primeira derivada da equação 1, Q é $1 \sim P$ e $I_i(\theta)$ é a informação captada pelo item no valor de traço latente θ .

Além disso, quando se trabalha com o modelo logístico de 3 parâmetros, pode-se reescrever a função de informação do item como

$$I_i(\theta) = D^2 a_i^2 \frac{Q_i(\theta)}{P'_i(\theta)} \left[\frac{P'_i(\theta) - c_i}{1 - c_i} \right]^2 \quad (2.3)$$

É interessante notar, por esta equação, o papel que cada parâmetro tem na informação do item, tornando a informação maior quanto maior for o parâmetro a_i , quanto mais próximo de b_i for o parâmetro θ e quanto mais se aproximar de 0 o parâmetro c_i .

Além disso, é chamado de Informação do Teste $I(\theta)$ a soma de todas as informações captadas por todos os itens que compõem o teste (Andrade et al., 2000). Isto é,

$$I(\theta) = \sum_{i=1}^I I_i(\theta) \quad (2.4)$$

Dessa forma, a informação do item calcula, na prática, para quais valores de θ aquele item conseguirá ser capaz de coletar alguma informação. Desta forma, a combinação da informação de um item para todos os valores de θ gera uma curva chamada de Curva de Informação do Item, ou CII. Essa curva geralmente tem formato de sino, podendo ou não ser simétrica.

A Figura 2.2 compara 4 itens e suas respectivas CCI e CII, mostrando bem as diferenças já citadas entre os parâmetros, suas interpretações e as diferentes curvas que eles geram.

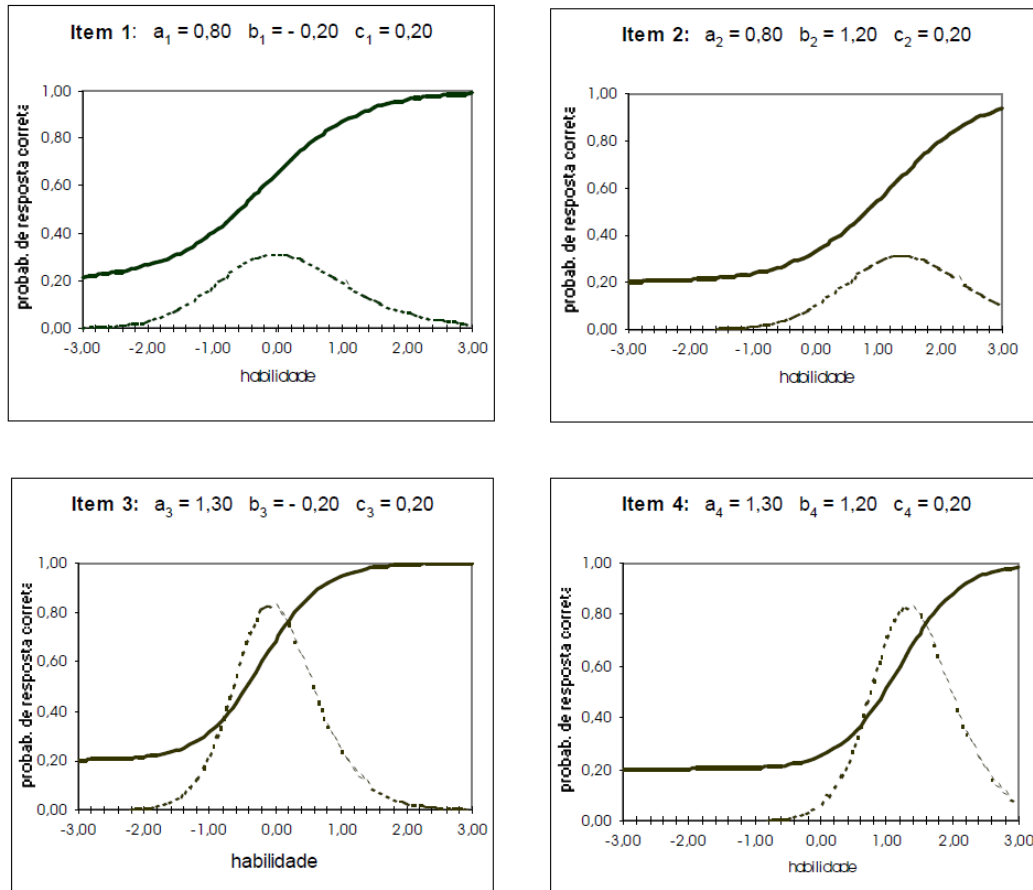


Figura 2.2: Exemplos de curvas características e de informação para quatro itens diferentes (Andrade et al., 2000).

É possível ver, por exemplo, como a CCI do item 3 tem um crescimento mais intenso do que o item 1, indicando que a a probabilidade de acerto aumenta bastante com um aumento pequeno do θ . Isso significa que o item 3 é mais adequado para diferenciar bem alunos que tem θ muito próximos entre si e com valores de θ próximos de b_i . Em outras palavras, o item é muito discriminante para alunos que tenham o traço latente parecido com a dificuldade do item.

Isso faz sentido, pois um aluno que tem o valor de $\theta = x$, por exemplo, provavelmente

vai acertar perguntas com b_i muito menor que x e vai errar itens com b_i muito maiores que x . Isso implica que o nico jeito de notar diferenas entre respondentes que tem valores de θ prximos  usando itens com valores de b_i tambm prximos de x , ou seja, itens que tem Informaco em torno do valor de θ desses alunos.

2.2 Estimaco no Paramtrica da CCI

2.2.1 Regresso Isotnica

Em 2014, Schlemper props a utilizaco da Regresso Isotnica como uma alternativa ao mtodo logstico de dois ou trs parmetros para a construco da CCI. A definico desta tcnica vem a seguir.

Definida formalmente por Barlow et al. (1972), a regresso isotnica parte do princpio que, em vrias situaes prticas, se pode obter uma informaco a priori de que os dados esto ou poderiam estar, de alguma forma, ordenados. Partindo disso, a ideia deste mtodo  parecida com a da regresso usual: minimizar a soma do quadrado dos erros.

Porm, com essa informaco extra, essa tarefa no se resume mais em estimar a e b para uma reta, mas sim em achar uma funo no decrescente que minimize essa soma. Outra diferena para a regresso linear simples  que a regresso isotnica  uma regresso no paramtrica, ou seja, no assume forma matemtica para a relao entre a varivel explicativa e a varivel resposta, apenas que ela  isotnica.

Seja X um conjunto finito de n elementos ordenados x_1, \dots, x_n , tais que $x_1 < x_2 < \dots < x_n$.

Definico 1: Uma funo f definida em X  chamada isotnica se, dados $x, y \in X$ e $x < y$ implica que $f(x) \leq f(y)$, ou seja, a funo  no decrescente.

Sejam g e w duas funes quaisquer e g^*  uma funo isotnica, todas definidas em X . Diz-se que g^*  uma regresso isotnica de g com peso w e com respeito  relao de

ordem $x_1 < x_2 < \dots < x_n$ se ela minimiza, na classe de funções isotônicas f em X , a soma

$$\sum_{i=1}^n [g(x_i) - f(x_i)]^2 w(x_i). \quad (2.5)$$

Para melhor se entender como encontrar a função g^* , primeiro, é preciso definir uma função importante: O Minorante Convexo Máximo (MCM).

Definição 2: O Minorante Convexo Máximo é definido como sendo a função

$H^* : [0, n] \rightarrow \mathbb{R}$ tal que

$$H^*(t) = \sup \left\{ H(t); H(i) \leq \sum_{j \leq i} g(x_j), 0 \leq i \leq n, H(0) = 0, H \text{convexa} \right\}$$

Para colocar em palavras, a função MCM é o gráfico do supremo de todas as funções convexas que estão abaixo do gráfico do Diagrama da Soma Acumulada (DSA). Trata-se, de uma linha que intercepta o DSA em alguns pontos, mas nunca está acima dele.

Para conseguir criar o DSA, é preciso plotar, em um gráfico de dispersão no plano cartesiano, os pares ordenados $P_k = (W_k, G_k)$, tais que:

$$G_k = \sum_{i=1}^k g(x_i) w(x_i), \quad (2.6)$$

e

$$W_k = \sum_{i=1}^k w(x_i), k = 1, 2, \dots, n. \quad (2.7)$$

Após plotados estes pontos, a ideia é ligá-los por dois métodos diferentes. O primeiro método, próprio do DSA, é fazer a inclinação do segmento que liga um ponto P_k ao ponto anterior ser

$$g(x_k) = \frac{G_k - G_{k-1}}{W_k - W_{k-1}}. \quad (2.8)$$

A segunda reta será o minorante e ligará um ponto P_{i-1} ao ponto P_k com $i < k$, ou seja, pulará alguns pontos para ir alternando entre mínimos locais. A inclinação deste

segmento, que representa a média ponderada, é dada por

$$A_{v\{x_i, x_{i+1}, \dots, x_k\}} = \frac{\sum_{r=i}^k g(x_r)w(x_r)}{\sum_{r=i}^k w(x_r)}. \tag{2.9}$$

A figura 2.3 mostra um exemplo de um gráfico DSA com o Minorante feito nos mesmos pontos.

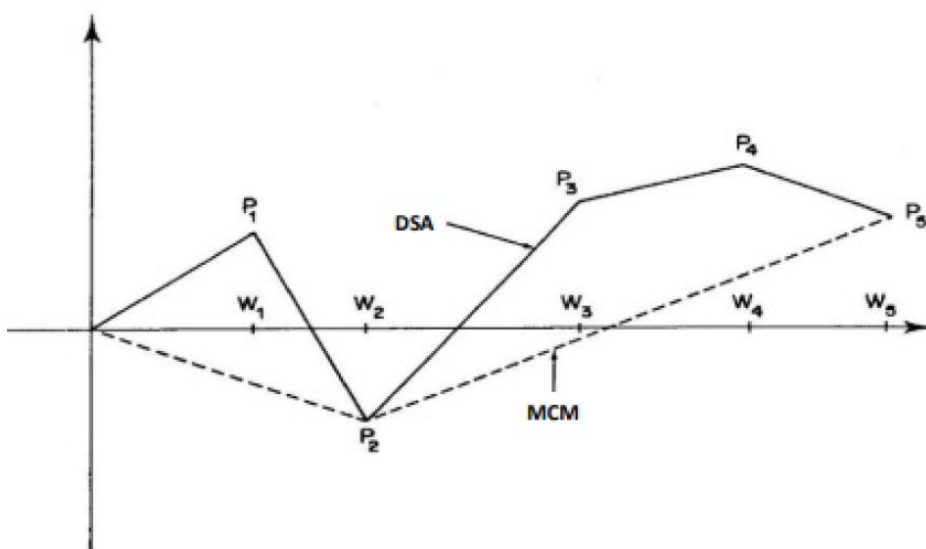


Figura 2.3: Exemplo de um DSA com os minorantes calculados.

Todas essas definições nos levam ao teorema central da Regressão Isotônica, descrito a seguir.

Teorema 1: Para um vetor de X ordenado, a inclinação g^* do MCM fornece a regressão isotônica de g , ou seja, se f é isotônica em X , então

$$\sum_x [g(x) - f(x)]^2 w(x) \geq \sum_x [g(x) - g^*(x)]^2 w(x) + \sum_x [g^*(x) - f(x)]^2 w(x). \tag{2.10}$$

A prova se encontra em Barlow et al. (1972).

2.2.2 Regressão Isotônica Suavizada

A regressão isotônica gera, então, uma função escada não decrescente que utiliza a informação presente nos pontos de θ para aumentar a probabilidade de resposta do usuário.

Para que essa curva se aproxime mais da curva da CCI, pode-se fazer uma suavização, tornando-a contínua. Uma possibilidade é se utilizar um Kernel de Nadaraya-Watson para alcançar esse objetivo.

Esse Kernel utilizará um parâmetro h para calcular uma média ponderada dos acertos e erros no intervalo entre dois pontos de θ com massa de probabilidade.

Esse parâmetro h , como visto na regressão de Nadaraya Watson, é chamado de parâmetro de suavização e controla a relação entre o viés e a variância do processo. Se o valor de h cresce, a curva fica mais suave, a variância do estimador como um todo diminui, e o seu vício aumenta.

Isso se dá porque, à medida que o parâmetro aumenta, a quantidade $\frac{q_j - \theta}{h}$ diminui e, dessa forma, os pesos dos valores que estão mais distantes do ponto θ aumentam.

Para este trabalho, o valor de h a ser escolhido será $\frac{1,1}{(n^{1/5})}$, conforme a literatura indica. Esse valor é ideal para evitar hipersuavização das curvas, o que acontece quando os valores de h são maiores do que o proposto.

Regressão de Nadaraya-Watson

Outra forma de aplicar a estimação não paramétrica da CCI é utilizando a regressão de Nadaraya Watson. Seja Y_i uma variável aleatória dicotômica que define se o respondente acertou o item i . Além disso, seja Θ a variável do traço latente.

A estimativa $\hat{\Theta}$ de Θ de se obter esse escore pode ser descrita por

$$\hat{P}_i^{non}(\hat{\Theta} = \theta) = \frac{\sum_j K \left[\frac{\theta - \hat{\theta}_j}{h} \right] Y_{ij}}{\sum_j K \left[\frac{\theta - \hat{\theta}_j}{h} \right]}, \quad (2.11)$$

sendo $\hat{\theta}_j$ é a estimação do traço latente do respondente e y_i é a variável indicadora do respondente j acertar o item i .

Para estimar a CCI, os y_i 's são calculados em volta de cada ponto de avaliação θ pelos seus pesos. Esses pesos são chamados de $K(u)$, um kernel simétrico não negativo com moda igual a 0, que é a diminuição monotônica do valor absoluto de u (Douglas e Cohen 2001). Neste estudo, a função Gaussiana é utilizada como o kernel na Equação 11, uma vez que é uma escolha típica para a estimativa da CCI não paramétrica (Douglas e Cohen 2001).

Na expressão dessa CCI, h é um parâmetro chamado discriminação que controla o intervalo de θ que será utilizado para estimar aquele ponto. Em outras palavras, esse valor controla o quão rápido o peso atinge o valor zero. Ramsay (1991) sugeriu que o valor ótimo de h para dados binários psicométricos depende do tamanho da amostra (N): $h = 1.1xn^{-0.2}$.

No presente estudo, este tamanho de amostra dependente do valor ótimo é utilizado para h porque Ramsay (1991) também mostrou que este valor funcionava bem sob a função do núcleo gaussiano.

2.3 Estatísticas para Adequabilidade de Ajuste

Como visto, a CCI pode ser estimada de muitas maneiras, tanto paramétricas, quanto não paramétricas. Por isso, é interessante encontrar formas de comparar essas várias estimativas para observar se a estimativa paramétrica da CCI está próxima das estimativas não paramétricas, o que indicaria que o modelo paramétrico ajustado é satisfatório para os dados coletados, independente de pressupostos do modelo.

Para isso, por exemplo, Douglas e Cohen (2001) propuseram um cálculo da distância entre a estimativa paramétrica da CCI obtida a partir da amostra coletada e a estimativa não paramétrica da CCI pelo método não paramétrico de Suavização por Kernel,

já descrito anteriormente.

Essa distância foi chamada de RISE, do inglês Root Integrated Squared Error. Na prática, essa medida integra a diferença entre as duas curvas para todos os valores de θ , como mostra a equação a seguir:

$$RISE = d(P, P^*) = \left\{ \int [P(\theta) - P^*(\theta)]^2 f(\theta) d\theta \right\}^{1/2}, \quad (2.12)$$

A mesma ideia de comparação de CCI's pode ser utilizada de várias formas, comparando o modelo paramétrico da CCI estimado a partir dos dados com a CCI estimada por métodos não paramétricos, muitos dos quais já foram discutidos anteriormente.

Cada uma dessas diferenças pode ser entendida como uma estatística de Teste para medir adequabilidade do ajuste do modelo paramétrico. Neste trabalho, serão utilizadas seis estatísticas como esta, as quais serão detalhadas a seguir:

Diferença entre a estimação não paramétrica por regressão isotônica e a estimação paramétrica da CCI.

$$d_i = \sqrt{\int |\hat{\pi}_{np}(\theta) - \hat{\pi}_p(\theta)| g(\theta) d\theta} \quad (2.13)$$

Diferença quadrática entre a estimação não paramétrica por regressão isotônica e a estimação paramétrica da CCI.

$$d_i = \sqrt{\int |\hat{\pi}_{np}(\theta) - \hat{\pi}_p(\theta)|^2 g(\theta) d\theta} \quad (2.14)$$

Diferença entre a estimação não paramétrica por regressão isotônica com suavização e a estimação paramétrica da CCI.

$$d_i = \sqrt{\int |\hat{\pi}_{np}^s(\theta) - \hat{\pi}_p(\theta)| g(\theta) d\theta} \quad (2.15)$$

Diferença quadrática entre a estimação não paramétrica por regressão isotônica com

suavização e a estimação paramétrica da CCI.

$$d_i = \sqrt{\int |\hat{\pi}_{np}^s(\theta) - \hat{\pi}_p(\theta)|^2 g(\theta) d\theta} \quad (2.16)$$

Diferença entre a estimação não paramétrica por regressão de Nadaraya-Watson e a estimação paramétrica da CCI.

$$d_i = \sqrt{\int |\hat{\pi}_{nw}^s(\theta) - \hat{\pi}_p(\theta)| g(\theta) d\theta} \quad (2.17)$$

Diferença quadrática entre a estimação não paramétrica por regressão de Nadaraya-Watson e a estimação paramétrica da CCI.

$$d_i = \sqrt{\int |\hat{\pi}_{nw}^s(\theta) - \hat{\pi}_p(\theta)|^2 g(\theta) d\theta} \quad (2.18)$$

Em 2014, Schlemper utilizou-se de ideias parecidas de comparar a CCI paramétrica com estimações não paramétricas dessa curva. A principal diferença deste trabalho para aquele é que os parâmetros que entram no processo de PPMC neste trabalho são retirados de uma distribuição a posteriori deles, enquanto no trabalho de Schlemper, foi utilizado um Bootstrap em cima do parâmetro estimado para fazer a reamostragem.

Essa mudança foi feita por permitir que o parâmetro assuma valores diferentes, reduzindo o viés causado pela replicação da estimação paramétrica do parâmetro que o Bootstrap faz.

Capítulo 3

Estimação Bayesiana

3.1 Teoria Bayesiana

A teoria Bayesiana era dominante no final do século XIX e início do século XX, e tem como ideia se aproveitar tanto da amostra coletada, quanto de conhecimentos prévios ou esperados da variável de interesse (AITKIN, 2010). Dessa forma, termos muito utilizados são distribuição *a priori* e *a posteriori*.

A primeira, se refere ao conhecimento que já se possui sobre o parâmetro, ou variáveis externas à pesquisa. Já a distribuição *a posteriori* se refere à distribuição resultante depois que as informações coletadas na amostra são aplicadas a essa *priori*. Dessa forma, a Distribuição *a posteriori* pode ser definida como

$$P(\theta|X) = \frac{L(\theta)P(\theta)}{\int L(\theta)P(\theta)d\theta} \quad (3.1)$$

sendo $L(\theta)$ a Função de Verossimilhança de θ e $P(\theta)$ a sua distribuição *a priori*.

Em 2000, Clement Stone propôs um abordagem que tem a sua Estatística do Teste baseada em probabilidade a posteriori, a ser calculada utilizando processos Bayesianos. Essa abordagem consiste em definir prioris para os parâmetros do modelo a ser utilizado e construir a CCI com base nas posterioris que serão obtidas.

No caso do modelo Logístico de três parâmetros, as prioris seguirão como a seguir:

$$a_i \sim \text{Lognormal}(0, 0.5); b_i \sim \text{Normal}(0, 2); c_i \sim \text{Beta}(50, 200); \theta_i \sim \text{Normal}(0, 1) \quad (3.2)$$

Desta forma, a probabilidade a posteriori é calculada usando uma adaptação da Fórmula de Bayes para probabilidade condicional, dada por:

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)}, \quad (3.3)$$

sendo $P(\theta|x)$ a distribuição a posteriori do parâmetro θ , $P(x|\theta)$ a probabilidade condicional de se responder a alternativa x tendo traço latente no valor de θ , $P(\theta)$ é a distribuição a priori de θ , que geralmente é definida como Normal(0,1) e $P(x)$ é a probabilidade estimada de resposta da alternativa x .

Para fazer esse cálculo, mais uma vez será preciso calcular com aproximação de quadratura. Ao usar q pontos de quadratura, a equação anterior fica

$$P(\theta_q|x_j) = \frac{P(x_j|\theta_q)w(\theta_q)}{\sum_{k=1}^q P(x_j|\theta_k)w(\theta_k)}, \quad (3.4)$$

sendo $w(\theta_q)$ o peso que corresponde à densidade da Distribuição Normal.

Quando o item é dicotômico, temos que essa probabilidade se torna uma proporção de acertos, ou seja

$$P_{posteriori}(\theta_q) = \frac{A_{iq}}{N_{iq}} = \frac{\sum_{j=1}^N P(\theta_q|x_j)Y_{ij}}{\sum_{j=1}^N P(\theta_q|x_j)}, \quad (3.5)$$

sendo Y_{ij} uma variável indicadora de acerto da resposta ao item i dada pelo indivíduo j , ou seja, $P_{posteriori}(\theta_q)$ é uma razão entre a soma das distribuições a posteriori de quem acertou o item sobre a soma da distribuição a posteriori de todos os respondentes nível

de habilidade q .

Essas probabilidades a posteriori, então, geram a CCI de cada item, finalizando a estimação bayesiana da curva.

3.2 Método de Monte Carlo via Cadeia de Markov (MCMC)

O MCMC é uma classe de métodos de simulação amostral que se utiliza da técnica das Cadeias de Markov para criar artificialmente uma amostra da distribuição de uma estatística qualquer. A principal diferença do MCMC para os processos de Monte Carlo usuais é que, nos processos usuais, as amostras gerais são independentes umas das outras, enquanto no MCMC, essas amostras são correlacionadas umas com as outras (Sharma, 2017).

Para gerar as amostras, o método gera uma cadeia de Markov não periódica e irredutível que tenha como distribuição estacionária a distribuição da estatística de interesse. Essa distribuição será chamada de $\pi(x)$.

Para se obter uma distribuição com essas características, é preciso encontrar uma distribuição de transição - ou Kernel de transição - $K(x, x + 1)$ que liga o estado presente x ao próximo passo da cadeia, o estado $x + 1$.

Para satisfazer a estacionariedade da cadeia, é necessário encontrar o K tal que $\pi K = \pi$, ou seja, dado uma observação $x \sim \pi(x)$, se encontrarmos um $x + 1$ tal que $x + 1 \sim K(x, x + 1)$, então também temos que $x + 1 \sim \pi(x + 1)$. Com essa condição satisfeita, ter-se-ia que aplicar o Kernel a um valor seria equivalente a aplicar a distribuição π , então a cadeia seria estacionária.

Desta forma, após o processo de convergência dessa cadeia, ou seja, após a cadeia atingir o seu processo estacionário, se pode tomar os próximos n valores e retirar, dessa forma, uma amostra da distribuição alvo. Os valores que a cadeia atingiu antes da convergência são chamados de Burnin e são descartados.

Além disso, para evitar que a amostra tenha qualquer problema de periodicidade, ao retirar valores da cadeia de Markov e adicionar na amostra, geralmente se salta algumas observações. Dessa forma, define-se um número k tal que os valores que compõem a amostra serão os valores $k, 2k, 3k, \dots$ até se atingir o tamanho da amostra desejado. Esse valor de k é chamado de *lag*.

3.3 Algoritmo Metropolis-Hastings

O algoritmo de MCMC mais geral é o chamado Algoritmo de Metropolis-Hastings, que é produto de dois artigos (Metropolis et al. 1953 e Hastings; 1970). A ideia é construir um Kernel de Transição que consiga passar de um ponto x para um ponto y dentro de um mesmo espaço E utilizando-se dois passos:

- Especificar uma proposta de uma distribuição $q(y|x)$.
- Retirar observações de $q(y|x)$ e aceitar ou rejeitar essas observações segundo a função $a(x, y) = \min \left[1, \frac{f(y)q(x|y)}{f(x)q(y|x)} \right]$.

Portanto, podemos definir $K(x, y) = q(y|x)a(x, y)$.

O algoritmo então, se resume a obter um valor inicial, que será passado como *input* e aplicar a função $q(y|x)$ nesse valor. Depois disso, se retira um valor único de uma Uniforme no intervalo $(0, 1)$. Por fim, compara-se essa Uniforme com a função $\frac{f(y)q(x|y)}{f(x)q(y|x)}$. Se for menor, aceita-se o valor de y como o próximo valor amostrado e se for maior, repete-se o valor anterior na amostra.

Dessa forma, tem-se que o Kernel gerado por Metropolis-Hastings é inversível, ou seja, $f(x)K(x, y) = f(y)K(y, x)$, que é uma propriedade desejável em um processo MCMC.

Capítulo 4

Posterior Predictive Model Checking (PPMC)

O PPMC foi apresentado inicialmente por Gelman et al. (1996). Esse método leva em consideração a incerteza da estimação do parâmetro ao integrar a sua distribuição a posteriori.

O PPMC é parecido com o Bootstrap no sentido de que os dois tem como objetivo gerar uma distribuição para o valor de referência. Porém, a diferença é que o Bootstrap usa apenas um conjunto de parâmetros a_i , b_i e c_i para cada item, enquanto o PPMC vai usar valores diferentes desses parâmetros cada vez que gerar uma observação, ou seja, cada vez que fizer uma reamostra. (Wang et al., 2020)

O uso deste método em TRI vem crescendo ao longo dos últimos anos. Alguns exemplos disso são Sinharay (2005), Kuhfeld (2019), Li et al. (2017) e Levi et al. (2009). Os parágrafos a seguir detalharão o funcionamento do PPMC.

Seja w um parâmetro desconhecido que faça parte do modelo. Para este trabalho, esse parâmetro pode ser pensado como os três parâmetros do modelo logístico. Seja y o dado observado, $P(w)$ a distribuição a priori do parâmetro w e $P(y|w)$ a distribuição de probabilidade de y dado que os parâmetros são verdadeiros - neste trabalho, o mo-

delo logístico de três parâmetros. Então, como mostrou Sinharay (2005), a distribuição preditiva a posteriori $P(w|y)$ pode ser calculada usando a equação a seguir:

$$P(w|y) \propto P(y|w)P(w) \quad (4.1)$$

Seja y^{rep} a amostra replicada, que pode ser interpretada como a amostra que seria observada na próxima vez que fossem coletados os mesmos dados das mesmas pessoas. A simulação desse y^{rep} é feita de forma que ele siga o mesmo modelo da amostra coletada e, substituindo os parâmetros calculados pela amostra inicial com os parâmetros encontrados pela equação anterior, temos que a distribuição a posteriori de y^{rep} é

$$P(y^{rep}|y) = \int P(y^{rep}|w)P(w|y)dw. \quad (4.2)$$

O próximo passo é calcular o PPP-valor - da sigla em inglês *Posterior Predictive P-Value* - para checar se a distribuição de y^{rep} está próxima o suficiente de y , ou seja, se a modelagem da reamostragem se ajustou à modelagem coletada nos dados reais (Gelman et al. 1996). Em 2000, Bayarri & Berger definiram o PPP-valor como

$$P(D(y^{rep}, w) \geq D(y, w)|y) = \int \int_{D(y^{rep}, w) \geq D(y, w)} P(y^{rep}|w)P(w|y)dw dy^{rep}, \quad (4.3)$$

sendo $D(y)$ uma medida de discrepância que será utilizada para medir o ajuste do modelo.

Apesar deste valor ser muito parecido com o p-valor que se utiliza em testes de hipóteses, é importante notar que o método PPMC não envolve testagem de hipótese nenhuma, apenas indicando se há ou não diferenças entre as medidas de interesse empíricas e preditas.

Essa diferença pode observada quando os valores do PPP-valor são extremos, ou seja, próximos de 0 ou 1, dependendo do tipo de medida de discrepância utilizada. Nesse

trabalho, valores extremos serão valores próximos de 0, visto que estamos analisando quantas reamostras ficaram maiores que um certo D .

Capítulo 5

Simulação e Aplicação a Dados Reais

Para tanto, foi construído um código, utilizando o software R, que completasse todos esses passos. Em um primeiro momento, esse código será aplicado a dados simulados. Nessa etapa, serão construídos duas provas diferentes: uma que tenha todos os itens se ajustando bem ao modelo de três parâmetros, e outra que terá oito itens que não se ajustem a esse modelo.

O esperado, portanto, é que o teste seja capaz de reconhecer, com certa confiabilidade, quais são os itens que tem esse problema de adequabilidade na segunda prova e não reconheça nenhum item na primeira. Apesar disso, se algum item for identificado como desajustado, pode-se analisar se ele tem algo específico que tenha causado isso, possibilitando o aprimoramento do teste.

Posteriormente aos dados simulados, o teste também será aplicado a dados reais, com uma aplicação em 1001 respondentes do turno da manhã do SARESP de 2007. Dessa forma, será possível entender como o teste se comporta em um ambiente não controlado e quais as diferenças entre um e outro.

Para o teste, o valor de M a ser utilizado para a obtenção de cada parâmetro é 25100. Esse valor foi escolhido para que pudéssemos ter uma amostra de tamanho 500, utilizando um lag de 50 e um burn-in de 100. Esse processo de MCMC será replicado 50

vezes para que podemos gerar valores de Erro Padrão para cada PPP.

Além disso, foram feitas 100 replicações do estudo para obter o erro tipo I empírico.

5.1 Aplicação a dados simulados

Para dados de simulação, foram considerados três conjuntos distintos de itens para se simular 1000 respostas a cada um desses itens.

Em um primeiro momento, foram geradas respostas para 48 itens com os seguintes valores para os parâmetros de discriminação (a), de dificuldade (b) e de acerto ao acaso (c). Seus valores estão apresentados na Tabela 5.1.

a	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5
b	-1,4	-1,0	-0,6	-0,2	0,2	0,6	1,0	1,4
c	0,2	0,25	0,2	0,25	0,2	0,25	0,2	0,25
a	0,8	0,8	0,8	0,8	0,8	0,8	0,8	0,8
b	-1,4	-1,0	-0,6	-0,2	0,2	0,6	1,0	1,4
c	0,2	0,25	0,2	0,25	0,2	0,25	0,2	0,25
a	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1
b	-1,4	-1,0	-0,6	-0,2	0,2	0,6	1,0	1,4
c	0,2	0,25	0,2	0,25	0,2	0,25	0,2	0,25
a	1,4	1,4	1,4	1,4	1,4	1,4	1,4	1,4
b	-1,4	-1,0	-0,6	-0,2	0,2	0,6	1,0	1,4
c	0,2	0,25	0,2	0,25	0,2	0,25	0,2	0,25
a	1,7	1,7	1,7	1,7	1,7	1,7	1,7	1,7
b	-1,4	-1,0	-0,6	-0,2	0,2	0,6	1,0	1,4
c	0,2	0,25	0,2	0,25	0,2	0,25	0,2	0,25

a	2,0	2,0	2,0	2,0	2,0	2,0	2,0	2,0
b	-1,4	-1,0	-0,6	-0,2	0,2	0,6	1,0	1,4
c	0,2	0,25	0,2	0,25	0,2	0,25	0,2	0,25

Tabela 5.1: Parâmetros reais para cada item ML3.

Para avaliar se o teste é capaz de reconhecer desajustes, foram adicionados aos 48 itens anteriores mais oito itens que não pertenciam ao ML3. Desses oito, os primeiros quatro itens eram tais que suas curvas características do item (CCI) são misturas (com pesos iguais) de duas CCI's com parâmetros:

$$a = 0,5 \text{ e } a = 3.$$

A Tabela 5.2 mostra os valores dos parâmetros para esses quatro itens.

Item	49	50	51	52
a_I	0,5	0,5	3	3
b_I	-1,5	-2	-1,5	-2
c_I	0,2	0,2	0,2	0,2
a_{II}	3	3	0,5	0,5
b_{II}	1,5	2	1,5	2
c_{II}	0,2	0,2	0,2	0,2

Tabela 5.2: Parâmetros dos 4 itens de mistura.

Além disso, mais quatro itens foram adicionados cujas curvas características do item são não monotônicas com CCI's

$$P_i(\theta) = c_i + (1 - c_i)/(1 + \exp(-0,5 \times a_i * ((\theta - d_i)^3 - (\theta - d_i) - b_i))).$$

Seus parâmetros podem ser vistos na Tabela 5.3.

Todas as estimativas de CCI destes últimos 8 itens podem ser vistas nas figuras 5.1 a 5.8.

Item	53	54	55	56
a_i	1,5	3	2	3
b_i	0	1	0	0
c_i	0,2	0,2	0,2	0,2
d_i	0	-1	0	1

Tabela 5.3: Parâmetros dos 4 itens não monotônicos.

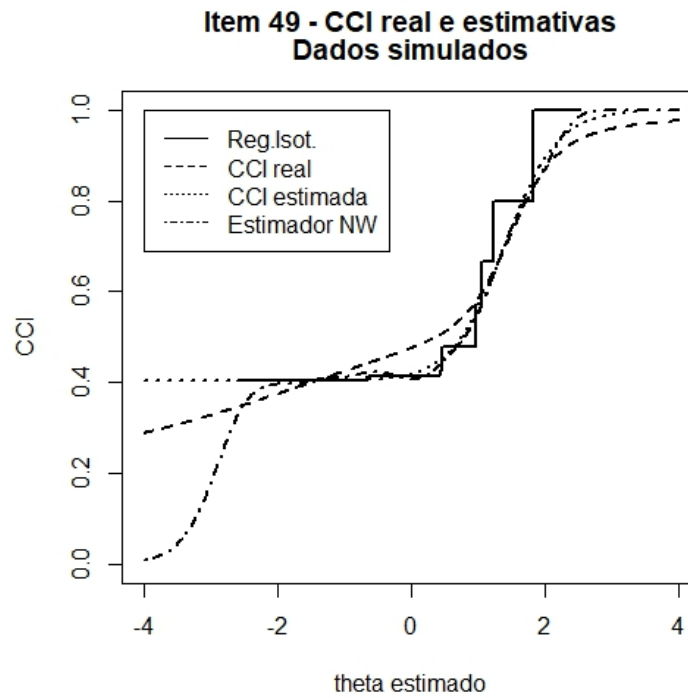


Figura 5.1: CCI's do Item 49.

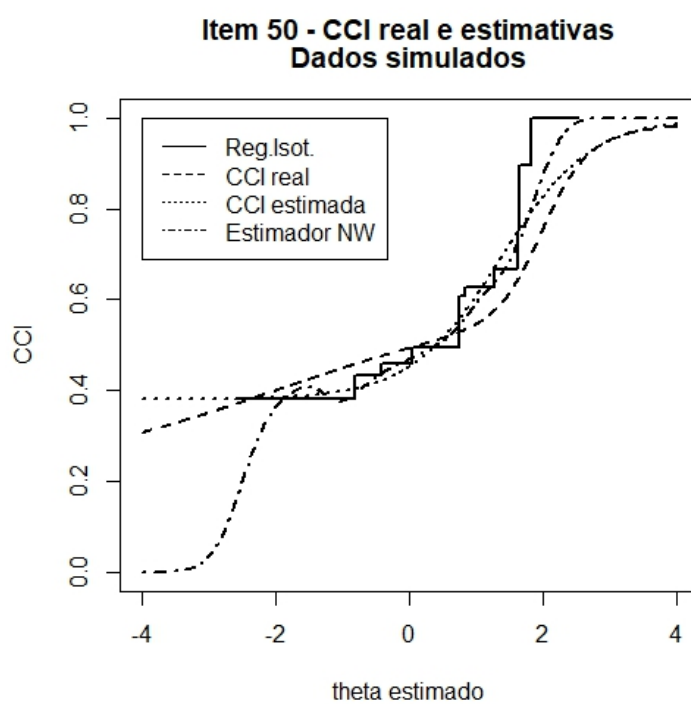


Figura 5.2: CCI's do Item 50.

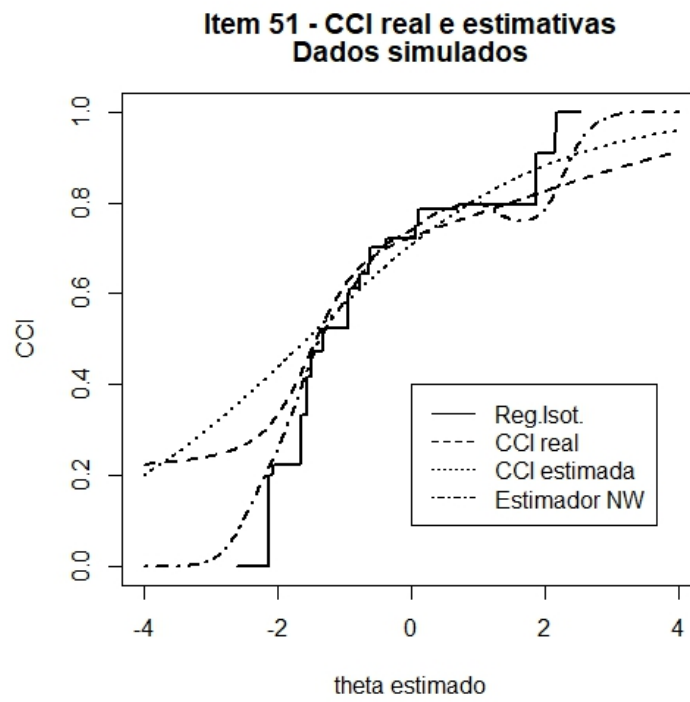


Figura 5.3: CCI's do Item 51.

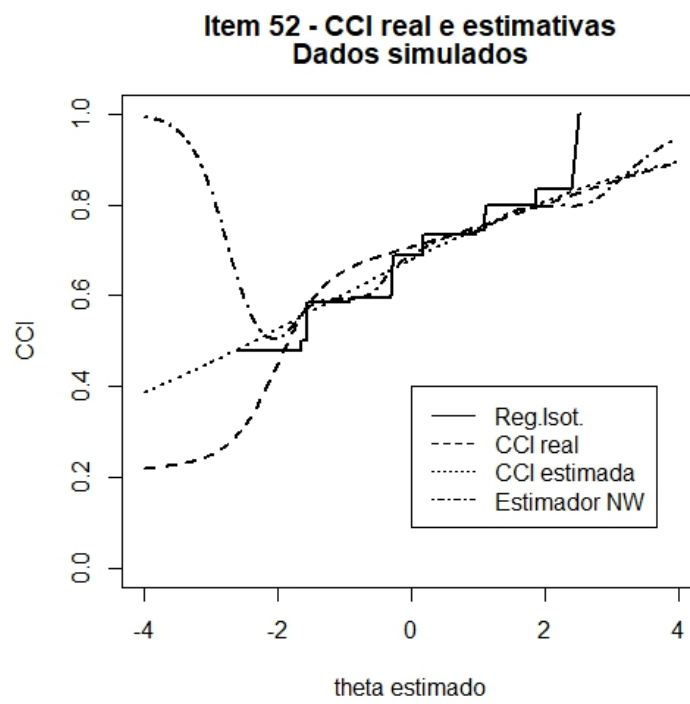


Figura 5.4: CCI's do Item 52.

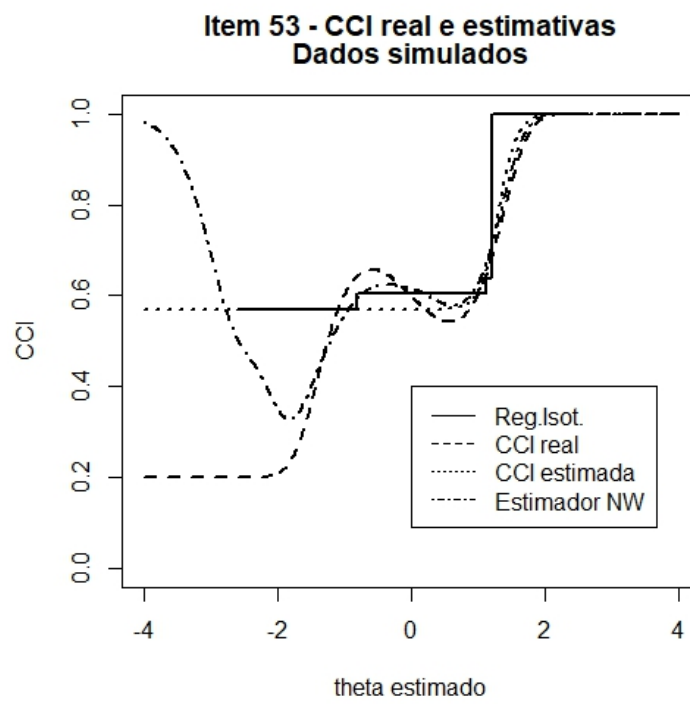


Figura 5.5: CCI's do Item 53.

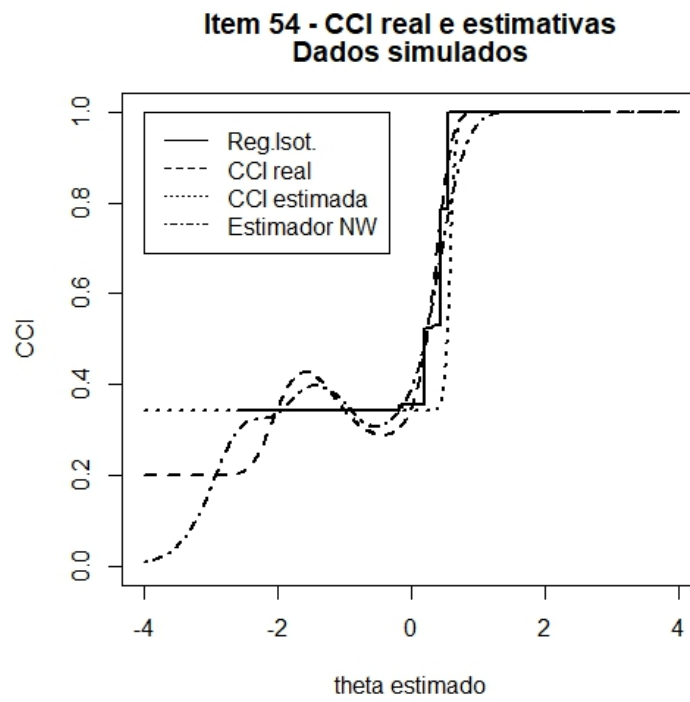


Figura 5.6: CCI's do Item 54.

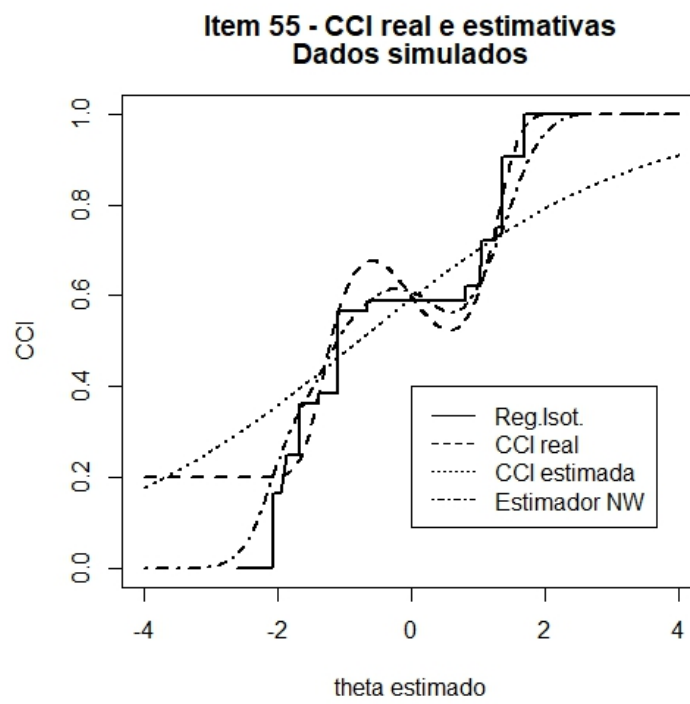


Figura 5.7: CCI's do Item 55.

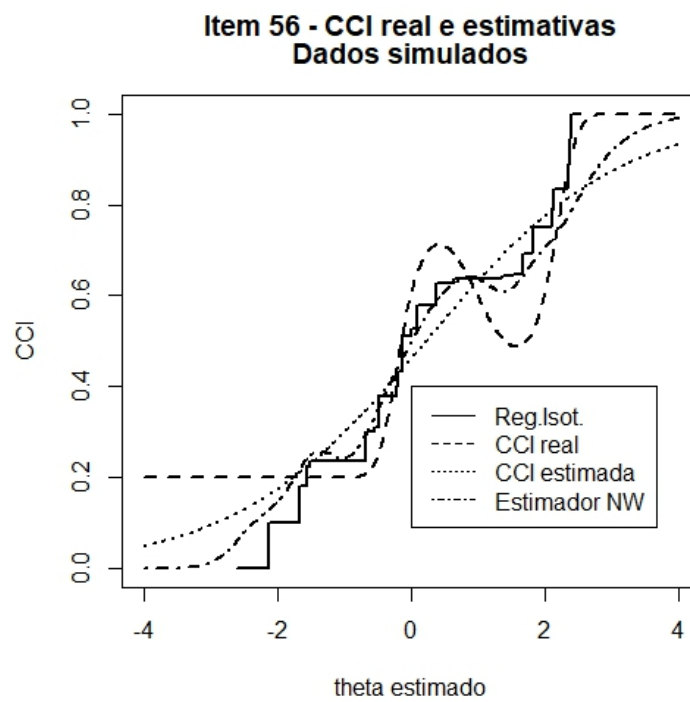


Figura 5.8: CCI's do Item 56.

A tabela 5.4 apresenta as estimativas dos parâmetros dos 56 itens.

Item	a	b	c
1	0.5610	-20.931	0,000005
2	0.8281	14.553	0,583211
3	0.4039	-20.130	0,001625
4	0.5216	-10.201	0,000001
5	0.6634	0.5712	0,280227
6	0.4383	-0.6862	7,31E-07
7	0.3819	-0.2636	2,42E-07
8	0.5028	0.3633	0,098922
9	0.6899	-18.024	0,042422
10	1.338	-0.4437	0,339918
11	0.7669	-0.6206	0,119385
12	0.7183	-0.0805	0,24825
13	0.6077	-0.4973	0,000004
14	0.9649	0.9628	0,32823
15	0.9949	0.9005	0,173087
16	0.7840	15.001	0,261484
17	10.969	-14.661	0,139228
18	11.357	-0.8818	0,318068
19	11.875	-0.3063	0,278748
20	12.207	-0.2041	0,229485
21	10.938	-0.0043	0,124905
22	12.208	0.5493	0,261985
23	13.621	11.778	0,257246
24	10.304	13.270	0,214614

25	13.594	-12.570	0,329654
26	12.399	-14.371	0,00095
27	15.607	-0.5603	0,219776
28	10.262	-0.6716	0,067744
29	13.065	0.1881	0,187963
30	11.977	0.2578	0,198575
31	15.463	10.451	0,214695
32	21.254	11.799	0,253875
33	17.580	-13.102	0,246164
34	24.157	-0.6219	0,41301
35	16.294	-0.6166	0,211816
36	18.721	-0.2174	0,218091
37	14.697	0.1767	0,177794
38	18.260	0.8063	0,274474
39	21.243	0.9746	0,205544
40	18.392	12.480	0,260873
41	27.728	-0.8998	0,437731
42	19.907	-0.9244	0,246423
43	22.500	-0.4686	0,217278
44	21.881	-0.0523	0,309219
45	24.597	0.2434	0,210077
46	22.286	0.4321	0,228205
47	23.559	0.9977	0,209691
48	31.800	11.767	0,28703
49	25.445	14.018	0,402903
50	14.780	13.679	0,380491

51	0.5639	-15.596	4,19E-01
52	0.3647	-13.377	0,15523
53	73.349	13.144	0,569375
54	249.180	0.5944	0,342615
55	0.4773	-0.7819	0,000001
56	0.6957	0.2246	2,31E-08

Tabela 5.4: Parâmetros estimados para os 56 itens da prova simulada.

É importante notar que muitos dos valores estimados pelo modelo de TRI não são adequados ao modelo, como é o caso dos valores de b e c para o item 2, por exemplo. Isso se dá, provavelmente, por conta de uma amostra pequena em relação ao quantitativo de itens na hora da calibração desses parâmetros.

Por causa disso, se o processo de calibração mais completo fosse feito, muitos desses itens seriam retirados. Porém, como o objetivo é analisar o teste em si e o código é bem demorado de se rodar, esse processo não será feito nesse trabalho.

Os valores PPP para os 56 itens e para as 6 estatísticas consideradas neste trabalho estão na tabela 5.5.

Item	RIA	RIQ	RIAS	RIQS	RNWA	RNWQ
1	0,5640 (0,0487)	0,0022 (0,0014)	0,4778 (0,0499)	0,0040 (0,0015)	0,9798 (0,0050)	0,7110 (0,0525)
2	0,6590 (0,0942)	0,0622 (0,0187)	0,4746 (0,0741)	0,0560 (0,0216)	0,9134 (0,0274)	0,9388 (0,0129)
3	0,3152 (0,0692)	0,0078 (0,0027)	0,2968 (0,0519)	0,0088 (0,0042)	0,9358 (0,0086)	0,8678 (0,0180)

4	0,5280 (0,0397)	0,0534 (0,0095)	0,4076 (0,0440)	0,0266 (0,0058)	0,9122 (0,0104)	0,8932 (0,0107)
5	0,8798 (0,0572)	0,5422 (0,0782)	0,6486 (0,0668)	0,2592 (0,0640)	0,9940 (0,0038)	0,9768 (0,0070)
6	0,0188 (0,0088)	0,0100 (0,0086)	0,1108 (0,0264)	0,0256 (0,0120)	0,2724 (0,0485)	0,4416 (0,0500)
7	0,1500 0 (0,0323)	,0264 (0,0161)	0,1780 (0,0330)	0,0202 (0,0130)	0,2342 (0,0546)	0,1886 (0,0620)
8	0,1656 (0,0562)	0,1218 (0,0474)	0,2558 (0,0488)	0,1962 (0,0605)	0,7052 (0,0574)	0,7714 (0,0453)
9	0,4220 (0,0593)	0,0442 (0,0143)	0,4476 (0,0732)	0,0360 (0,0126)	0,9946 (0,0022)	0,9968 (0,0018)
10	0,7062 (0,0597)	0,4188 (0,0804)	0,6764 (0,0973)	0,2882 (0,1045)	0,9930 (0,0053)	0,9890 (0,0081)
11	0,1750 (0,0366)	0,1788 (0,0416)	0,2314 (0,0619)	0,2184 (0,0760)	0,8312 (0,0390)	0,8462 (0,0223)
12	0,6790 (0,0741)	0,3926 (0,0761)	0,6858 (0,0891)	0,2286 (0,0729)	0,9808 (0,0068)	0,9836 (0,0036)
13	0,6490 (0,0431)	0,2242 (0,0510)	0,6422 (0,0653)	0,1522 (0,0526)	0,8872 (0,0167)	0,9330 (0,0120)
14	0,5698 (0,0689)	0,3718 (0,0979)	0,6466 (0,0915)	0,4230 (0,1244)	0,8836 (0,0287)	0,8848 (0,0248)
15	0,4810 (0,0846)	0,4732 (0,0661)	0,7998 (0,0663)	0,5740 (0,1034)	0,9710 (0,0116)	0,9910 (0,0044)
16	0,3224	0,2912	0,3628	0,3576	0,6676	0,8386

	(0,0855)	(0,0845)	(0,0822)	(0,1068)	(0,0541)	(0,0243)
17	0,1310 (0,0200)	0,0166 (0,0076)	0,2892 (0,0421)	0,0194 (0,0058)	0,8686 (0,0239)	0,7292 (0,0320)
18	0,3456 (0,0263)	0,0766 (0,0145)	0,2896 (0,0496)	0,0704 (0,0254)	0,9056 (0,0237)	0,6676 (0,0529)
19	0,2370 (0,05453)	0,2164 (0,0448)	0,6130 (0,0897)	0,3288 (0,0922)	0,9934 (0,0034)	0,9976 (0,0023)
20	0,1516 (0,0474)	0,1352 (0,0502)	0,5744 (0,1212)	0,4532 (0,1474)	0,9674 (0,0144)	0,9898 (0,0051)
21	0,8434 (0,0475)	0,7700 (0,0612)	0,7428 (0,0836)	0,6518 (0,1167)	1,0000 (0,0000)	0,9994 (0,0009)
22	0,8090 (0,0616)	0,7848 (0,0631)	0,8406 (0,0683)	0,7174 (0,1248)	0,9994 (0,0009)	0,9992 (0,0013)
23	0,3950 (0,1062)	0,1184 (0,0458)	0,5676 (0,1018)	0,3224 (0,0991)	0,9378 (0,0315)	0,9582 (0,0198)
24	0,4960 (0,0932)	0,4084 (0,0678)	0,7874 (0,0727)	0,6046 (0,0906)	0,9936 (0,0032)	0,9856 (0,0067)
25	0,7150 (0,0254)	0,2868 (0,0403)	0,7060 (0,0522)	0,1086 (0,0164)	0,9898 (0,0030)	0,8298 (0,0384)
26	0,3612 (0,0280)	0,0248 (0,0046)	0,3726 (0,0559)	0,0688 (0,0270)	0,9968 (0,0018)	0,9670 (0,0083)
27	0,2728 (0,0494)	0,2662 (0,0374)	0,4744 (0,0892)	0,4160 (0,1074)	0,9758 (0,0098)	0,9830 (0,0084)
28	0,0838 (0,0201)	0,0328 (0,0120)	0,1846 (0,0452)	0,0978 (0,0402)	0,7406 (0,0652)	0,8264 (0,0360)
29	0,0282	0,0222	0,1702	0,1792	0,4914	0,5402

	(0,0127)	(0,0127)	(0,0480)	(0,0654)	(0,1108)	(0,0817)
30	0,2472 (0,0648)	0,2318 (0,0459)	0,5156 (0,0675)	0,4336 (0,0757)	0,9962 (0,0024)	0,9942 (0,0032)
31	0,5072 (0,0833)	0,5778 (0,0605)	0,7232 (0,0702)	0,7340 (0,0711)	0,9670 (0,0132)	0,9826 (0,0095)
32	0,5712 (0,1069)	0,4788 (0,0987)	0,6970 (0,0875)	0,5616 (0,1275)	0,9972 (0,0027)	0,9998 (0,0006)
33	0,0190 (0,0098)	0,0200 (0,0064)	0,2566 (0,0516)	0,0380 (0,0117)	0,8694 (0,0385)	0,9320 (0,0216)
34	0,1202 (0,0218)	0,0554 (0,0113)	0,7606 (0,0491)	0,2708 (0,0560)	0,9856 (0,0072)	0,8802 (0,0448)
35	0,1554 (0,0341)	0,1282 (0,0146)	0,5098 (0,0657)	0,2634 (0,0467)	0,9860 (0,0061)	0,9834 (0,0069)
36	0,2796 (0,0433)	0,3114 (0,0396)	0,5490 (0,0794)	0,5360 (0,1010)	1,0000 (0,0000)	1,0000 (0,0000)
37	0,4488 (0,0617)	0,4994 (0,0716)	0,5354 (0,0965)	0,5040 (0,1254)	0,9830 (0,0060)	0,8696 (0,0262)
38	0,2638 (0,0774)	0,2874 (0,0746)	0,5196 (0,0856)	0,5252 (0,1000)	0,8952 (0,0257)	0,7144 (0,0539)
39	0,5920 (0,0628)	0,4886 (0,0368)	0,7714 (0,0443)	0,6088 (0,0538)	0,9768 (0,0129)	0,9808 (0,0106)
40	0,6350 (0,0559)	0,4436 (0,0676)	0,8158 (0,0470)	0,5164 (0,0868)	0,8692 (0,0195)	0,9560 (0,0110)
41	0,8572 (0,0299)	0,3318 (0,0370)	0,9204 (0,0158)	0,2106 (0,0223)	0,9922 (0,0055)	0,9602 (0,0218)
42	0,9000	0,7234	0,8068	0,5446	0,9826	0,9188

	(0,0214)	(0,0382)	(0,0367)	(0,0454)	(0,0061)	(0,0267)
43	0,4292 (0,0576)	0,3864 (0,0496)	0,8594 (0,0225)	0,7204 (0,0428)	0,9658 (0,0103)	0,9838 (0,0080)
44	0,1280 (0,0420)	0,0508 (0,0139)	0,6332 (0,0810)	0,4364 (0,0858)	0,9678 (0,0154)	0,9142 (0,0231)
45	0,6854 (0,0585)	0,7148 (0,0417)	0,9764 (0,0108)	0,8734 (0,0445)	0,9984 (0,0019)	0,9912 (0,0051)
46	0,4888 (0,0949)	0,3612 (0,0583)	0,8480 (0,0592)	0,7710 (0,0915)	0,9914 (0,0055)	0,9658 (0,0175)
47	0,4950 (0,0911)	0,3052 (0,1161)	0,7732 (0,0778)	0,6828 (0,1172)	0,9888 (0,0070)	0,8492 (0,0334)
48	0,6188 (0,1125)	0,2846 (0,0946)	0,5892 (0,0867)	0,4100 (0,0917)	0,9774 (0,0129)	0,9942 (0,0041)
49	0,7842 (0,0619)	0,2224 (0,0540)	0,7616 (0,0576)	0,2662 (0,0495)	0,9998 (0,0006)	0,9994 (0,0009)
50	0,5746 (0,1086)	0,2306 (0,0766)	0,7196 (0,0871)	0,2794 (0,0946)	0,9334 (0,0172)	0,8180 (0,0247)
51	0,0098 (0,0060)	0,0002 (0,0006)	0,0066 (0,0055)	0,0000 (0,0000)	0,0820 (0,0390)	0,0314 (0,0100)
52	0,6958 (0,1032)	0,2794 (0,0696)	0,4790 (0,0869)	0,2104 (0,0709)	0,9886 (0,0052)	0,8728 (0,0276)
53	0,1622 (0,0517)	0,0042 (0,0049)	0,0302 (0,0120)	0,0050 (0,0049)	0,0924 (0,0216)	0,0126 (0,0084)
54	0,0014 (0,0018)	0,0000 (0,0000)	0,0320 (0,0103)	0,0000 (0,0000)	0,7540 (0,1378)	0,2100 (0,0626)
55	0,0000	0,0000	0,0002	0,0000	0,0028	0,0042

	(0,0000)	(0,0000)	(0,0006)	(0,0000)	(0,0024)	(0,0045)
56	0,0042	0,0176	0,0146	0,0342	0,1370	0,3082
	(0,0032)	(0,0110)	(0,0064)	(0,0199)	(0,0351)	(0,0457)

Tabela 5.5: Valores do PPP para os 56 itens em cada uma das 6 estatísticas do teste na prova simulada.

Vemos que, para os itens 49 a 52 (em que a CCI é a mistura de outras duas CCI's), apenas o item 51 obteve valores baixos de PPP para algumas estatísticas. Já para os itens 53 a 56 (em que a CCI é não monotônica), observamos valores baixos para algumas estatísticas para todos os itens.

Para os itens 1 a 48, cujas CCI's seguem o modelo logístico de 3 parâmetros, apenas os itens 6 e 33 apresentaram três estatísticas com valores menores que 0,05 para o PPP. O gráfico do Item 6 pode ser visto na Figura 5.9.

É interessante notar que esse item tem um valor de c estimado muito distante do valor real dele, o que se reflete em todas as regressões não paramétricas. Também é interessante notar que as diferenças que causaram esse valor de PPP baixo se encontram majoritariamente nos valores extremos de Theta.

5.2 Aplicação a dados reais

Para a segunda parte, foram aplicados os mesmos métodos à prova do SARESP. A tabela 5.6 apresenta os valores de PPP.

Todos os itens apresentaram valores PPP baixos para as estatísticas baseadas na regressão isotônica, enquanto apenas os itens 5 e 15 apresentaram valor PPP baixo para alguma estatística baseada na regressão de Nadaraya-Watson. Os gráficos das estimativas das CCI's destes dois itens são apresentados nas figuras 5.10 e 5.11.

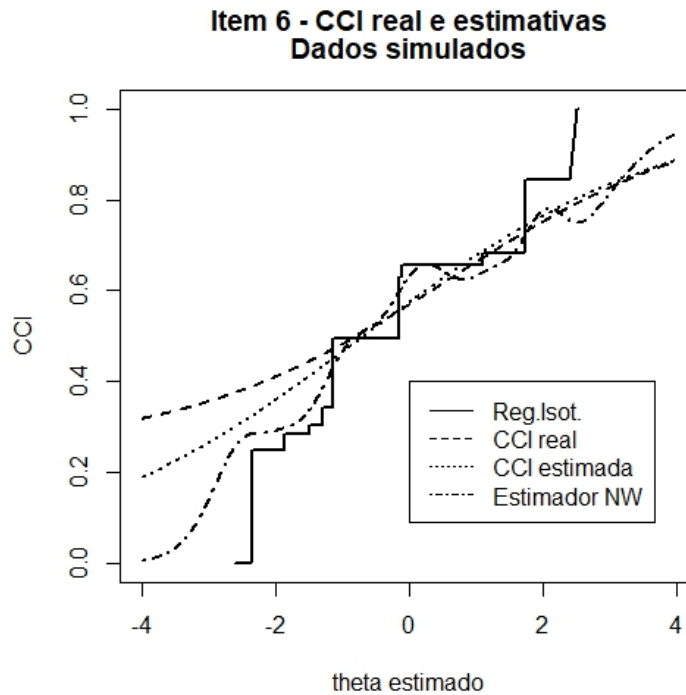


Figura 5.9: Item 6.

A tabela 5.7 apresenta as estimativas dos parâmetros dos itens.

Para esses resultados, assim como nos dados simulados, é importante notar que alguns valores de parâmetros ainda estão inadequados para o modelo, como os valores de c para os itens 17 e 18, por exemplo. Apesar de em bem menor quantidade, é possível notar que há uma alta probabilidade de termos esses itens retirados pelo processo completo de calibração.

Além disso, é possível ver que esses dois itens tem curvas de CCI muito deslocadas para os lados, com o valor do parâmetro b muito baixo no item 5 e muito alto no item 15.

Além disso, é interessante pontuar que, ao se ordenar os respondentes da menor proficiência para a maior, é possível ver que das 10 pessoas com menor proficiência, 7 acertaram o item 15, mesmo ele sendo um item com dificuldade muito alta.

Item	RIA	RIQ	RIAS	RIQS	RNWA	RNWQ
1	0,274	0,012	0,182	0,004	0,966	0,912
2	0,002	0,006	0,020	0,012	0,976	0,984
3	0,028	0,002	0,066	0,000	0,990	0,996
4	0,002	0,002	0,010	0,002	0,838	0,900
5	0,000	0,000	0,000	0,000	0,154	0,008
6	0,048	0,000	0,044	0,000	0,890	0,880
7	0,056	0,008	0,020	0,006	0,862	0,884
8	0,006	0,004	0,008	0,008	0,154	0,136
9	0,010	0,000	0,028	0,000	0,738	0,590
10	0,208	0,004	0,104	0,002	0,990	1,000
11	0,002	0,000	0,002	0,000	0,338	0,380
12	0,030	0,004	0,036	0,002	0,900	0,980
13	0,002	0,004	0,006	0,002	0,864	0,870
14	0,040	0,002	0,042	0,000	0,874	0,902
15	0,256	0,002	0,104	0,000	0,152	0,000
16	0,108	0,006	0,066	0,002	0,956	0,978
17	0,092	0,000	0,048	0,000	0,934	0,990
18	0,002	0,000	0,002	0,000	0,948	0,992
19	0,020	0,008	0,024	0,002	0,894	0,906
20	0,042	0,000	0,018	0,000	0,888	0,970
21	0,008	0,000	0,030	0,004	0,762	0,780
22	0,074	0,000	0,044	0,000	0,778	0,542
23	0,192	0,008	0,104	0,006	0,398	0,576
24	0,402	0,002	0,244	0,004	0,860	0,604
25	0,010	0,018	0,014	0,012	0,902	0,988
26	0,002	0,000	0,000	0,000	0,734	0,756
27	0,048	0,000	0,018	0,000	0,880	0,946
28	0,002	0,008	0,008	0,012	0,936	0,966
29	0,000	0,000	0,000	0,000	0,482	0,266
30	0,000	0,000	0,002	0,000	0,758	0,926

Tabela 5.6: Valores do PPP para os 56 itens em cada uma das 6 estatísticas do Teste na prova do SARESP.

Essas inconsistências podem ajudar a explicar os resultados de PPP obtidos pelo teste nesses itens específicos.

Além disso, é interessante notar como, para os dados reais, os estimadores de diferenças quadráticas baseados na Regressão Isotônica tiveram valores consistentemente menores do que os estimadores de diferença simples.

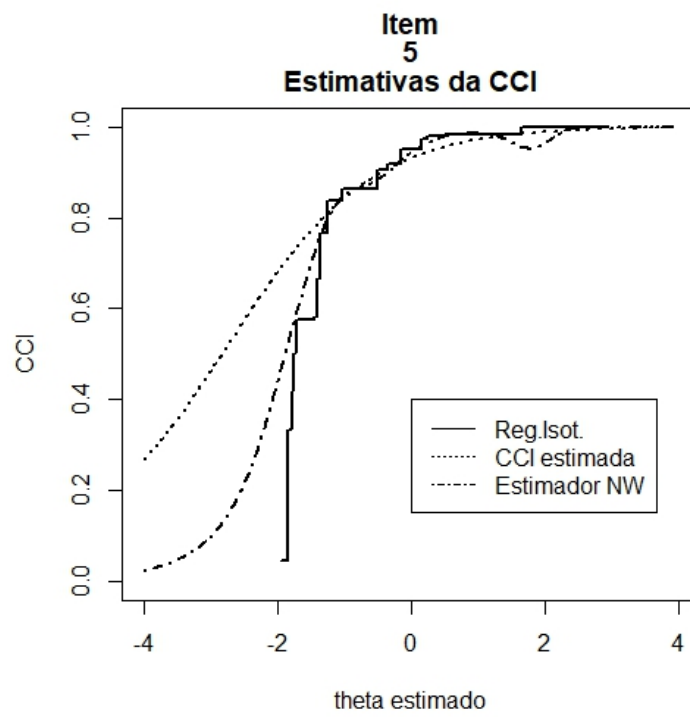


Figura 5.10: CCI's do Item 5.

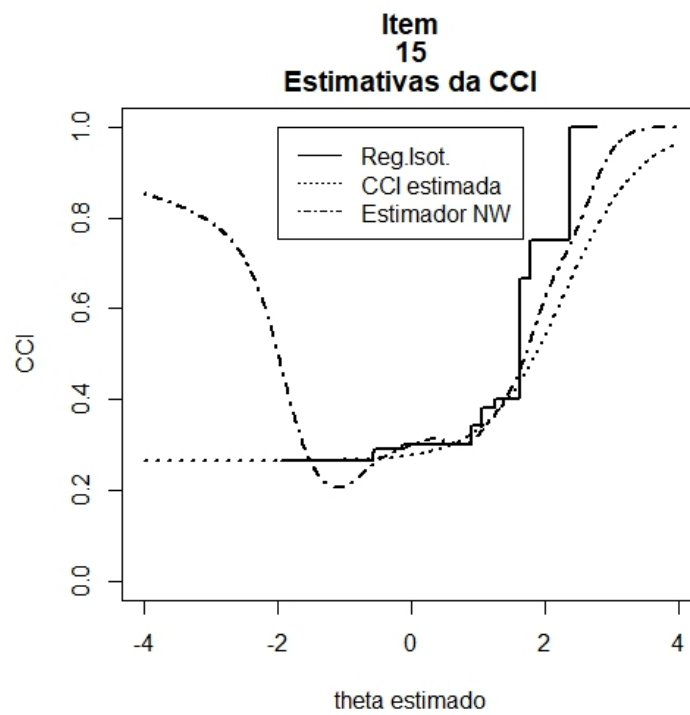


Figura 5.11: CCI's do Item 15.

Item	a	b	c
1	0,6315	1,9770	0,2425
2	1,8945	0,2532	0,1911
3	1,7179	1,1143	0,1560
4	1,0824	0,2527	0,1422
5	0,9375	-2,7291	0,0445
6	0,5703	-0,4024	0,0406
7	0,8111	0,7837	0,0973
8	0,5204	1,3510	0,0019
9	1,5216	1,0433	0,1749
10	1,4679	1,2921	0,3088
11	0,6025	-0,6377	0,0110
12	0,7708	0,9713	0,3231
13	1,1677	-0,1060	0,0392
14	1,2276	0,2330	0,2429
15	1,7473	2,2958	0,2649
16	1,6065	1,7083	0,1577
17	2,3149	0,7974	0,4926
18	3,1399	0,0014	0,7314
19	0,9553	0,4331	0,0751
20	1,3236	0,5333	0,5914
21	0,5840	0,9836	0,0026
22	2,8331	2,1955	0,1726
23	2,2500	2,3506	0,2147
24	0,7959	2,5865	0,4017
25	0,9639	0,9990	0,0032
26	1,3596	-0,2468	0,1921
27	1,9911	0,6057	0,5781
28	1,5150	0,3593	0,0661
29	1,2200	-1,0847	0,0027
30	1,3626	-0,6196	0,4261

Tabela 5.7: Parâmetros estimados dos 30 itens da Prova SARESP.

Capítulo 6

Conclusão e Trabalhos Futuros

6.1 Conclusão

Observa-se que o teste se comportou bem em não rejeitar a monotonicidade da CCI para os itens que realmente seguiam este modelo. Porém, apesar disso, os dados foram feitos em cima de uma amostra pequena, o que prejudicou os valores dos parâmetros de calibração como um todo.

Dessa forma, com esses itens sendo retirados num processo completo de calibração, não se teve tempo de analisar o que aconteceria com o teste se pegássemos um teste apenas com itens que tivessem valores de parâmetros adequados para o modelo.

Porém, apesar de ter conseguido detectar alguns itens que não seguiam o modelo logístico, a Estatística do teste que utiliza Regressão Isotônica não conseguiu detectar a maioria dos itens que não tem CCI's monotônicas.

Além disso, para os dados reais, essas mesmas estatísticas rejeitaram a grande maioria dos itens, principalmente as estatísticas de diferença quadrática. Dessa forma, se torna necessário que mais estudos sejam feitos sobre o tema.

6.2 **Trabalhos Futuros**

O próximo passo principal é refazer o trabalho com uma prova que tenha mais respondentes, principalmente a aplicação a dados reais. Uma opção interessante seria usar os dados abertos do ENEM para isso e fazer o processo completo de calibração, com retirada de itens com valores de parâmetros inadequados e análise de DIF, caso tenham grupos diferentes.

Além disso, sabendo-se da existência de muitos outros testes de adequabilidade e suas práticas comuns, é importante tentar otimizar a ideia presente neste trabalho, aplicando técnicas diferentes e comparando os resultados.

Uma possibilidade para fazer isso seria comparar esses resultados com outros tipos de processos, substituindo PPMC por algum outro método para gerar valores de a , b e c . Uma forma de fazer isso poderia ser simular uma prova com um valor n de respondentes para cada trinca de parâmetros obtidas no processo MCMC.

Além disso, uma outra sugestão poderia ser definir um intervalo quantílico para o θ e fazer os valores das diferenças considerarem apenas os valores dentro desse intervalo. Esse intervalo poderia ser de 5% a 95%, por exemplo, o que faria os pontos da CCI comparados terem mais respondentes e, portanto, teriam suas estimações paramétricas e não paramétricas mais precisas.

Capítulo 7

Referências Bibliográficas

Aitkin, Murray (2010).. Statistical inference: an integrated Bayesian/likelihood approach. **CRC Press**,

Andrade, D. F., Tavares, H. R. e Valle, R. C. (2000). **Teoria da resposta ao item: conceitos e aplicações**. Associação Brasileira de Estatística, Caxambu

Andrich, D. (1978). A rating formulation for ordered response categories. **Psychometrika**, 43(4), 561-573.

Baker, F. B., R. C. 1985. **The basics of item response theory** ERIC Clearinghouse on Assessment and Evaluation, College Park, MD,

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. **Psychometrika**, 37(1), 29-51.

Bolt, D. M. (2001). Conditional covariance-based representation of multidimensional test structure. **Applied Psychological Measurement**, 25, 244-257.

Brunk, H. D., et al. (1972). Statistical inference under order restrictions.(the theory and application of isotonic regression). **Missouri Univ Columbia Dept of Statistics**,

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. **Statistical theories of mental test scores**, pages 395–479.

Douglas, J., Cohen, A. (2001). Nonparametric item response function estimation for assessing parametric model fit. **Applied Psychological Measurement**, 25(3), 234–243.

Gelman, A., Meng, X. L., Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. **Statistica Sinica**, 6, 733–759.

Hastings WK. (1970). Monte Carlo sampling methods using Markov chains and their applications. **Biometrika** 57:97–109

Kuhfeld, M. (2019). A posterior predictive model checking method assuming posterior normality for item response theory. **Applied psychological measurement**, 43(2), 125–142.

Levy, R., Mislevy, R. J., Sinharay, S. (2009). Posterior predictive model checking for multidimensionality in item response theory. **Applied Psychological Measurement**, 33(7), 519–537.

Li, T., Xie, C., Jiao, H. (2017). Assessing fit of alternative unidimensional polytomous IRT models using posterior predictive model checking. **Psychological Methods**, 22(2), 397

Masters, G. N. (1982). A Rasch model for partial credit scoring. **Psychometrika**, 47(2), 149–174.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. (1953). Hastings WK. 1970. **Biometrika** 57:97–109. **The journal of chemical physics** 21:1087–1092

Muraki, E., Bock, R. D. (1996). **Irt based test scoring and item analysis for graded open-ended exercises and performance tasks.**

Orlando, M., Thissen, D. (2000). Likelihood-based item- t indices for dichotomous item response theory models. **Applied Psychological Measurement**, 24, 50-65.

Ramsay, J. (1991). "Kernel smoothing approaches to nonparametric item characteristic curve estimation". **Psychometrika**, 56, 611-630.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. **Psychometrika monograph supplement**.

Schlemper, B. N. (2014). Testes de adequabilidade de ajuste em Teoria da Resposta ao Item. Master's thesis, Universidade de Brasília.

Sinharay, S. (2005). Assessing fit of unidimensional item response theory models using a Bayesian approach. **Journal of Educational Measurement**, 42(4), 375-394.

Sharma, Sanjib (2017). "Markov chain Monte Carlo methods for Bayesian data analysis in astronomy." arXiv preprint arXiv:1706.01629.

Stone, C. A. (2000). Monte Carlo based null distribution for an alternative goodness-of-fit test statistic in IRT models. **Journal of Educational Measurement**, 37(1), 58-75.

Wang, W., Kingston, N. (2020). Using Bayesian Nonparametric Item Response Function Estimation to Check Parametric Model Fit. **Applied Psychological Measurement**, 44(5), 331-345.

Yen, W. (1981). Using simulation results to choose a latent trait model. **Applied Psychological Measurement**, 5, 245-262.

Apêndice A

Código em R

```
#####
```

```
### Código para estudo das seguintes estatísticas:
```

```
###
```

```
### 1) baseada no ENPMV da CCI de cada item obtido via
```

```
regressão ###
```

```
### isotônica para dados com status corrente, proposta por
```

```
###
```

```
### Antonio Eduardo Gomes;
```

```
###
```

```
### 2) baseada no estimador de Nadarajah–Watson para regressão
```

```
###
```

```
### linear simples, proposta por Ramsay (1991)
```

```
###
```

```
#####
```

```
#####
```

```
### FUNÇÕES para a estatística proposta por A.E.Gomes ###
```

```
#####
```

```
#####

### Função "npml1"
###
### Calcula o ENPMV de F (função de distribuição do tempo de
falha) ###
### para dados com censura intervalar, caso 1 (current status
data).###
### O vetor "fest" contém as estimativas de F(x) em cada ponto
###
### ordenado de "x".
###
#####

npml1 <- function(n,g) {
  w <- rep(1,n)
  gw <- g*w
  gwac <- cumsum(gw)
  wac <- cumsum(w)

  matslopes <- cbind(c(1:n),gwac/wac)
  ind <- indaux <- 1

  ### "fest" é o vetor com os valores estimados de F (EMV não
paramétrico)
  fest <- rep(0,n)

  while(ind < n) {
    ind <- max(matslopes[matslopes[,2]==min(matslopes[,2]),1])
    fest[indaux:ind] <- min(matslopes[,2])
    indaux <- ind+1
    matslopes <- cbind(c(indaux:n),(gwac[indaux:n]-gwac[ind])/(
      wac[indaux:n]-wac[ind]))
  }
  # print(matslopes)
}
```

```

    }
    return(fest)
}
#####

### Função "integ.abs"
###
### calcula a integral do valor absoluto da diferença entre
###
### o modelo de 3 parâmetros ajustado e o estimador não
### paramétrico ###
#####

integ.abs <- function(x){
  tf <- logical(0)
  for (j in 1:(length(t.salto)-1)) tf <- cbind(tf,((x>t.salto[j]
    ])*(x<=t.salto[j+1])))
  tf <- cbind(tf,(x>t.salto[length(t.salto)]))
  fe <- tf%%p.salto
  cci <- cf+(1-cf)/(1+exp(-af*(x-bf)))
  dif.abs <- sum(abs(cci-fe)*dnorm(x,0,1))*(x[2]-x[1])
  return(dif.abs)
}

#####

### Função "integ.quad"
###
### calcula a integral do quadrado da diferença entre
###
### o modelo de 3 parâmetros ajustado e o estimador não
### paramétrico ###
#####

```

```

integ.quad <- function(x){
  tf <- logical(0)
  for (j in 1:(length(t.salto)-1)) tf <- cbind(tf,((x>t.salto[j]
    ])*(x<=t.salto[j+1]))
  tf <- cbind(tf,(x>t.salto[length(t.salto)]))
  fe <- tf%%p.salto
  cci <- cf+(1-cf)/(1+exp(-af*(x-bf)))
  dif.quad <- sqrt(sum(((cci-fe)^2)*dnorm(x,0,1))*(x[2]-x[1]))
  return(dif.quad)
}

```

```
#####
```

```
### Função "integ.suav.abs"
```

```
###
```

```
### calcula a integral do valor absoluto da diferença entre
```

```
###
```

```
### o modelo de 3 parâmetros ajustado e o estimador não
```

```
paramétrico ###
```

```
### suavizado via núcleo estimador (kernel)
```

```
###
```

```
#####
```

```

integ.suav.abs <- function(x){
  mat1 <- matrix(rep(x,length(t.salto)),ncol=length(t.salto))
  mat2 <- matrix(rep(t.salto,rep(length(x),length(t.salto))),
    ncol=length(t.salto))
  mat.aux <- (mat1-mat2)/h
  F.suave <- (pnorm(mat.aux,0,1))%%(diff(c(0,p.salto)))
  cci <- cf+(1-cf)/(1+exp(-af*(x-bf)))
  dif.suav.abs <- sum((abs(cci-F.suave))*dnorm(x,0,1))*(x[2]-x
    [1])
  return(dif.suav.abs)
}

```

```

#####

### Função "integ.suav.quad"
###
### calcula a integral do quadrado da diferença entre
###
### o modelo de 3 parâmetros ajustado e o estimador não
### paramétrico ###
### suavizado via núcleo estimador (kernel)
###
#####

integ.suav.quad <- function(x){
  mat1 <- matrix(rep(x,length(t.salto)),ncol=length(t.salto))
  mat2 <- matrix(rep(t.salto,rep(length(x),length(t.salto))),
    ncol=length(t.salto))
  mat.aux <- (mat1-mat2)/h
  F.suave <- (pnorm(mat.aux,0,1))%*%(diff(c(0,p.salto)))
  cci <- cf+(1-cf)/(1+exp(-af*(x-bf)))
  dif.suav.quad <- sqrt(sum(((cci-F.suave)^2)*dnorm(x,0,1))*(x
    [2]-x[1]))
  return(dif.suav.quad)
}

#####
### FUNÇÃO para a estatística proposta por Ramsay (1991) ###
#####

#####
### Função "integ.abs.nw"
###
### calcula a integral do valor absoluto da diferença entre
###

```

```

#### o modelo de 3 par metros ajustado e o estimador proposto
####
#### por Ramsay (1991) para os dados originais #####
#####
integ.abs.nw <- function(x){
  mat1 <- matrix(rep(x,no.resp),ncol=no.resp)
  mat2 <- matrix(rep(theta.est,rep(length(x),no.resp)),ncol=no.
    resp)
  mat.aux <- (mat1-mat2)/h
  kern <- dnorm(mat.aux,0,1)
  fe <- (kern%*%resp)/(apply(kern,1,sum))
  cci <- cf+(1-cf)/(1+exp(-af*(x-bf)))
  dif.abs <- sum(abs(cci-fe)*dnorm(x,0,1))*(x[2]-x[1])
  return(dif.abs)
}

#####
#### Função "integ.quad.nw"
####
#### calcula a integral do quadrado da diferen a entre
####
#### o modelo de 3 par metros ajustado e o estimador proposto
####
#### por Ramsay (1991) para os dados originais #####
#####
integ.quad.nw <- function(x){
  mat1 <- matrix(rep(x,no.resp),ncol=no.resp)
  mat2 <- matrix(rep(theta.est,rep(length(x),no.resp)),ncol=no.
    resp)
  mat.aux <- (mat1-mat2)/h
  kern <- dnorm(mat.aux,0,1)
  fe <- (kern%*%resp)/(apply(kern,1,sum))
  cci <- cf+(1-cf)/(1+exp(-af*(x-bf)))
  dif.quad <- sqrt(sum(((cci-fe)^2)*dnorm(x,0,1))*(x[2]-x[1]))

```



```

    return(dif.quad)
  }

#####
### Função "integ.abs.nw.ppp"
###
### calcula a integral do valor absoluto da diferença entre
###
### o modelo de 3 parâmetros ajustado e o estimador proposto
###
### por Ramsay (1991) para os dados das amostras MCMC      ###
#####
integ.abs.nw.ppp <- function(x){
  mat1 <- matrix(rep(x,no.resp),ncol=no.resp)
  mat2 <- matrix(rep(am.theta[j,],rep(length(x),no.resp)),ncol=
    no.resp)
  mat.aux <- (mat1-mat2)/h
  kern <- dnorm(mat.aux,0,1)
# fe <- (kern%*%resp)/(apply(kern,1,sum))
fe <- (kern%*%dados.m[,i])/(apply(kern,1,sum))
cci <- cf+(1-cf)/(1+exp(-af*(x-bf)))
dif.abs <- sum(abs(cci-fe)*dnorm(x,0,1))*(x[2]-x[1])
  return(dif.abs)
}

#####
### Função "integ.quad.nw.ppp"
###
### calcula a integral do quadrado da diferença entre
###
### o modelo de 3 parâmetros ajustado e o estimador proposto
###
### por Ramsay (1991) para os dados das amostras MCMC      ###
#####

```

```

integ.quad.nw.ppp <- function(x){
  mat1 <- matrix(rep(x,no.resp),ncol=no.resp)
  mat2 <- matrix(rep(am.theta[j,],rep(length(x),no.resp)),ncol=
    no.resp)
  mat.aux <- (mat1-mat2)/h
  kern <- dnorm(mat.aux,0,1)
# fe <- (kern%c*%resp)/(apply(kern,1,sum))
fe <- (kern*%dados.m[,i])/(apply(kern,1,sum))
cci <- cf+(1-cf)/(1+exp(-af*(x-bf)))
dif.quad <- sqrt(sum(((cci-fe)^2)*dnorm(x,0,1))*(x[2]-x[1]))
return(dif.quad)
}

#####
### SIMULA O ###
#####

library(irtoys)
library(ltm)
library(mirt)

## Definição dos valores dos parâmetros para geração dos
  dados

set.seed(23456) # semente

no.item <- 56 # numero de itens
no.resp <- 1000 # numero de respondentes
aa <- rep(seq(0.5,2,0.3),rep(8,6)) # vetor de parâmetros de
  discriminacao
bb <- rep(seq(-1.4,1.4,0.4),6) # vetor de parâmetros de
  dificuldade
cc <- rep(c(0.2,0.25),24)

```

```

pa <- cbind(aa,bb,cc)

# proficiencias reais

theta <- rnorm(no.resp) # proficiencias reais

# geracao dos dados

dados.sim <- sim(ip=pa,x=theta)

##### Itens adicionais que nao seguem o ML3

at1 <- c( 0.5, 0.5, 3, 3)
at2 <- c( 3, 3, 0.5, 0.5)
bt1 <- c(-1.5,-2, -1.5,-2)
bt2 <- c( 1.5, 2, 1.5, 2)
ct <- 0.2

## Mistura de duas CCI's com ML3
for (i in 1:4) {
  prob.H1 <- ct + (1-ct)*(1/(1+exp(-at1[i]*(theta-bt1[i])))
    +1/(1+exp(-at2[i]*(theta-bt2[i]))))/2
  dados.sim <- cbind(dados.sim,sign(sign(prob.H1-runif(no.resp
    ,0,1))+1))
}

##### Nonmonotonic

at <- c( 1.5, 3, 2, 3)
bt <- c( 0, 1, 0, 0)
dt <- c( 0, -1, 0, 1)

# i=1 OK
# i=2 OK

```

```

# i=3 OK
# i=4 OK

for (i in 1:4) {
  prob.H1 <- ct + (1-ct)/(1+exp(-0.5*at[i]*((theta-dt[i])^3-(
    theta-dt[i])-bt[i])))
  dados.sim <- cbind(dados.sim, sign(sign(prob.H1-runif(no.resp
    ,0,1))+1))
}

#### Ajuste dos modelos param tricos

dados.tpm <- tpm(dados.sim)
par.est <- coef(dados.tpm) # cc, bb, aa

#### estimacao da proficiencia
theta.est <- eap(dados.sim, cbind(par.est[,3], par.est[,2], par.
  est[,1]), qu=normal.qu())

#### ordenacao dos dados dos respondentes em ordem crescente pelo
  valor estimado de theta
dados.sim.y <- dados.sim
dados.sim <- dados.sim[order(theta.est[,1]),]
theta <- theta[order(theta.est[,1])]
theta.est <- theta.est[order(theta.est[,1]),]
theta.est <- theta.est[,1]

mat.enpmv <- matrix(0, no.resp, no.item)
dist.abs <- dist.quad <- dist.suav.abs <- dist.suav.quad <- rep
  (0, no.item)

#### dist.abs = distancia L1 entre a CCI param trica ajustada e
  o ENPMV
#### dist.quad = distancia L2 entre a CCI param trica ajustada

```

```

e o ENPMV
#### dist.suav.abs = dist ncia L1 entre a CCI param trica
ajustada e o ENPMV suavizado via kernel
#### dist.suav.quad = dist ncia L2 entre a CCI param trica
ajustada e o ENPMV suavizado via kernel

h <- 1.1/(no.resp^0.2)

for (i in 1:no.item) {
  mat.enpmv[,i] <- npml1(no.resp,dados.sim[,i])
  mat.enpmv[mat.enpmv[,i]<=par.est[i,1],i] <- par.est[i,1]

#### C lculo das estat sticas de teste

  ind.aux <- cumsum(table(mat.enpmv[,i]))+1
  ind.aux <- c(1,ind.aux[ind.aux < no.resp])
  t.salto <- theta.est[c(ind.aux,no.resp)] # valores de escore.
  padrao onde ocorrem saltos do enpmv
  p.salto <- mat.enpmv[c(ind.aux,no.resp),i] # valores do enpmv
  nos pontos da proficiencia estimada onde ocorrem os saltos

  af <- par.est[i,3]
  bf <- par.est[i,2]
  cf <- par.est[i,1]

  x <- seq(-4,4,0.01)

  dist.abs[i] <- integ.abs(x)
  dist.quad[i] <- integ.quad(x)
  dist.suav.abs[i] <- integ.suav.abs(x)
  dist.suav.quad[i] <- integ.suav.quad(x)

}

```

```

#### C lculo das estat sticas para o estimador proposto por
      Ramsay(1991)

h <- 1.1/(no.resp^0.2)

dist.abs.nw <- dist.quad.nw <- rep(0,no.item)

for (i in 1:no.item) {

  af <- par.est[i,3]
  bf <- par.est[i,2]
  cf <- par.est[i,1]

  resp <- dados.sim[,i]
  dist.abs.nw[i] <- integ.abs.nw(x)
  dist.quad.nw[i] <- integ.quad.nw(x)

}

#####
#### Gr ficos ilustrativos
#####

# mistura
at1 <- c( 0.5, 0.5, 3, 3)
at2 <- c( 3, 3, 0.5, 0.5)
bt1 <- c(-1.5,-2, -1.5,-2)
bt2 <- c( 1.5, 2, 1.5, 2)
ct <- 0.2

# Nonmonotonic
at <- c( 1.5, 3, 2, 3)
bt <- c( 0, 1, 0, 0)
dt <- c( 0, -1, 0, 1)

```

```

i <- 33 # n mero do item utilizado como exemplo no gr fico

resp <- dados.sim[, i]
mat1 <- matrix(rep(x, no.resp), ncol=no.resp)
mat2 <- matrix(rep(theta.est, rep(length(x), no.resp)), ncol=no.
  resp)
mat.aux <- (mat1-mat2)/h
kern <- dnorm(mat.aux, 0, 1)
fe <- (kern%*%resp)/(apply(kern, 1, sum))
cci <- cc[i] + (1-cc[i])/(1+exp(-aa[i]*(x-bb[i])))
#cci <- ct + (1-ct)*(1/(1+exp(-at1[i-48]*(x-bt1[i-48]))) + 1/(1+
  exp(-at2[i-48]*(x-bt2[i-48]))))/2
#cci <- ct + (1-ct)/(1+exp(-0.5*at[i-52]*((x-dt[i-52])^3-(x-dt[i
  -52])-bt[i-52])))
cci.est <- par.est[i, 1] + (1-par.est[i, 1])/(1+exp(-par.est[i
  , 3]*(x-par.est[i, 2])))

plot(theta.est, mat.enpmv[, i], type="l", xlim=c(-4, 4), ylim=c(0, 1),
  lwd=2,
xlab=c("theta estimado"), ylab=c("CCI"), main=c("Item 33 - CCI
  real e estimativas", "Dados simulados"))
lines(x, cci, lty=2, lwd=2)
lines(x, cci.est, lty=3, lwd=2)
lines(x, fe, lty=4, lwd=2)
legend(0, 0.4, c("Reg. Isot.", "CCI real", "CCI estimada", "Estimador
  NW"), lty = c(1, 2, 3, 4))

#####
## MCMC
#####

### omega = valor atual do parametro
### phi = novo valor do parametro

```

```
#####
### Funcao "qtr.a" calcula o valor do kernel de transicao
#####
qtr.a <- function(par.novo, par.atual, dp) {
  dlnorm(par.novo, par.atual, dp)
}
#####
### Funcao "qtr.b" calcula o valor do kernel de transicao
#####
qtr.b <- function(par.novo, par.atual, dp) {
  dnorm(par.novo, par.atual, dp)
}
#####
### Funcao "qtr.c" calcula o valor do kernel de transicao
#####
qtr.c <- function(par.novo, par.atual, dp) {
  dbeta(par.novo, aux.aaa, aux.bbb)
}
#####
### Funcao "qtr.theta" calcula o valor do kernel de transicao
#####
qtr.theta <- function(par.novo, par.atual, dp) {
  dnorm(par.novo, par.atual, dp)
}

#####
### Funcao "pri.a" calcula a priori de "a"
#####
pri.a <- function(par) {
  dlnorm(par, 0, 1)
}

#####
```



```

#### Funcao "pri.b" calcula a priori de "b"
#####
pri.b <- function(par) {
  dnorm(par,0,2)
}

#####
#### Funcao "pri.c" calcula a priori de "c"
#####
pri.c <- function(par) {
  dbeta(par,51,204)
}

#####
#### Funcao "pri.theta" calcula a priori de "theta"
#####
pri.theta <- function(par) {
  dnorm(par,0,1)
}

#####
#### Funcao "lik" calcula a verossimilhanca para cada item
#####
lik <- function(dados,aa,bb,cc,theta) {
  prob <- veross <- matrix(0,no.resp,no.item)
  for (i in 1:no.item) {
    prob[,i] <- cc[i]+(1-cc[i])/(1+exp(-aa[i]*(theta-bb[i])))
    veross[,i] <- (prob[,i]^dados[,i])*((1-prob[,i])^(1-dados[,i]
    ))
  }
  veross.prod <- apply(veross,2,prod)
  return(veross.prod)
}

```

```
#####

### Funcao "lik.theta" calcula a verossimilhanca para
    respondente
#####

lik.theta <- function(dados,aa,bb,cc,theta) {
  prob <- veross <- matrix(0,no.resp,no.item)
  for (k in 1:no.resp) {
    prob[k,] <- cc+(1-cc)/(1+exp(-aa*(theta[k]-bb)))
    veross[k,] <- (prob[k,]^dados[k,])*((1-prob[k,])^(1-dados[k,]))
  }
  veross.prod <- apply(veross,1,prod)
  return(veross.prod)
}

#####
no.repet <- 25100
mat.a.mcmc <- mat.b.mcmc <- mat.c.mcmc <- matrix(0,no.repet,no.item)
mat.theta.mcmc <- matrix(0,no.repet,no.resp)

j <- 0
vec.a.j <- as.numeric(par.est[,3])
vec.b.j <- as.numeric(par.est[,2])
vec.c.j <- as.numeric(par.est[,1])
vec.theta.j <- as.numeric(theta.est)

sigma.a <- 0.01
sigma.b <- 0.01
sigma.c <- 0.001
sigma.theta <- 0.01
```

```

while (j<no.repet) {
  j <- j+1
print(j)

#####
## geracao de a
#####

vec.a.star <- rnorm(no.item,vec.a.j,0.01)
vec.a.star[vec.a.star <=0.4] <- 0.4
vec.a.star[vec.a.star >=3] <- 3

aux.a <- (lik(dados.sim,vec.a.star,vec.b.j,vec.c.j,vec.theta.j)
  )*pri.a(vec.a.star)*
  qtr.a(vec.a.star,vec.a.j,sigma.a))/
  (lik(dados.sim,vec.a.j,vec.b.j,vec.c.j,vec.theta.j)*
  pri.a(vec.a.j)*
  qtr.a(vec.a.j,vec.a.star,sigma.a))

p.aceit <- apply(cbind(rep(1,no.item),aux.a),1,min)

z <- runif(no.item,0,1)
vec.a.j <- vec.a.star^sign(sign(p.aceit-z)+1)*vec.a.j^sign(
  sign(z-p.aceit)+1)

mat.a.mcmc[j,] <- vec.a.j

#####
## geracao de b
#####

vec.b.star <- rnorm(no.item,vec.b.j,0.01)

aux.b <- (lik(dados.sim,vec.a.j,vec.b.star,vec.c.j,vec.theta.j)

```

```

    )*pri.b(vec.b.star)*
      qtr.b(vec.b.star,vec.b.j,sigma.b))/
      (lik(dados.sim,vec.a.j,vec.b.j,vec.c.j,vec.theta.j)*
        pri.b(vec.b.j)*
        qtr.b(vec.b.j,vec.b.star,sigma.b))

p.aceit <- apply(cbind(rep(1,no.item),aux.b),1,min)

z <- runif(no.item,0,1)
vec.b.j <- vec.b.star^sign(sign(p.aceit-z)+1)*vec.b.j^sign(
  sign(z-p.aceit)+1)

mat.b.mcmc[j,] <- vec.b.j

#####
## geracao de c
#####

vec.c.star <- abs(rnorm(no.item,vec.c.j,0.01))

aux.c <- (lik(dados.sim,vec.a.j,vec.b.j,vec.c.star,vec.theta.j)
  )*pri.c(vec.c.star)*
  qtr.c(vec.c.star,vec.c.j,sigma.c))/
  (lik(dados.sim,vec.a.j,vec.b.j,vec.c.j,vec.theta.j)*
    pri.c(vec.c.j)*
    qtr.c(vec.c.j,vec.c.star,sigma.c))

p.aceit <- apply(cbind(rep(1,no.item),aux.c),1,min)

z <- runif(no.item,0,1)
vec.c.j <- vec.c.star^sign(sign(p.aceit-z)+1)*vec.c.j^sign(
  sign(z-p.aceit)+1)

mat.c.mcmc[j,] <- vec.c.j

```

```
#####
## geracao de theta
#####

vec.theta.star <- rnorm(no.resp, vec.theta.j, 0.01)

aux.theta <- (lik.theta(dados.sim, vec.a.j, vec.b.j, vec.c.j,
  vec.theta.star)*pri.theta(vec.theta.star)*
  qtr.theta(vec.theta.star, vec.theta.j, sigma.
    theta))/
  (lik.theta(dados.sim, vec.a.j, vec.b.j, vec.c.j,
    vec.theta.j)*pri.theta(vec.theta.j)*
  qtr.theta(vec.theta.j, vec.theta.star, sigma.
    theta))

p.aceit <- apply(cbind(rep(1, no.resp), aux.theta), 1, min)

z <- runif(no.resp, 0, 1)
vec.theta.j <- vec.theta.star^sign(sign(p.aceit-z)+1)*vec.
  theta.j^sign(sign(z-p.aceit)+1)

mat.theta.mcmc[j,] <- vec.theta.j

} # j = 1, no.repet

#####
## Determina o do burn-in
#####

## Determina o do periodo

a.cor <- function(x, k) {
  ## x = vetor de dados
```

```
## k = lag maximo
vec <- rep(0,k)
for (ind in 1:k) vec[ind] <- cor(x[1:(length(x)-ind)],x[(ind+1)
  :length(x)])
return(vec)
}

#### Verificacao da autocorrelacao para parametros dos itens "a",
  "b" e "c"

x <- mat.c.mcmc
attach(x)

no.item <- ncol(x)
lag.max <- 100
auto.cor <- matrix(0,lag.max,no.item)

compr <- rep(0,no.item)
for (i in 1:no.item) {
  auto.cor[,i] <- a.cor(x[,i],lag.max)
  compr[i] <- sum(sign(sign(auto.cor[,i]-0.2)+1))
}

max(compr) # maior valor de lag com autocorrelacao maior que 0.2
  (valido para "a", "b" e "c")

#### Verificacao da autocorrelacao para a proficiencia "theta"

x <- mat.theta.mcmc
attach(x)

no.resp <- ncol(x)
lag.max <- 100
auto.cor <- matrix(0,lag.max,no.resp)
```

```

compr <- rep(0,no.resp)
for (i in 1:no.resp) {
  auto.cor[,i] <- a.cor(x[,i],lag.max)
  compr[i] <- sum(sign(sign(auto.cor[,i]-0.2)+1))
}

max(compr) # maior valor de lag com autocorrelacao maior que 0.2
            (valido para "theta")

i <- 10
plot(auto.cor[,i],type="l",main=c("Respondente", i, "
  Autocorrelações"),xlab=c("Lag"),ylab=c("autocorrelação"),
ylim=c(0,1))
abline(h=0.2)

#####

lag.def <- 50
burn.in <- 100 # graficamente, houve indicacao de burn-in apenas
               para "b".
#####

#####
### Amostra MCMC
#####

lag.def <- 50
burn.in <- 100 # graficamente, houve indicacao de burn-in apenas
               para "b".
aux.ind <- burn.in + c(1:((no.repet-burn.in)/lag.def))*lag.def

no.amost <- (no.repet - burn.in)/lag.def # tamanho da amostra
MCMC

```

```

am.a <- as.matrix(mat.a.mcmc[aux.ind,])
am.b <- as.matrix(mat.b.mcmc[aux.ind,])
am.c <- as.matrix(mat.c.mcmc[aux.ind,])
am.theta <- as.matrix(mat.theta.mcmc[aux.ind,])

mat.dist.abs.m <- mat.dist.quad.m <- mat.dist.suav.abs.m <- mat.
  dist.suav.quad.m <- matrix(0,no.amost,no.item)
mat.dist.abs.nw.m <- mat.dist.quad.nw.m <- matrix(0,no.amost,no.
  item)

for (j in 1:no.amost) {

  print(j)

  #### Definição dos valores dos parâmetros para geração da
    amostra MCMC
  par.mcmc <- cbind(am.a[j,],am.b[j,],am.c[j,])
  theta.m <- sort(am.theta[j,])
  dados.m <- sim(ip=par.mcmc,x=theta.m)

  #### matriz com as estimativas no ENPMV via regressão para
    cada item
  mat.enpmv <- matrix(0,no.resp,no.item)
  dist.abs.m <- dist.quad.m <- dist.suav.abs.m <- dist.suav.quad
    .m <- rep(0,no.item)

  #### dist.abs.m = distância L1 entre a CCI paramétrica
    ajustada e o ENPMV
  #### dist.quad.m = distância L2 entre a CCI paramétrica
    ajustada e o ENPMV
  #### dist.suav.abs.m = distância L1 entre a CCI paramétrica
    ajustada e o ENPMV suavizado via kernel
  #### dist.suav.quad.m = distância L2 entre a CCI paramétrica
    ajustada e o ENPMV suavizado via kernel

```



```

# parametro de suavizacao utilizado na estatistica que utiliza
  estimadores suavizados
h <- 1.1/(no.resp^0.2)

dist.abs.nw.m <- dist.quad.nw.m <- rep(0,no.item)

for (i in 1:no.item) {

  mat.enpmv[,i] <- npml1(no.resp,dados.m[,i])
  mat.enpmv[mat.enpmv[,i]<=par.mcmc[i,3],i] <- par.mcmc[i,3]

  ### C lculo das estat sticas de teste para a amostra MCMC

  ind.aux <- cumsum(table(mat.enpmv[,i]))+1
  ind.aux <- c(1,ind.aux[ind.aux < no.resp])
  t.salto <- theta.m[c(ind.aux,no.resp)] # valores de escore.
  padrao onde ocorrem saltos do ENPMV
  p.salto <- mat.enpmv[c(ind.aux,no.resp),i] # valores do
  enpmv nos pontos da proficiencia estimada
  # onde ocorrem os
  saltos

  af <- par.mcmc[i,1]
  bf <- par.mcmc[i,2]
  cf <- par.mcmc[i,3]

  x <- seq(-4,4,0.01)

  dist.abs.m[i] <- integ.abs(x)
  dist.quad.m[i] <- integ.quad(x)
  dist.suav.abs.m[i] <- integ.suav.abs(x)
  dist.suav.quad.m[i] <- integ.suav.quad(x)

} # i in 1:no.item

```

```

#### C lculo das estat sticas para o estimador proposto por
      Ramsay(1991) para a amostra MCMC

for (i in 1:no.item) {

  af <- par.mcmc[i,1]
  bf <- par.mcmc[i,2]
  cf <- par.mcmc[i,3]

  resp <- dados.m[,i]
  dist.abs.nw.m[i] <- integ.abs.nw.ppp(x)
  dist.quad.nw.m[i] <- integ.quad.nw.ppp(x)

} # i in 1:no.item

mat.dist.abs.m[j,] <- dist.abs.m
mat.dist.quad.m[j,] <- dist.quad.m
mat.dist.suav.abs.m[j,] <- dist.suav.abs.m
mat.dist.suav.quad.m[j,] <- dist.suav.quad.m
mat.dist.abs.nw.m[j,] <- dist.abs.nw.m
mat.dist.quad.nw.m[j,] <- dist.quad.nw.m

} # j in 1:no.amost

#### C lculo do PPP para a estat stica proposta por A.E.Gomes

#### Distancia absoluta

aux.mat <- t(matrix(rep(dist.abs,no.amost),ncol=no.amost))
aux.dif <- mat.dist.abs.m-aux.mat
mat.ppp <- sign(sign(aux.dif)+1)
p.ppp.dist.abs <- apply(mat.ppp,2,sum)/no.amost

```

```
#### Distancia quadr tica
```

```
aux.mat <- t(matrix(rep(dist.quad,no.amost),ncol=no.amost))
aux.dif <- mat.dist.quad.m-aux.mat
mat.ppp <- sign(sign(aux.dif)+1)
p.ppp.dist.quad <- apply(mat.ppp,2,sum)/no.amost
```

```
#### Distancia absoluta suavizada
```

```
aux.mat <- t(matrix(rep(dist.suav.abs,no.amost),ncol=no.amost))
aux.dif <- mat.dist.suav.abs.m-aux.mat
mat.ppp <- sign(sign(aux.dif)+1)
p.ppp.dist.suav.abs <- apply(mat.ppp,2,sum)/no.amost
```

```
#### Distancia quadratica suavizada
```

```
aux.mat <- t(matrix(rep(dist.suav.quad,no.amost),ncol=no.amost))
aux.dif <- mat.dist.suav.quad.m-aux.mat
mat.ppp <- sign(sign(aux.dif)+1)
p.ppp.dist.suav.quad <- apply(mat.ppp,2,sum)/no.amost
```

```
#### C lculo do p-valor para a estat stica proposta por Ramsay
(1991)
```

```
#### Distancia absoluta
```

```
aux.mat <- t(matrix(rep(dist.abs.nw,no.amost),ncol=no.amost))
aux.dif <- mat.dist.abs.nw.m-aux.mat
mat.ppp <- sign(sign(aux.dif)+1)
p.ppp.dist.abs.nw <- apply(mat.ppp,2,sum)/no.amost
```

```
#### Distancia quadr tica
```

```
aux.mat <- t(matrix(rep(dist.quad.nw,no.amost),ncol=no.amost))
```

```
aux.dif <- mat.dist.quad.nw.m-aux.mat
mat.ppp <- sign(sign(aux.dif)+1)
p.ppp.dist.quad.nw <- apply(mat.ppp,2,sum)/no.amost

print(cbind(p.ppp.dist.abs,p.ppp.dist.quad,p.ppp.dist.suav.abs,
p.ppp.dist.suav.quad,p.ppp.dist.abs.nw,p.ppp.dist.quad.nw))

print(cbind(p.ppp.dist.abs[49:56],p.ppp.dist.quad[49:56],p.ppp.
dist.suav.abs[49:56],
p.ppp.dist.suav.quad[49:56],p.ppp.dist.abs.nw[49:56],p.ppp.dist.
quad.nw[49:56]))
par.est[49:56,]
```