



Universidade de Brasília  
Departamento de Estatística

**UTILIZAÇÃO DA TEORIA DA RESPOSTA AO ITEM NA  
AVALIAÇÃO DAS QUESTÕES DE FÍSICA DO VESTIBULAR DA UnB**

**Rubens Mateus da Cunha**

Projeto apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília  
2022**

**Rubens Mateus da Cunha**

**UTILIZAÇÃO DA TEORIA DA RESPOSTA AO ITEM NA  
AVALIAÇÃO DAS QUESTÕES DE FÍSICA DO VESTIBULAR DA UnB**

Orientador(a): Prof. Antonio Eduardo Gomes

Projeto apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília  
2022**

## Resumo

O objetivo deste trabalho é realizar um estudo dos itens de física do vestibular do segundo semestre de 2014 da Universidade de Brasília - UnB, realizado pelo CEBRASPE. Tal estudo foi feito aplicando técnicas da Teoria de Resposta ao Item - TRI, como modelos logísticos e de resposta gradual. Todo o trabalho foi realizado com a utilização do *software* estatístico R.

**Palavras chaves:** Teoria da resposta ao item, Modelos de logísticos, Modelo de resposta gradual, Estudo dos itens.

## Abstract

The objective of this work is to carry out a study of the physics items of the entrance exam of the second semester of 2014 of the University of Brasília - UnB, carried out by CEBRASPE. This study was carried out by applying techniques from the Item Response Theory - IRT, such as logistical and gradual response models. All the work carried out using the statistical software R.

**Keywords:** Item response theory, Logistic models, Graded response model, Item study.

## Sumário

<b>1 Introdução</b> . . . . .	4
<b>2 Metodologia</b> . . . . .	5
2.1 Banco de Dados . . . . .	5
2.2 Teoria de resposta ao item . . . . .	5
2.2.1 Modelos logísticos para itens dicotômicos ou dicotomizados . . . . .	5
2.2.2 Modelo de resposta gradual . . . . .	6
2.2.3 Curva característica do item . . . . .	7
2.2.4 Estimação dos parâmetros dos itens . . . . .	7
2.2.5 Estimação da habilidade . . . . .	8
2.3 Coeficiente de correlação ponto-bisserial . . . . .	9
<b>3 Resultados</b> . . . . .	10
3.1 Análise descritiva . . . . .	10
3.2 Ajuste dos modelos logísticos . . . . .	16
3.2.1 Ajuste do modelo logístico de 3 parâmetros . . . . .	16
3.2.2 Ajuste do modelo logístico de 2 parâmetros . . . . .	21
3.3 Modelo de resposta gradual . . . . .	32
<b>4 Conclusão</b> . . . . .	37
<b>Referências</b> . . . . .	38

# 1 Introdução

A teoria de resposta ao item (TRI) surgiu com o intuito de solucionar algumas limitações apresentadas na Teoria Clássica dos Teste (TCT). Os primeiros modelos de TRI surgiram na década de 50. Andrade, Tavares e Valle (2000), definem a TRI como um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da habilidade (ou habilidades) do respondente. Essa relação é sempre expressa de tal forma que quanto maior a habilidade maior a probabilidade de acerto no item.

Uma das principais características da TRI é que ela tem como elementos centrais os itens, e não a prova como um todo. Dessa forma, é possível propor modelos para traços latentes, ou seja, características do indivíduo que não podem ser observadas diretamente. A TRI permite a comparação do desempenho de indivíduos independentemente se realizaram a mesma avaliação, além de possibilitar a separação por níveis de habilidades ou escala de conhecimento.

No Brasil a TRI é utilizada desde 1995, aplicada por exemplo nas provas do Sistema Nacional de Avaliação da Educação Básica (Saeb), Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo - SARESP e na maior e mais popular prova realizada no país, o Exame Nacional do Ensino Médio (ENEM).

O vestibular da Universidade de Brasília é um exemplo das diversas provas realizadas em todo o país, mas ao contrário das avaliações supracitadas, tal vestibular leva em consideração se o item foi respondido corretamente ou não para calcular o escore final. Diante disso e das várias possibilidades que a TRI oferece para a obtenção de resultados além da resposta correta, um estudo mais aprofundado dos itens pode apresentar resultados interessantes.

## 2 Metodologia

O desenvolvimento desse trabalho se deu pela aplicação de técnicas da TRI nos itens do vestibular da UnB. Tal estudo ocorreu em três principais etapas. A primeira foi a análise descritiva, com intuito de entender como cada item se comportava inicialmente. Na segunda, ajustaram-se os modelos logísticos de 2 e 3 parâmetros. Por fim, na terceira etapa ajustou-se um modelo de resposta gradual.

O banco de dados e as técnicas utilizadas nesse trabalho serão descritos a seguir. Ressalta-se ainda que todo o projeto foi realizado com o *software* estatístico R, com o auxílio de Anjos e Andrade (2012) e Baker e Kim (2017)

### 2.1 Banco de Dados

O banco de dados utilizado contém as respostas de parte dos candidatos do vestibular da UNB 2014/2 e foi disponibilizado pelo Centro Brasileiro de Pesquisa em Avaliação e Seleção e de Promoção de Eventos (CEBRASPE). Estão presentes as respostas de 7232 candidatos para cada uma das 34 questões de física do caderno de provas 1. Tais itens são caracterizados como tipo A (certo/errado) e tipo C (múltipla escolha com quatro alternativas).

### 2.2 Teoria de resposta ao item

A escolha das técnicas da TRI dependem de características da população estudada. Para esse trabalho, temos apenas um grupo de estudo, representado pelos candidatos que participaram do vestibular e responderam às mesmas questões. Consideraremos, inicialmente, os itens dicotômicos ou dicotomizados.

#### 2.2.1 Modelos logísticos para itens dicotômicos ou dicotomizados

Segundo Andrade, Tavares e Valle (2000) os modelos logísticos são os mais utilizados na TRI. Os três principais modelos utilizados se diferenciam pelo número de parâmetros adotados para descrever o item. O modelo de 3 parâmetros é o mais completo e o mais utilizado.

##### Modelo logístico de 3 parâmetros

O modelo logístico unidimensional de 3 parâmetros (ML3) considera a dificuldade,

a discriminação e a probabilidade de acerto ao acaso do item, e é dado por,

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-D_{ai}(\theta_j - b_i)}} \quad (2.2.1)$$

com  $i=1,2,\dots,I$  e  $j=1,2,\dots,n$ , onde

- $U_{ij}$ : é uma variável dicotômica que assume os valores 1, quando o indivíduo  $j$  responde corretamente o item  $i$ , ou 0 quando o indivíduo  $j$  não responde corretamente ao item  $i$ .
- $\theta_j$ : representa a habilidade (traço latente) do  $j$ -ésimo indivíduo.
- $P(U_{ij} = 1|\theta_j)$ : é a probabilidade de um indivíduo  $j$  com habilidade  $\theta_j$  responder corretamente o item  $i$  e é chamada de Função de Resposta do Item - FRI.
- $b_i$ : é o parâmetro de dificuldade (ou de posição) do item  $i$ , medido na mesma escala da habilidade.
- $a_i$ : é o parâmetro de discriminação (ou de inclinação) do item  $i$ , com valor proporcional à inclinação da Curva Característica do Item - CCI no ponto  $b_i$ .
- $c_i$ : é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item  $i$  (muitas vezes referido como a probabilidade de acerto casual).
- $D$ : é um fator de escala, constante e igual a 1. Utiliza-se o valor 1,7 quando deseja-se que a função logística forneça resultados semelhantes ao da função ogiva normal.

### Modelo logístico de 2 parâmetros

Para o modelo de dois parâmetros a probabilidade de acerto ao acaso é retirada, sendo consideradas então a dificuldade e a discriminação do item. O modelo logístico unidimensional de 2 parâmetros (ML2) é dado por:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-D_{ai}(\theta_j - b_i)}}, \quad (2.2.2)$$

com  $i=1,2,\dots,I$  e  $j=1,2,\dots,n$ .

### 2.2.2 Modelo de resposta gradual

O Modelo de Resposta Gradual, desenvolvido por Samejima (1969), assume que as categorias de resposta de um item podem ser ordenadas entre si. Além disso, tenta obter mais informações das respostas dos indivíduos do que simplesmente se eles deram respostas corretas ou não. Tal modelo é dado por



$$P_{i,k}^+(\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k})}},$$

com  $i=1,2,\dots,I$  e  $j=1,2,\dots,n$  e  $k=0,1,\dots,m_i$  onde:

- $b_{i,k}$ : é o parâmetro de dificuldade da  $k$ -ésima categoria do item  $i$ .

### 2.2.3 Curva característica do item

Pasquali (2009) diz que a relação entre o desempenho e os traços latentes pode ser descrita por uma equação matemática monotônica, chamada Curva Característica do Item - CCI. A CCI é única para cada item, tem um formato de "S" com inclinação e deslocamento na escala de habilidade definidos pelos parâmetros do item. Além disso, pressupõe que os indivíduos com maior habilidade possuem maior probabilidade de acertar o item, como é possível observar na Figura 1 a seguir.

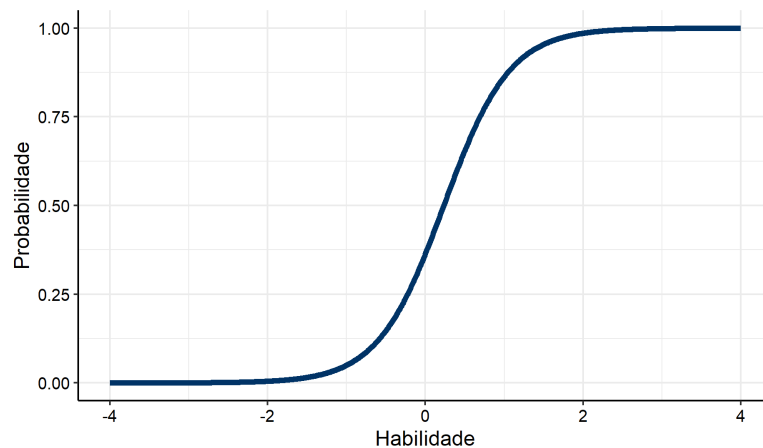


Figura 1: Exemplo de uma Curva Característica do Item - CCI

### 2.2.4 Estimação dos parâmetros dos itens

A estimação dos parâmetros dos itens pelo método de máxima verossimilhança, usando o ml3 para efeito de exemplificação, usa as seguintes suposições:

- as respostas oriundas de indivíduos diferentes são independentes;
- os itens respondidos de forma independente por cada indivíduo (independência local), fixada sua habilidade.

Assim, pode-se escrever a verossimilhança, como

$$L(\zeta) = \prod_{j=1}^n P(U_{j\cdot} = u_{j\cdot} | \theta_j, \zeta) \quad (2.2.3)$$

$$L(\zeta) = \prod_{j=1}^n \prod_{i=1}^I P(U_{ji} = u_{ji} | \theta_j, \zeta_i) \quad (2.2.4)$$

$$P(U_{ji} = u_{ji} | \theta_j, \zeta_i) = P(j_i = 1 | \theta_j, \zeta_i)^{u_{ji}} P(j_i = 0 | \theta_j, \zeta_i)^{1-u_{ji}} = P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} \quad (2.2.5)$$

- $\zeta$ : representa os parâmetros dos itens.

Dessa forma, a verossimilhança pode ser escrita como

$$L(\zeta) = \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} \quad (2.2.6)$$

Com o passo a passo demonstrado em Andrade, Tavares e Valle (2000), as equações de estimação dos parâmetros de discriminação ( $a_i$ ), dificuldade ( $b_i$ ) e de acerto ao acaso ( $c_i$ ) são,

$$a_i : D(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji})(\theta_j - b_i)W_{ji} = 0, \quad (2.2.7)$$

$$b_i : -Da_i(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji})W_{ji} = 0 \quad (2.2.8)$$

$$c_i : \sum_{j=1}^n (u_{jj} - P_{ji}) \frac{W_{ji}}{P_{ji}^*} = 0 \quad (2.2.9)$$

### 2.2.5 Estimação da habilidade

Assumindo que os parâmetros dos itens foram estimados, como mostrado anteriormente, e seguindo os mesmos pressupostos de independência entre as respostas diferentes e a independência local, pode se escrever a log-verossimilhança como função de  $\theta$ ,

$$\log L(\theta) = \sum_{j=1}^n \sum_{i=1}^I \{u_{ji} \log P_{ji} + (1 - u_{ji}) \log Q_{ji}\}. \quad (2.2.10)$$

O EMV de  $\theta_j$  é o valor que maximiza a verossimilhança, ou equivalentemente, é

a solução da equação

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = 0, \quad j = 1, \dots, n. \quad (2.2.11)$$

A partir da resolução dessa equação, como mostrado em Andrade, Tavares e Valle (2000), segue que a equação de estimação é

$$\theta_j : D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ji} = 0. \quad (2.2.12)$$

### 2.3 Coeficiente de correlação ponto-bisserial

O coeficiente de correlação ponto-bisserial é calculado individualmente para cada item. Calculado pelo coeficiente de correlação de Pearson entre a variável indicadora de respostas positivas para o item e os escores, sendo portanto, dado por

$$\begin{aligned} \rho_i^{PB} &= \frac{\sum_{j=1}^n U_{ij} T_j - (\sum_{j=1}^n U_{ij})(\sum_{j=1}^n T_j)/n}{\sqrt{[\sum_{j=1}^n (U_{ij} - \bar{U}_i)^2][\sum_{j=1}^n (T_j - \bar{T})^2]}} \\ &= \frac{n_i \bar{T}_A - n_i \bar{T}}{n \sqrt{p_i(1 - p_i)} S_T} = \left( \frac{\bar{T}_A - \bar{T}}{S_T} \right) \sqrt{\frac{p_i}{1 - p_i}}, \end{aligned} \quad (2.3.1)$$

em que

- $n_i$  é o número de respondentes com resposta positiva para o  $i$ -ésimo item;
- $\bar{T}_A$  é o escore médio dos respondentes com resposta positiva para o  $i$ -ésimo item;
- $\bar{T}$  é o escore médio dos respondentes;
- $S_T$  é o desvio-padrão dos escores;
- $p_i$  é a proporção de indivíduos que deram resposta positiva para o  $i$ -ésimo item.

Espera-se que cada item apresente um valor positivo para o coeficiente, caso contrário indicaria que o item é incoerente com o teste.

### 3 Resultados

#### 3.1 Análise descritiva

Neste capítulo será apresentada uma análise preliminar dos itens, antes de ajustar os modelos da TRI. Na Figura 2, estão distribuídas as respostas segundo item.

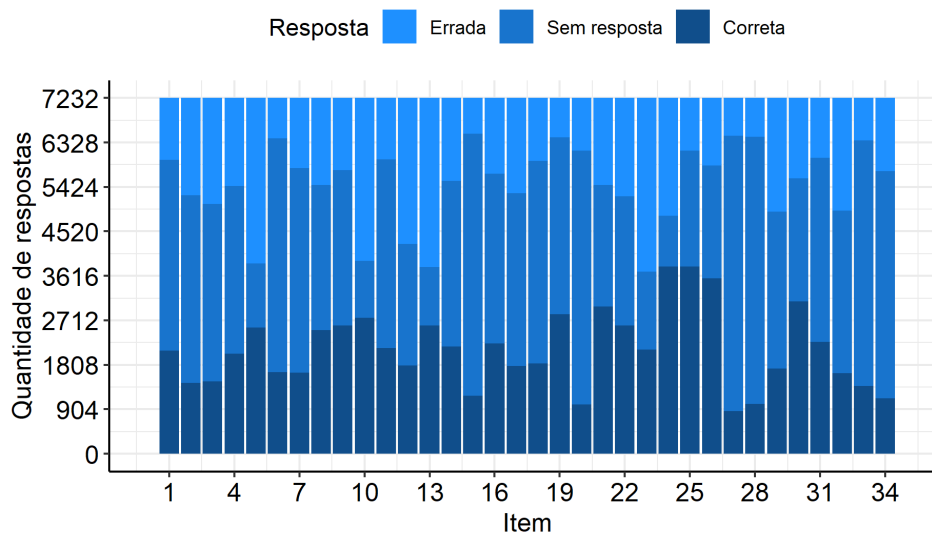


Figura 2: Distribuição das repostas segundo item

Nota-se que os itens 5, 10, 12, 13 e 23 apresentam os maiores números de respostas erradas. Por outro lado, os itens 24, 25, 26 e 30 possuem a maior quantidade de respostas corretas. A maioria dos itens possuem uma quantidade considerável de não resposta, com exceção do itens 5,10,13 e 22.

Os itens que possuem muitas repostas erradas ou que estão sem repostas, possuem potencial de serem considerados difíceis, assim como os itens com muitas repostas corretas, de serem fáceis. Tais fatos serão investigados nos próximos capítulos.

Analisando a proporção de respostas corretas na Figura 3 a seguir, a grande maioria dos itens possuem menos de 50% de acertos, com exceção apenas de dois deles.

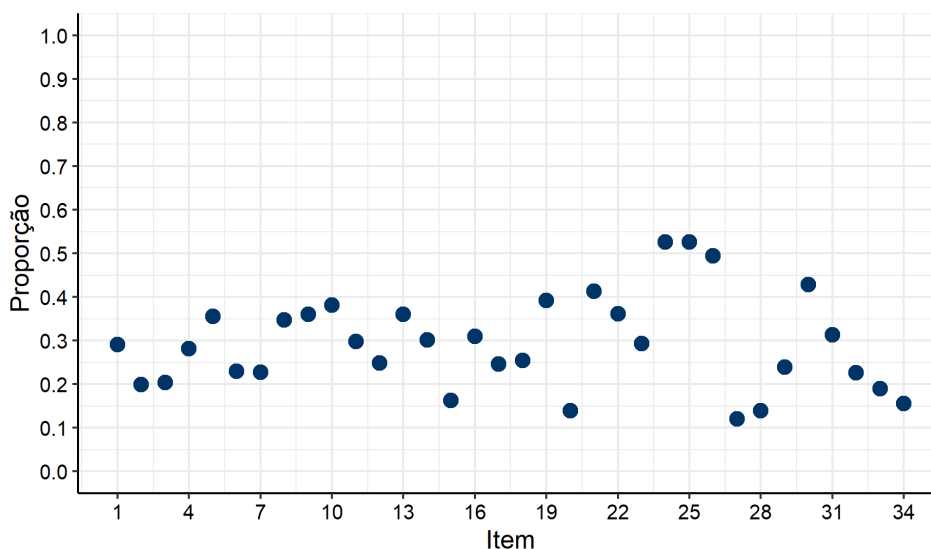


Figura 3: Proporção de acertos segundo item

Como apresentado na Seção 2.1, estão sendo estudados os 34 itens de física do vestibular. Dado isso, o maior número de respostas corretas que um candidato poderia ter é 34, mas nenhum conseguiu.

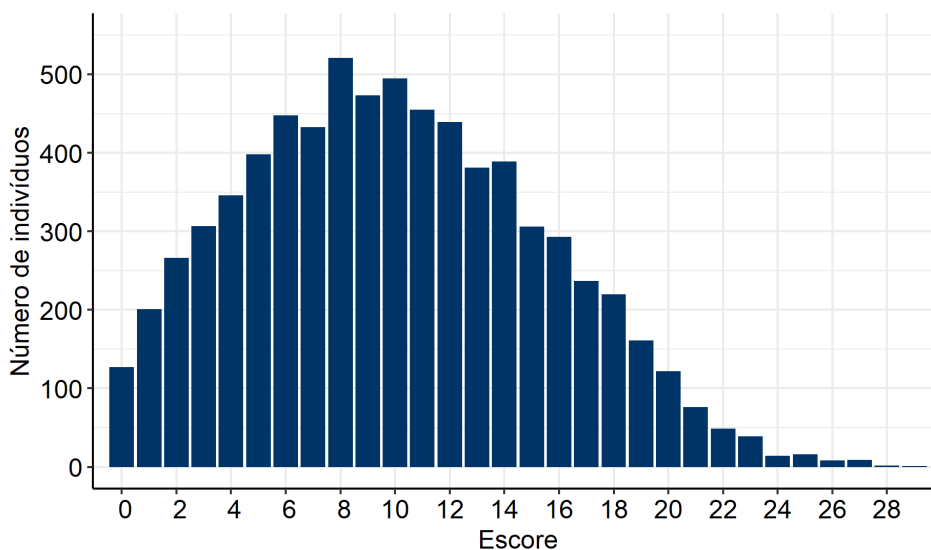


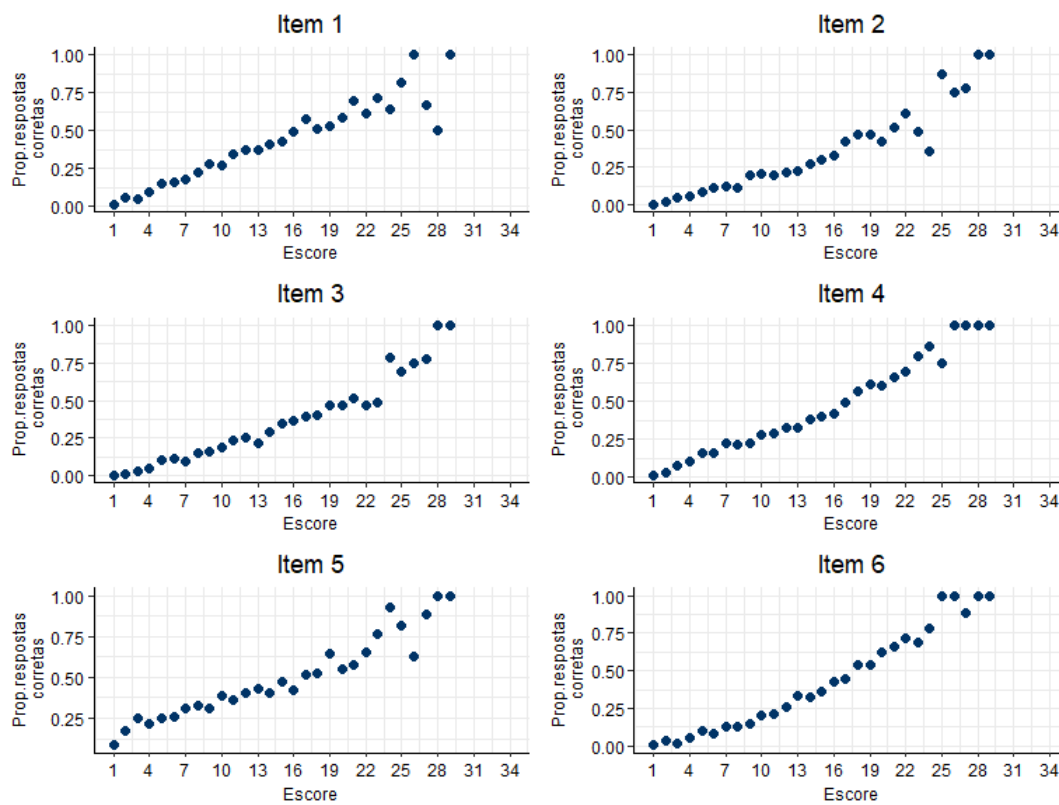
Figura 4: Distribuição da quantidade de respostas corretas

Observando a Figura 4, o maior número de acertos foi 29, sendo apenas um candidato com esse feito. A maioria acertou entre 8 e 10 itens, ressalta-se ainda que mais de 100 candidatos erraram todas as questões que responderam.

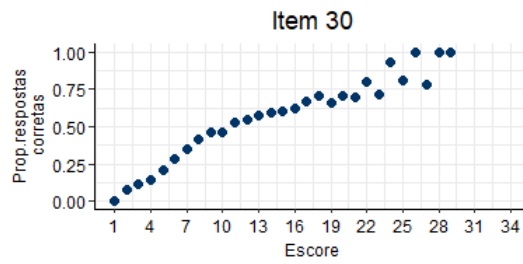
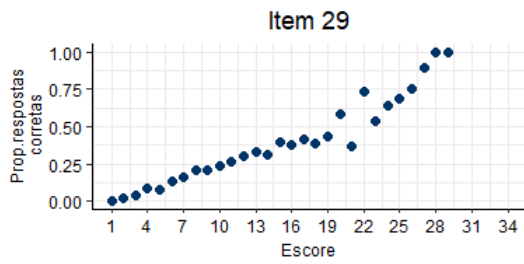
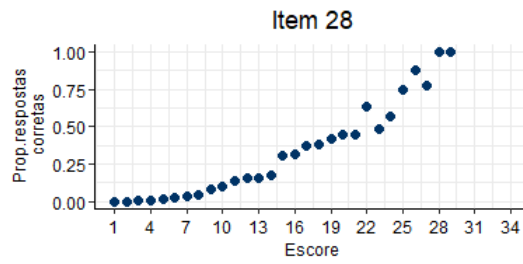
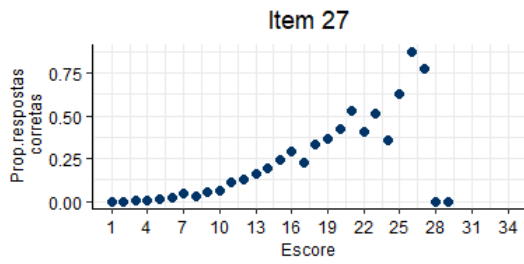
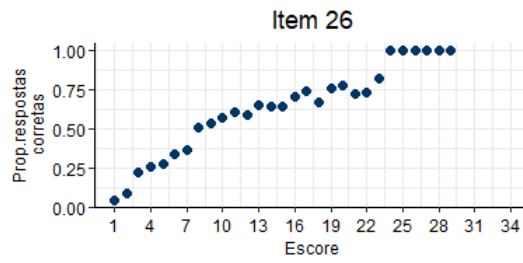
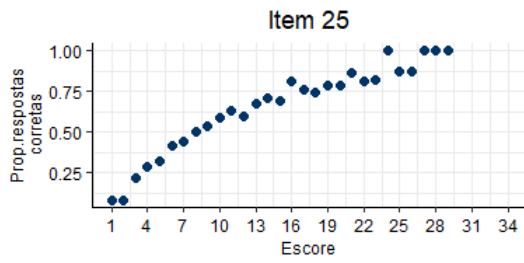
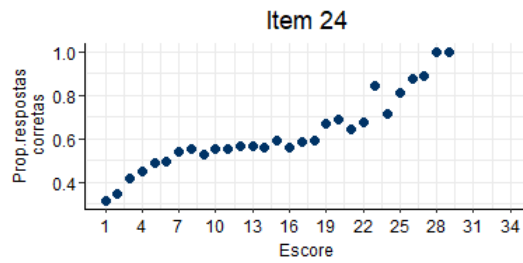
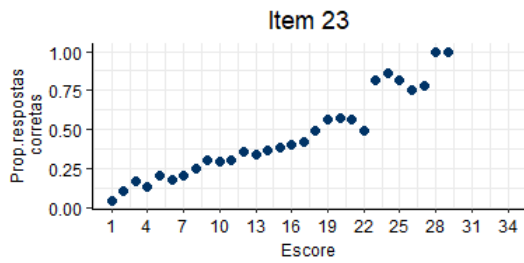
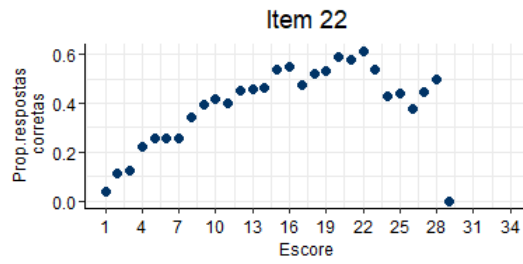
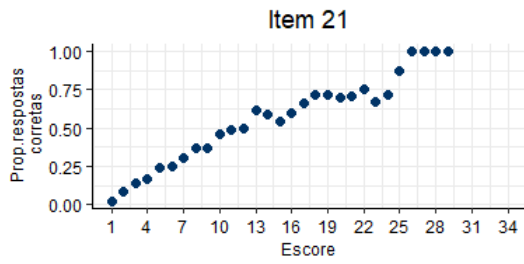
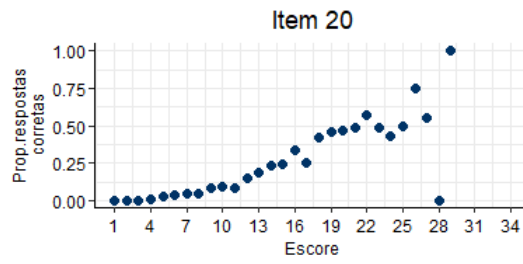
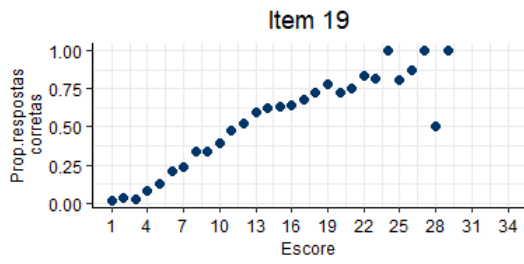
Na Figura 5, para cada item, os respondentes são agrupados pelo número de respostas corretas que cada obteve na prova (no eixo das abscissas) e, dentro de cada grupo, observamos a proporção de indivíduos que acertou o item (eixo das ordenadas). Isto

fornece uma visualização não paramétrica da CCI. É possível notar um comportamento irregular nos gráficos para indivíduos com escores maiores ou iguais a 24 em todos os itens.

Isto ocorre porque há poucos indivíduos que acertaram pelo menos 24 questões, como pode-se ver na Figura 4. Ainda assim, é possível observar uma baixa variação na proporção de respostas corretas para os itens 12 e 22 para respondentes com número de respostas corretas entre 12 e 20, indicando um possível baixo poder de discriminação entre respondentes com diferentes níveis de conhecimento, fato confirmado também para os itens 5, 13, 23 e 24 nas estimativas do valor do parâmetro de discriminação no modelo logístico de 3 parâmetros observados na Tabela 1.









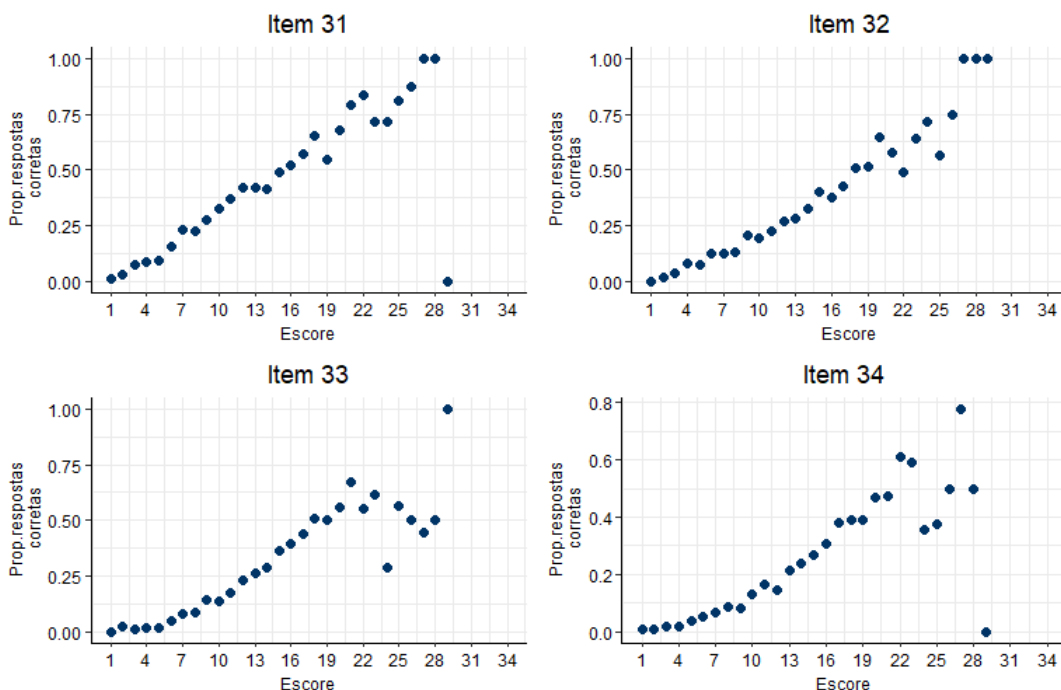


Figura 5: CCI não paramétrica

Para finalizar este capítulo de análise preliminar dos itens, observa-se na Figura 6 o resultado do cálculo do coeficiente de correlação ponto-bisserial.

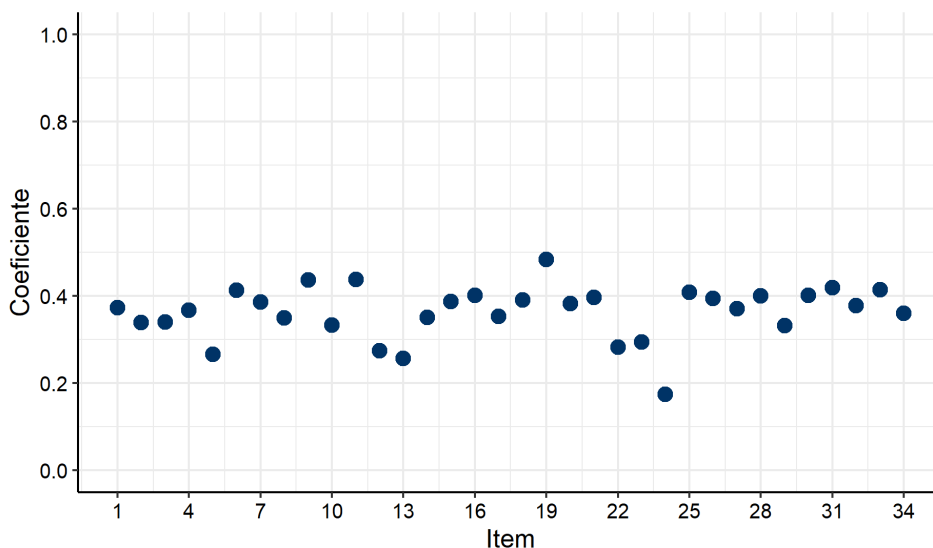


Figura 6: Coeficiente de correlação ponto-bisserial de cada item

Analisando os valores do coeficiente, nota-se que nenhum item obteve valor negativo, ou seja, não são incoerentes com o teste. Mesmo obtendo um resultado acima de zero, o item 24 chama atenção por apresentar o menor valor para o coeficiente, além de estar mais afastado dos outros resultados.

## 3.2 Ajuste dos modelos logísticos

### 3.2.1 Ajuste do modelo logístico de 3 parâmetros

O primeiro modelo ajustado foi o logístico de 3 parâmetros, os resultados estão dispostos na Tabela 1 a seguir.

Item	Acerto ao Acaso	Dificuldade	Discriminação
1	0,00	1,11	0,97
2	0,00	1,73	0,94
3	0,00	1,81	0,86
4	0,00	1,35	0,79
5	0,00	1,67	0,37
6	0,00	1,38	1,07
7	0,00	1,34	1,15
8	0,00	1,03	0,68
9	0,00	0,74	0,95
10	0,00	0,91	0,58
11	0,00	1,05	0,99
12	0,00	2,26	0,52
13	0,00	1,42	0,42
14	0,00	1,21	0,79
15	0,00	1,48	1,52
16	0,00	1,02	0,95
17	0,00	1,43	0,92
18	0,00	1,30	1,00
19	0,00	0,49	1,24
20	0,00	1,62	1,54
21	0,00	0,49	0,86
22	0,00	1,13	0,54
23	0,00	1,69	0,56
24	0,00	-0,61	0,18
25	0,00	-0,12	0,88
26	0,00	0,05	0,80
27	0,00	1,65	1,72
28	0,00	1,52	1,73
29	0,00	1,63	0,81
30	0,00	0,42	0,82
31	0,00	0,93	1,05
32	0,00	1,46	1,02
33	0,00	1,32	1,54
34	0,00	1,73	1,23

Tabela 1: Resultado do modelo logístico de 3 parâmetros

Analisando os resultados, o primeiro que chama atenção é o fato de todos os itens estarem com probabilidade de acerto ao acaso zerada. O parâmetro de dificuldade caracteriza o item 12 como muito difícil. Por outro lado, os itens 24 e 25, de acordo com o parâmetro, são os mais fáceis. Além disso, nota-se pela Figura 7 abaixo, que a relação entre a proporção de resposta e o parâmetro de dificuldade estão coerentes.

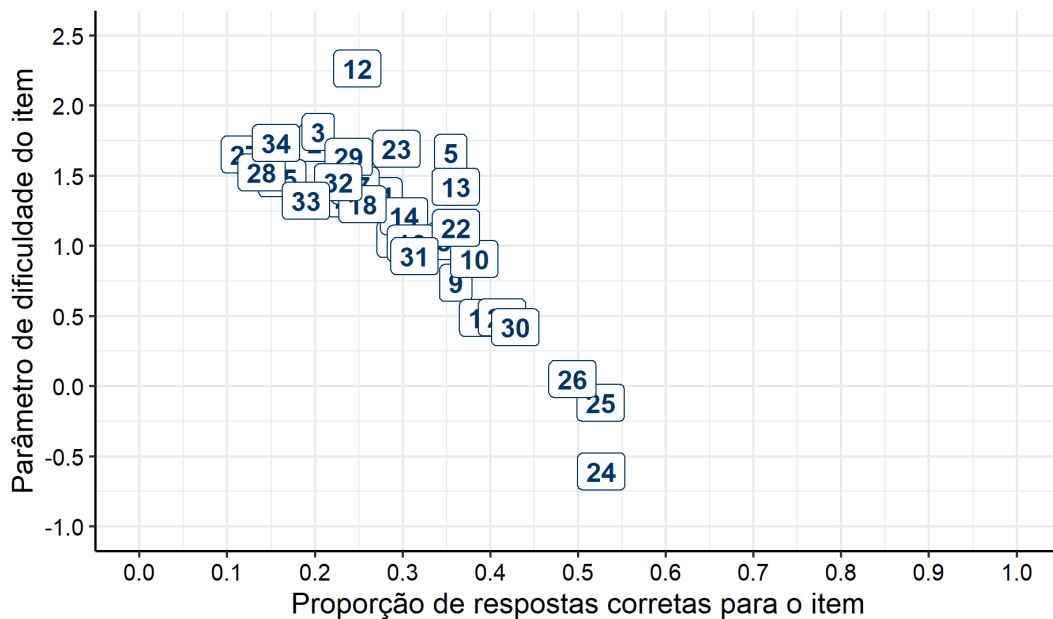
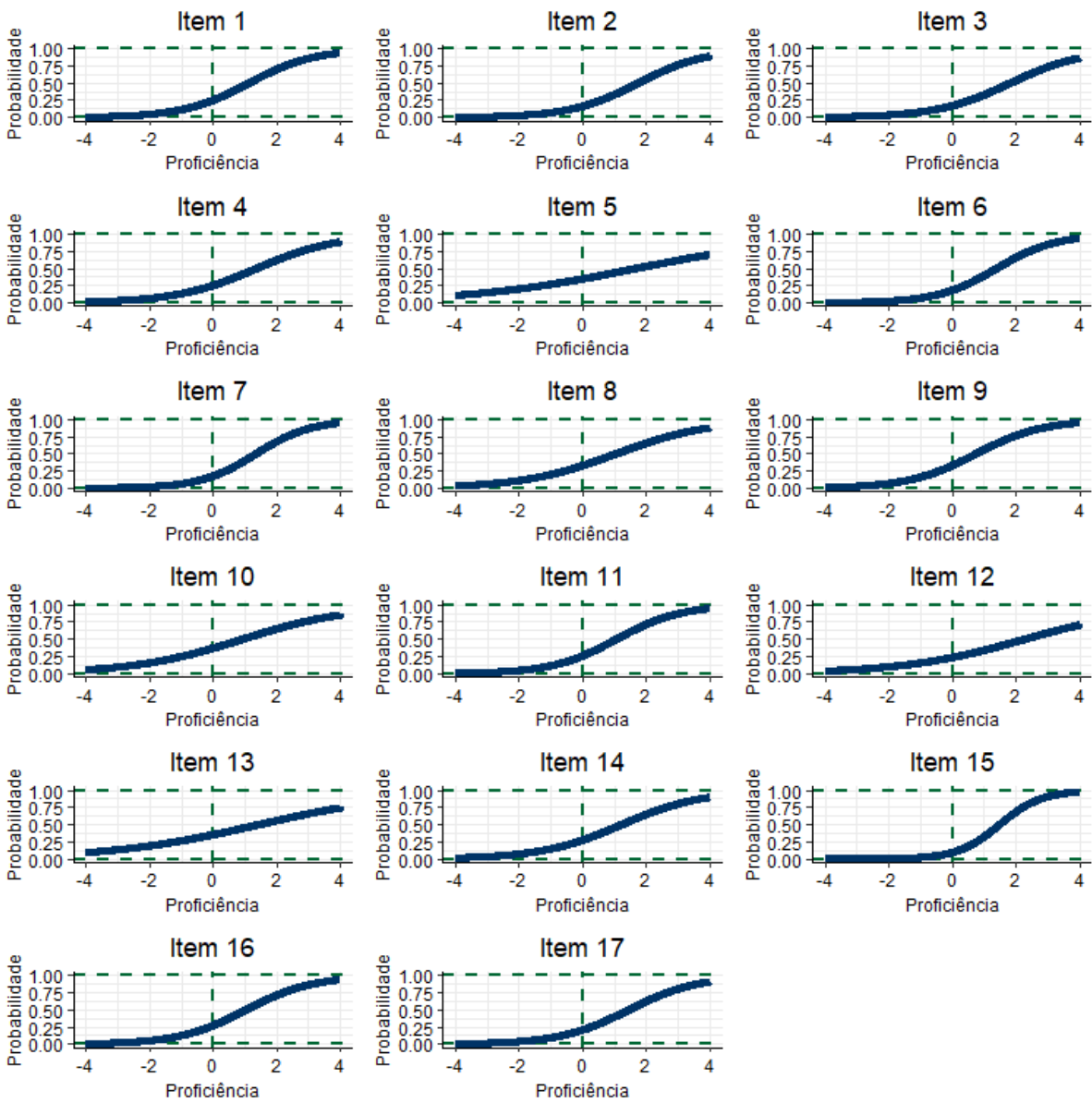


Figura 7: Proporção de respostas corretas e o parâmetro de dificuldade do item

Olhando para os valores de discriminação, os itens 5, 10, 12, 13, 22, 23 e 24 apresentaram pouco poder de discriminação, ou seja, candidatos com habilidades diferentes tem aproximadamente a mesma probabilidade de responder corretamente o item. Por outro lado, os itens 15, 20, 27, 28 e 33 apresentaram os maiores valores para o parâmetro de discriminação.

Observando as curvas características de cada item na Figura 8, tem-se para o grupo de itens com baixo poder de discriminação uma curva menos inclinada. Já para os itens com valores maiores de discriminação, as curvas são bem íngremes em alguns momentos.



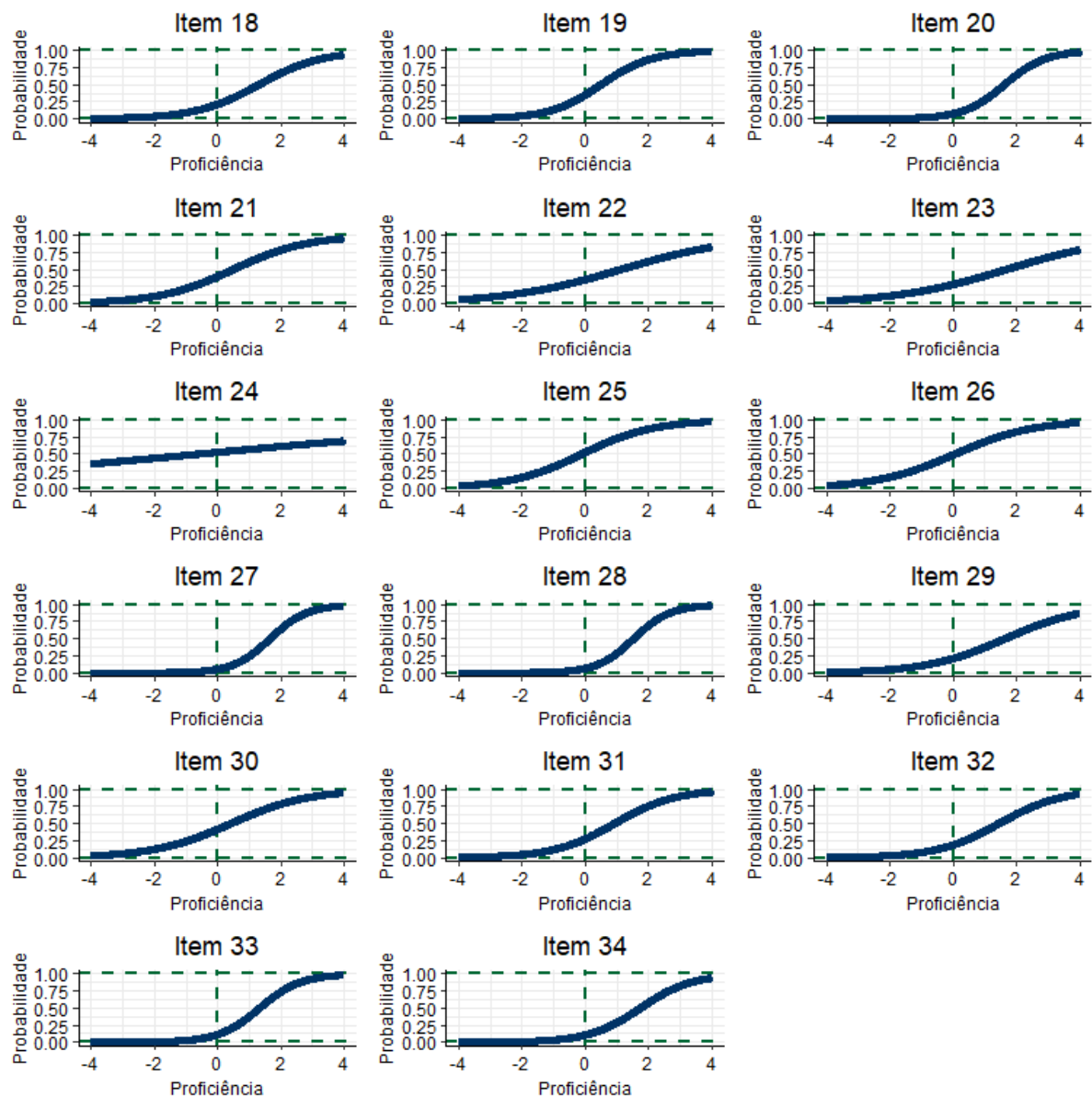


Figura 8: Curvas características dos itens

Comparando o escore padronizado calculado utilizando técnicas da TRI, os escores padronizados calculados utilizando a fórmula do vestibular e a proficiência estimada, nota-se pela Figura 9 que a proficiência está mais coerente com os escores calculados pela TRI em comparação com os estimados pelo vestibular, o que faz sentido, pois no ajuste do modelo logístico, consideramos "não resposta" e "resposta errada" como uma categoria única, que equivale a não responder.

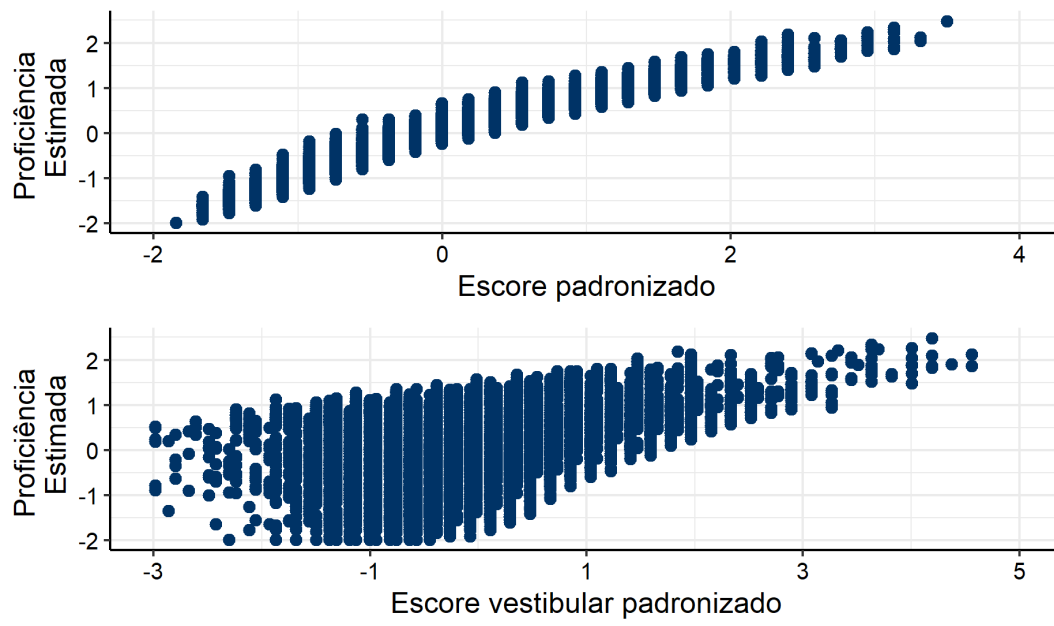


Figura 9: Proficiência Estimada e os escores padronizados

### 3.2.2 Ajuste do modelo logístico de 2 parâmetros

O segundo modelo ajustado foi o logístico de 2 parâmetros, os resultados podem ser conferidos na Tabela 2 abaixo.

Item	Dificuldade	Discriminação
1	1,11	0,97
2	1,73	0,94
3	1,81	0,86
4	1,35	0,79
5	1,66	0,37
6	1,38	1,07
7	1,34	1,15
8	1,03	0,68
9	0,74	0,95
10	0,91	0,58
11	1,05	0,99
12	2,26	0,52
13	1,42	0,42
14	1,21	0,79
15	1,48	1,52
16	1,02	0,95
17	1,43	0,92
18	1,30	1,00
19	0,49	1,24
20	1,62	1,54
21	0,49	0,86
22	1,13	0,54
23	1,69	0,56
24	-0,61	0,18
25	-0,12	0,88
26	0,05	0,80
27	1,65	1,72
28	1,52	1,73
29	1,63	0,81
30	0,42	0,82
31	0,93	1,05
32	1,46	1,02
33	1,32	1,54
34	1,73	1,23

Tabela 2: Resultados do modelo logístico de 2 parâmetros

Analisando o resultado do ajuste do modelo logístico de 2 parâmetros, nota-se que são os mesmos valores do ML3 apresentados na Seção 3.2.1. Tal fato já era esperado, uma vez que no ML3 o parâmetro de acerto ao acaso foi muito próximo de zero para todos

os itens. Dessa forma as análises dos parâmetros são as mesmas para os dois modelos, como é possível observar na Figura 10 abaixo.

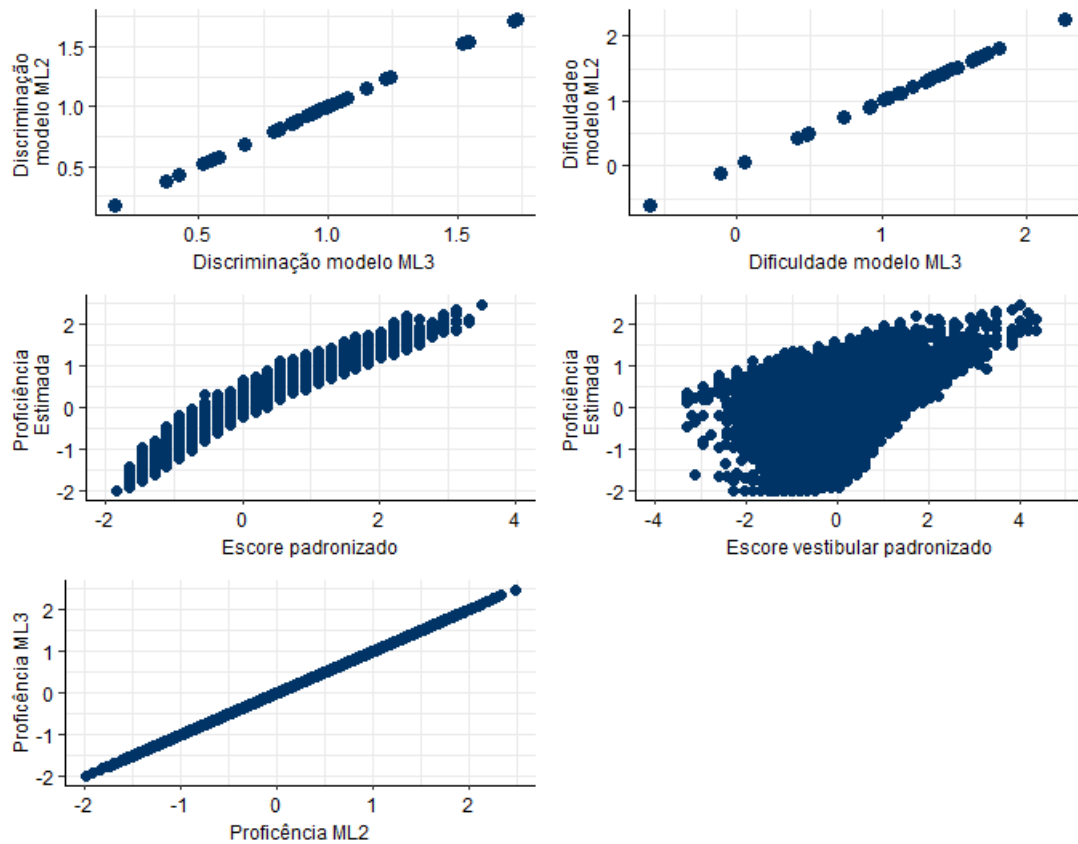


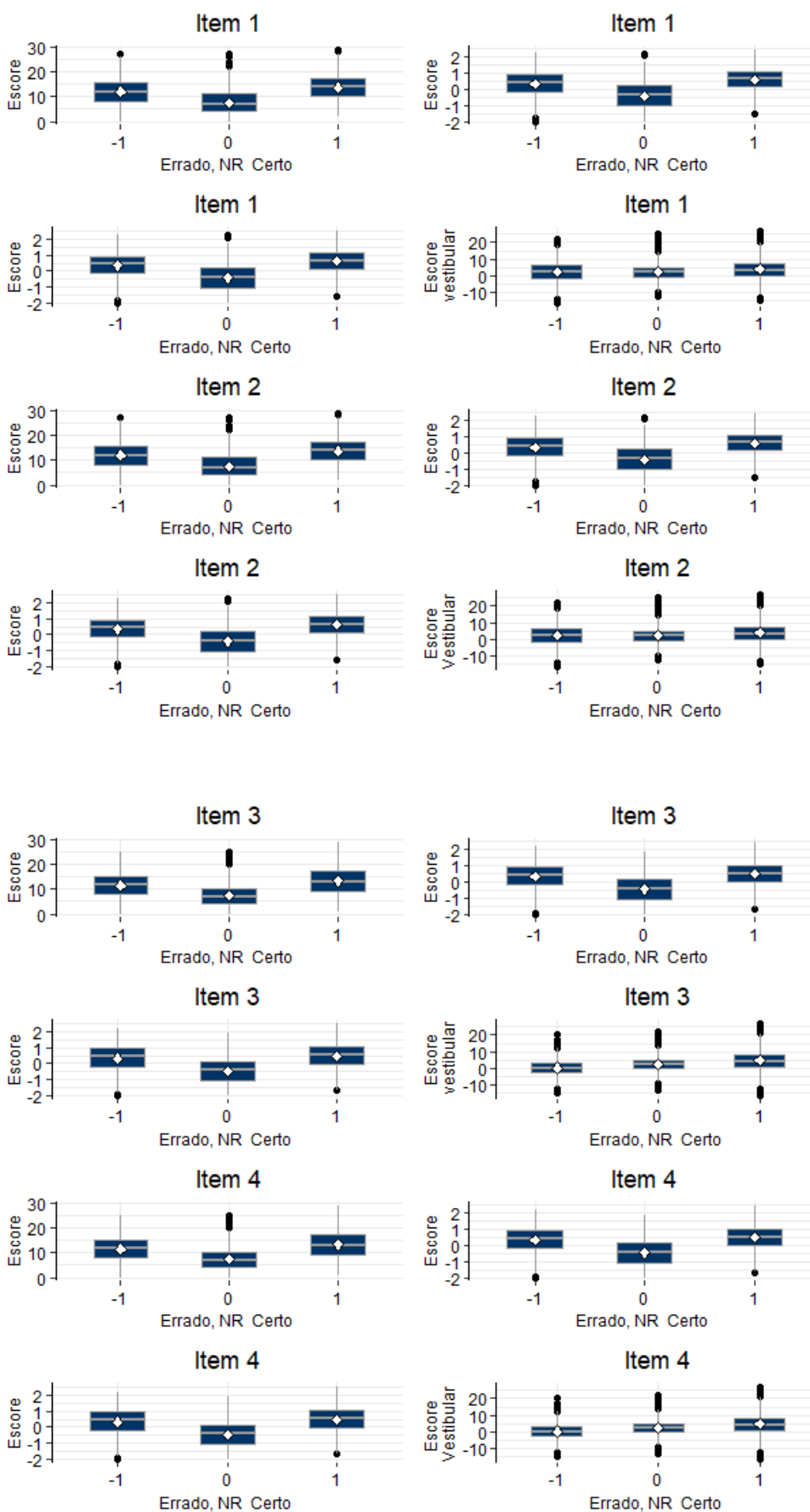
Figura 10: Comparação dos parâmetros dos modelos logísticos de 2 e 3 parâmetros

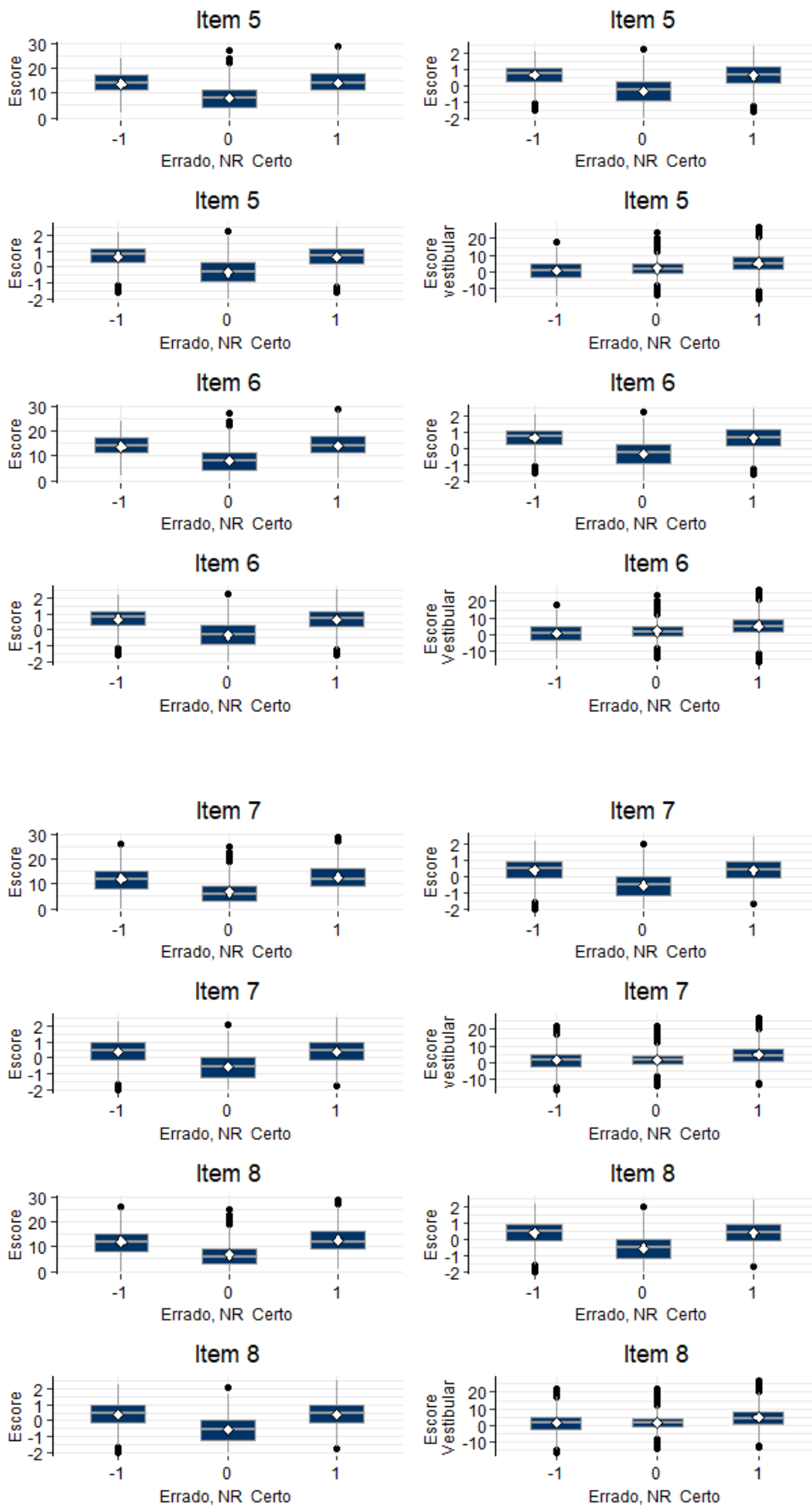
Comparando os escores padronizados, a proficiência estimada via ML3, a proficiência estimada via ML2 e o escore calculado pelo vestibular, é possível identificar um comportamento em comum para todos os itens.

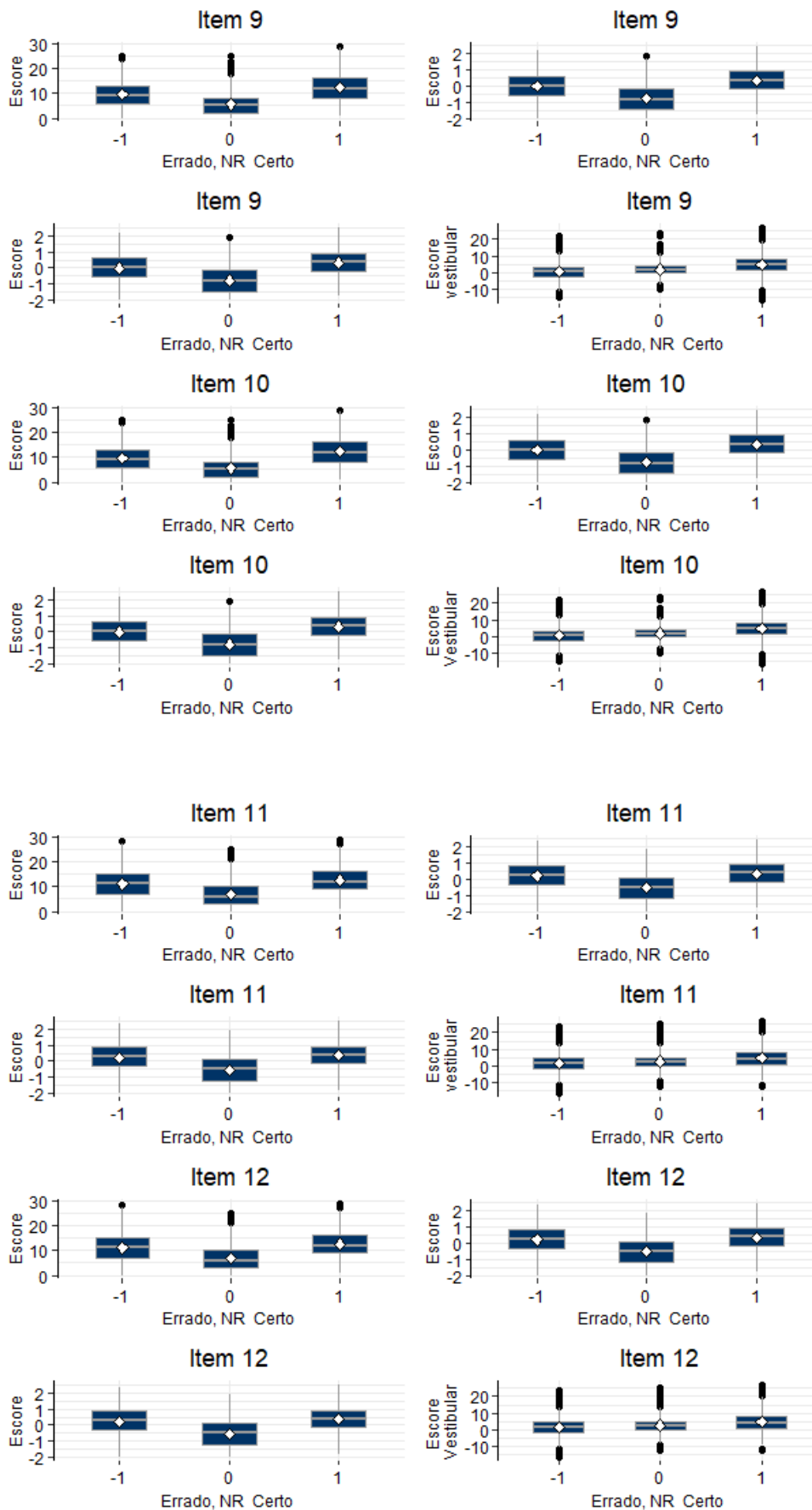
Analisando a Figura 11, nota-se que para o cálculo do escore feito pela TRI, o candidato que errou um determinado item, no geral tem um escore mais alto que o candidato que não respondeu o item. Já no escore calculado pelo vestibular, o valor no geral é maior para quem não respondeu o item.

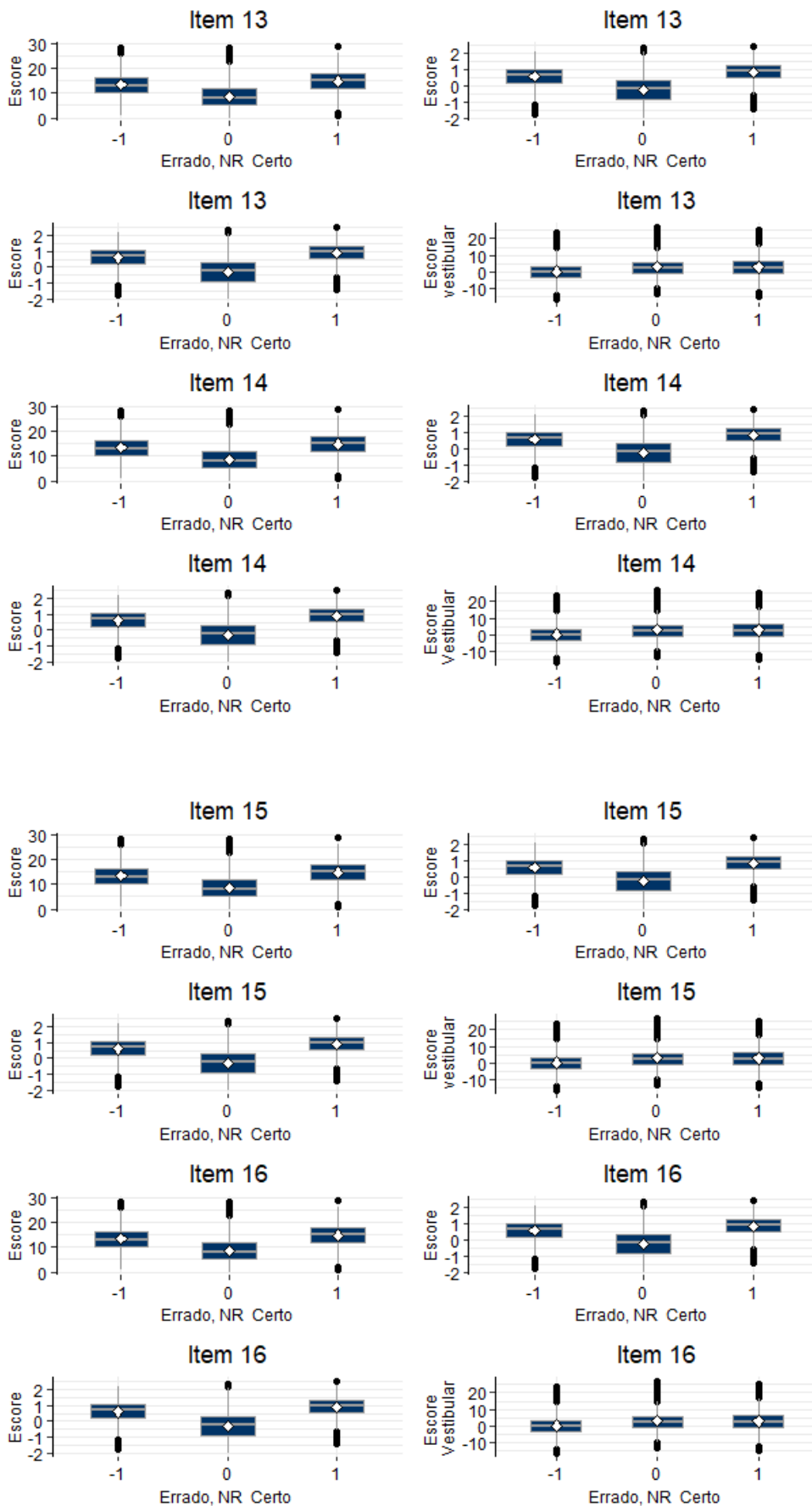
Utilizando o item 1 como exemplo, os candidatos que erraram o item 1, no geral possuem um escore final maior em comparação com quem não respondeu tal item (primeiros 3 boxplots). Já o escore no vestibular será em geral mais alto pra quem não respondeu o item 1 (último boxplot).

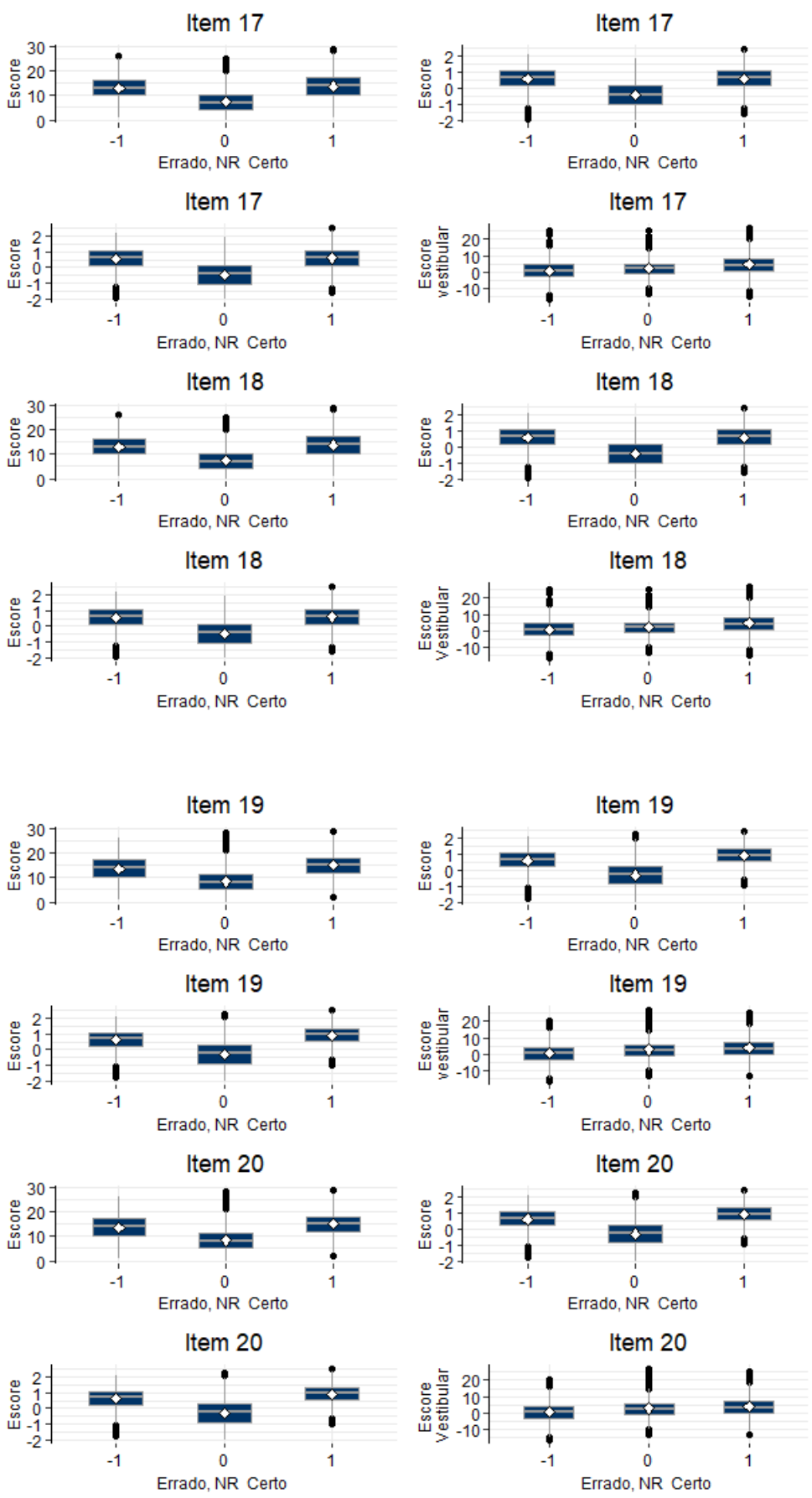


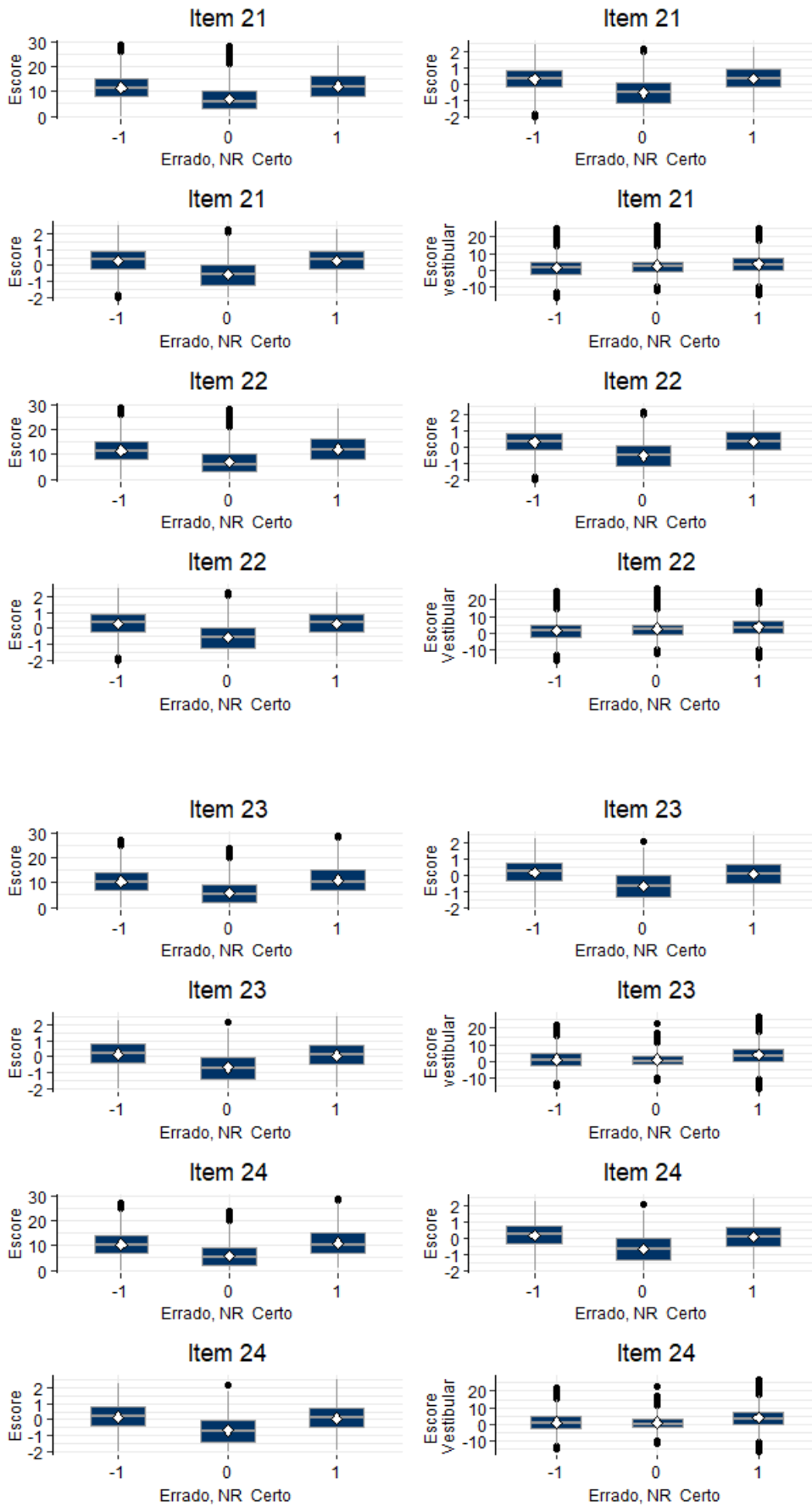


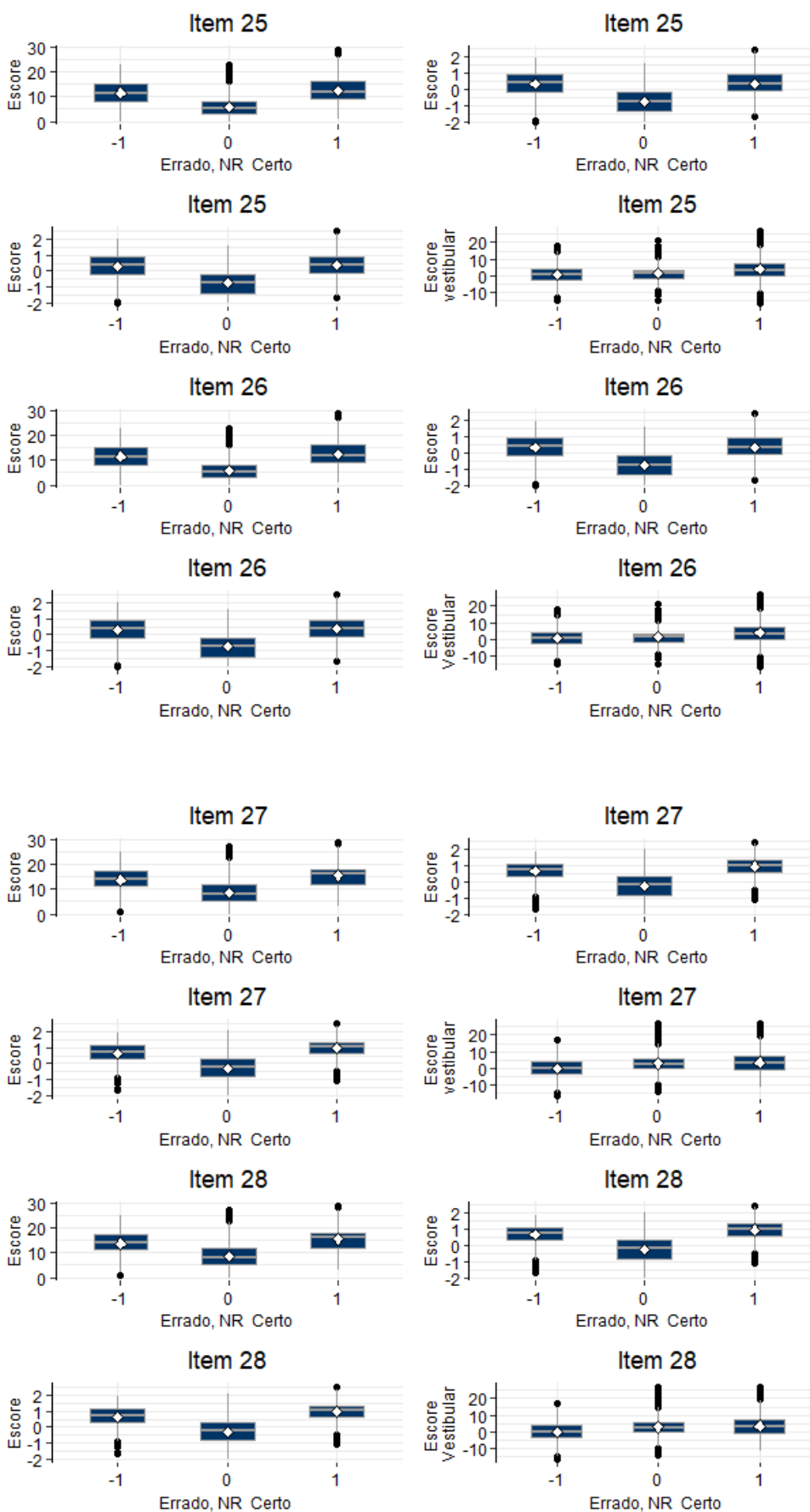


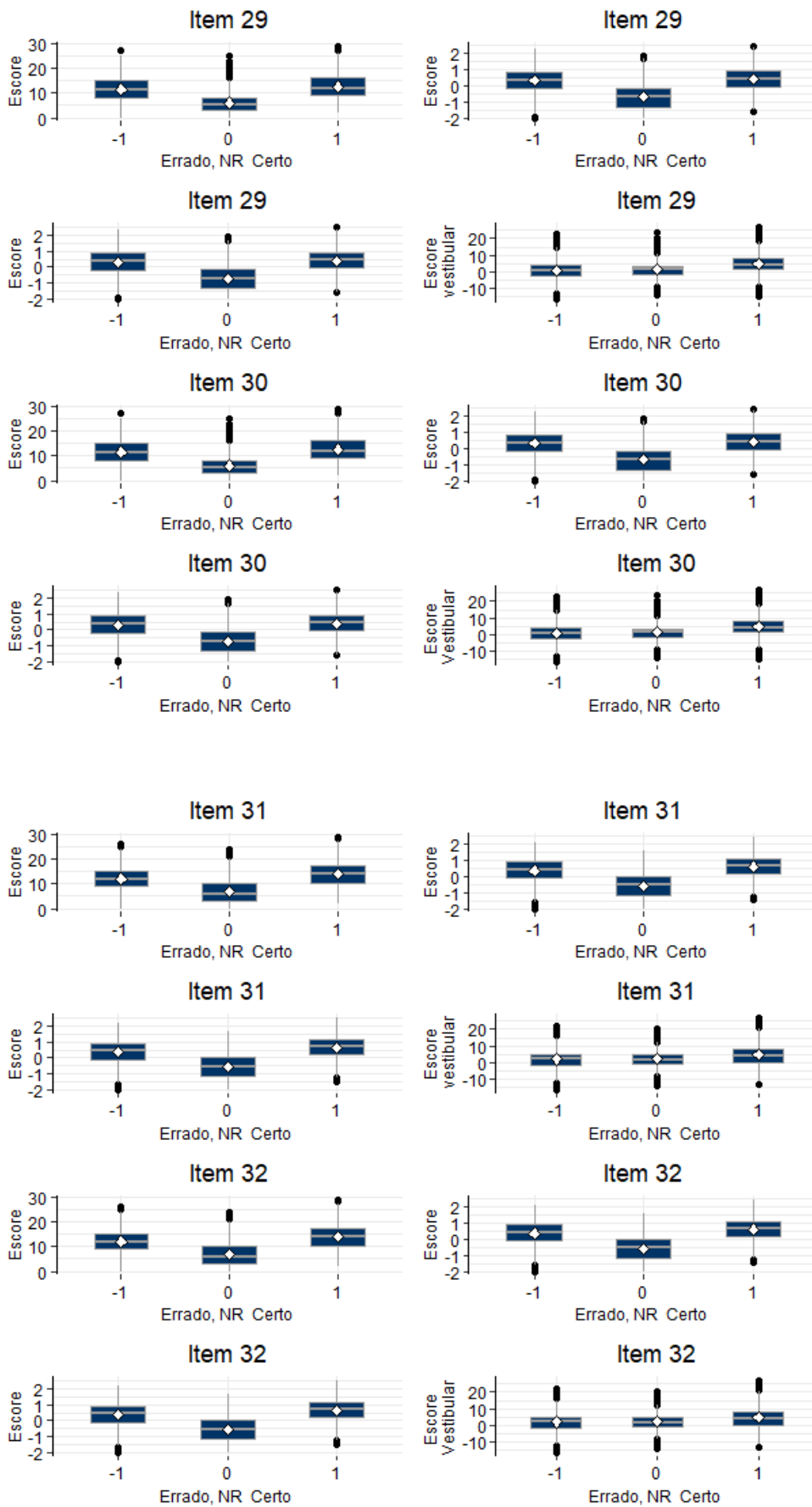














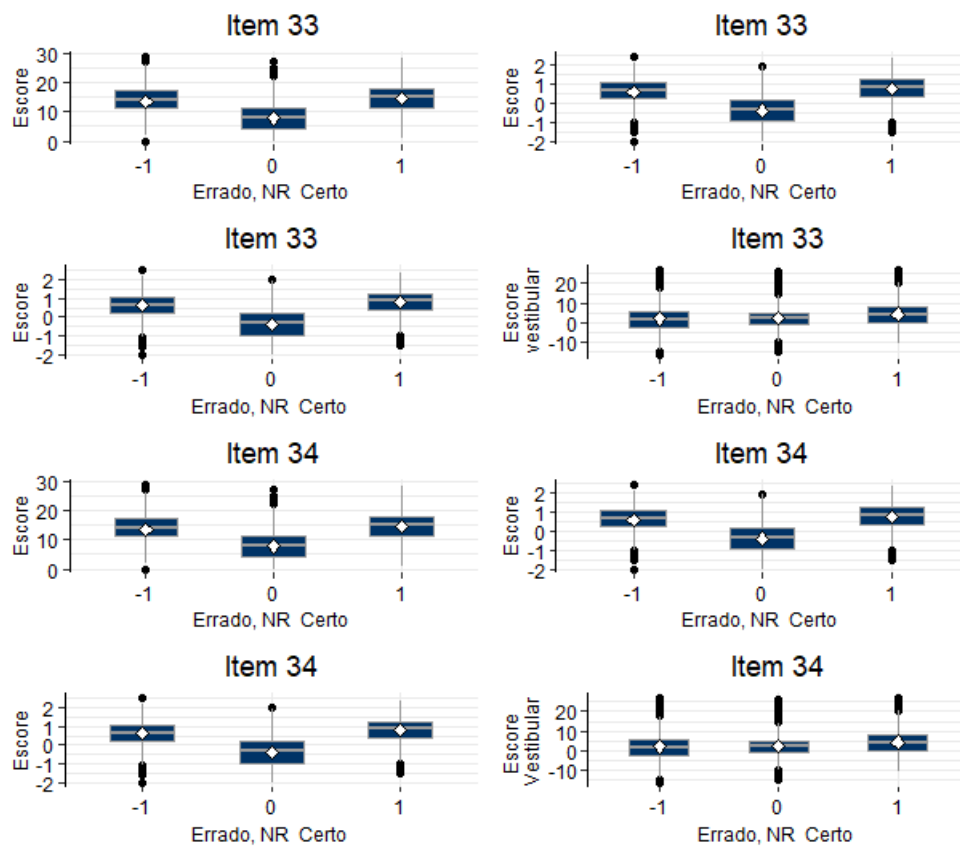


Figura 11: Comparação entre os escores calculados via TRI e os calculados pelo vestibular

### 3.3 Modelo de resposta gradual

Para ajustar o modelo de resposta gradual, considerou-se a seguinte ordenação para as alternativas dos itens: resposta errada, não resposta e resposta certa.

Analisando os resultados dos parâmetros apresentados na Figura 12, nota-se alguns problemas no modelo ajustado dessa forma. O parâmetro de discriminação apresenta alguns valores negativos, os parâmetros de dificuldade apresentam valores estimados muito altos, além de que alguns valores para o primeiro parâmetro de dificuldade ( $b_1$ ) serem maiores que o segundo ( $b_2$ ).

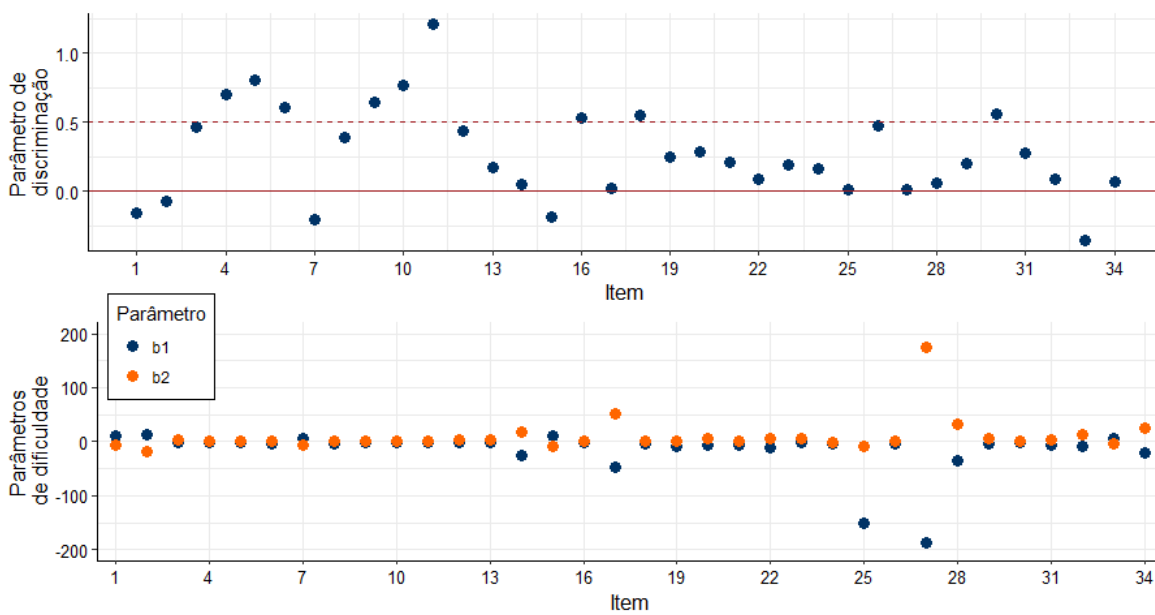


Figura 12: Resultado do MRG com a ordenação resposta errada, não resposta e resposta certa

Comparando as proficiências estimadas via modelo de resposta gradual e via modelos logísticos (Figura 13), os valores não possuem relação clara.

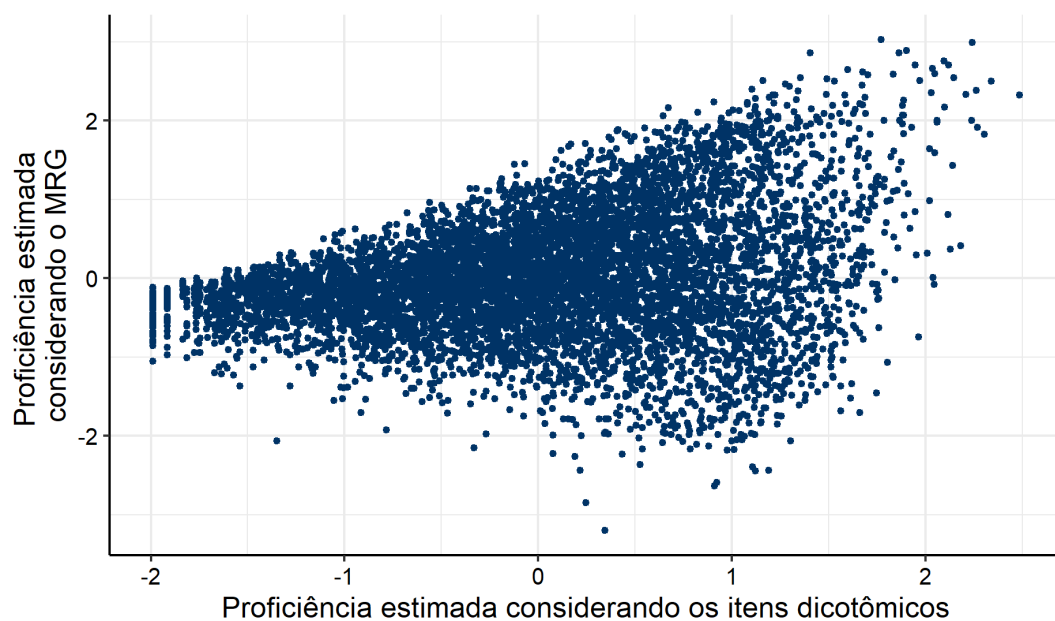


Figura 13: Comparação da proficiência estimada via modelos logísticos e via MRG com ordenação resposta errada, não resposta e resposta certa

Diante desses fatores que indicaram o MRG com a ordenação resposta errada, não resposta e resposta certa, como um modelo mal ajustado, uma nova ordem será considerada. Para o segundo ajuste, a ordenação adotada foi não resposta, resposta errada e resposta certa.

Comparando as proficiências estimadas pelo MRG com nova ordenação e a proficiência estimada pelo modelo logístico (Figura 14), os valores são mais coerentes.

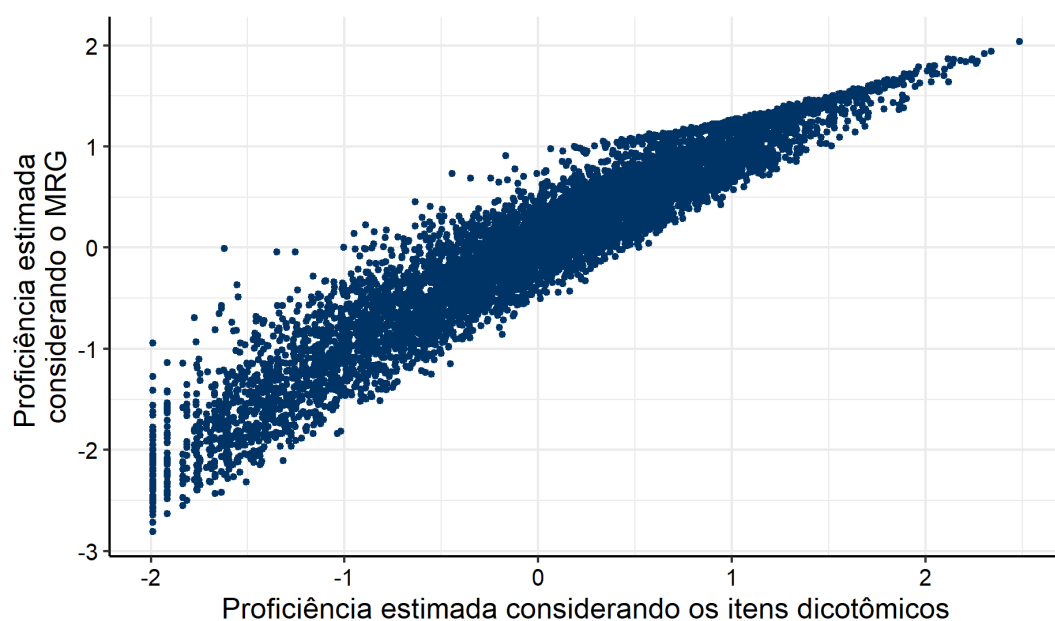


Figura 14: Comparação da proficiência estimada via modelos logísticos e via MRG com ordenação não resposta, resposta errada e resposta certa

Observando os resultados dos parâmetros do modelo com essa ordenação, dispostos na Figura 15, os valores dos parâmetros de discriminação são positivos para todos os itens. Assim como no modelo logístico, o item 24 apresenta o menor valor de discriminação, abaixo de 0,5.

Para os parâmetros de dificuldade, os valores são coerentes e todos os itens apresentam o segundo parâmetro de dificuldade ( $b_2$ ) maior que o primeiro ( $b_1$ ). Os itens 5, 10, 12, 13, 23 e 24, apresentam as maiores diferenças entre os parâmetros de dificuldade, tais itens apresentam os menores valores de não resposta (Figura 2).

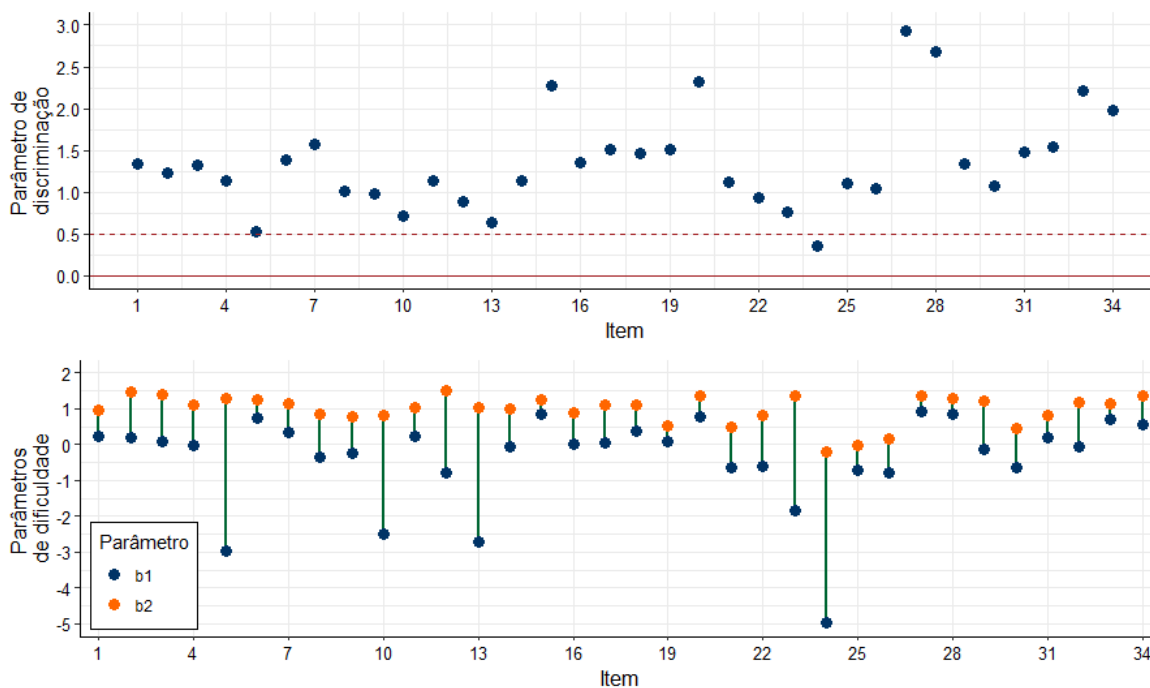


Figura 15: Resultado do MRG com a ordenação não resposta, resposta errada e resposta certa

As Figuras 16 e 17 mostram que os itens 5, 13 e 24 apresentam baixo poder de discriminação, indicando que o uso desses itens em futuras provas do vestibular deve ser desconsiderado.

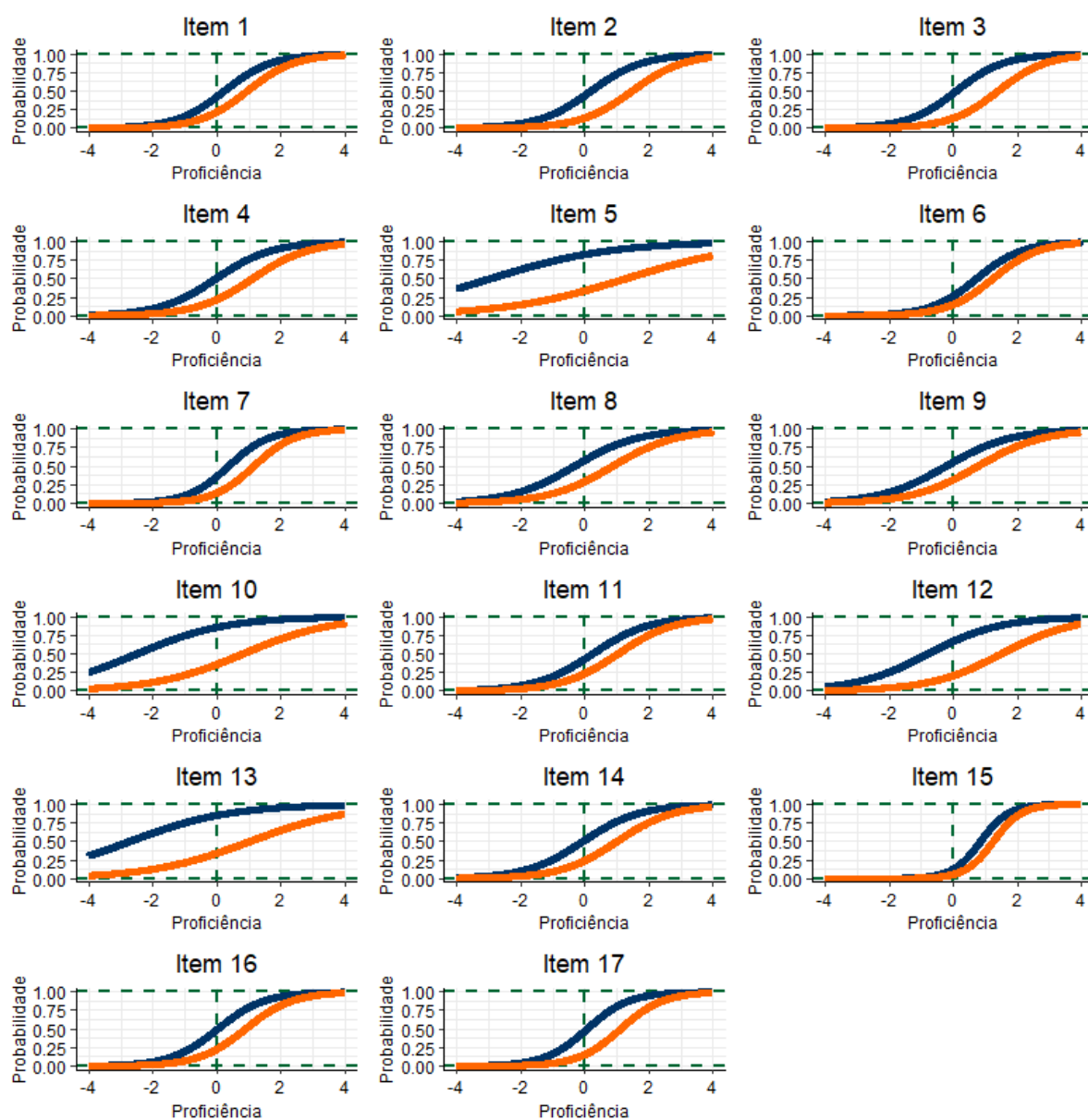


Figura 16: CCI do MRG para os itens 1 a 17

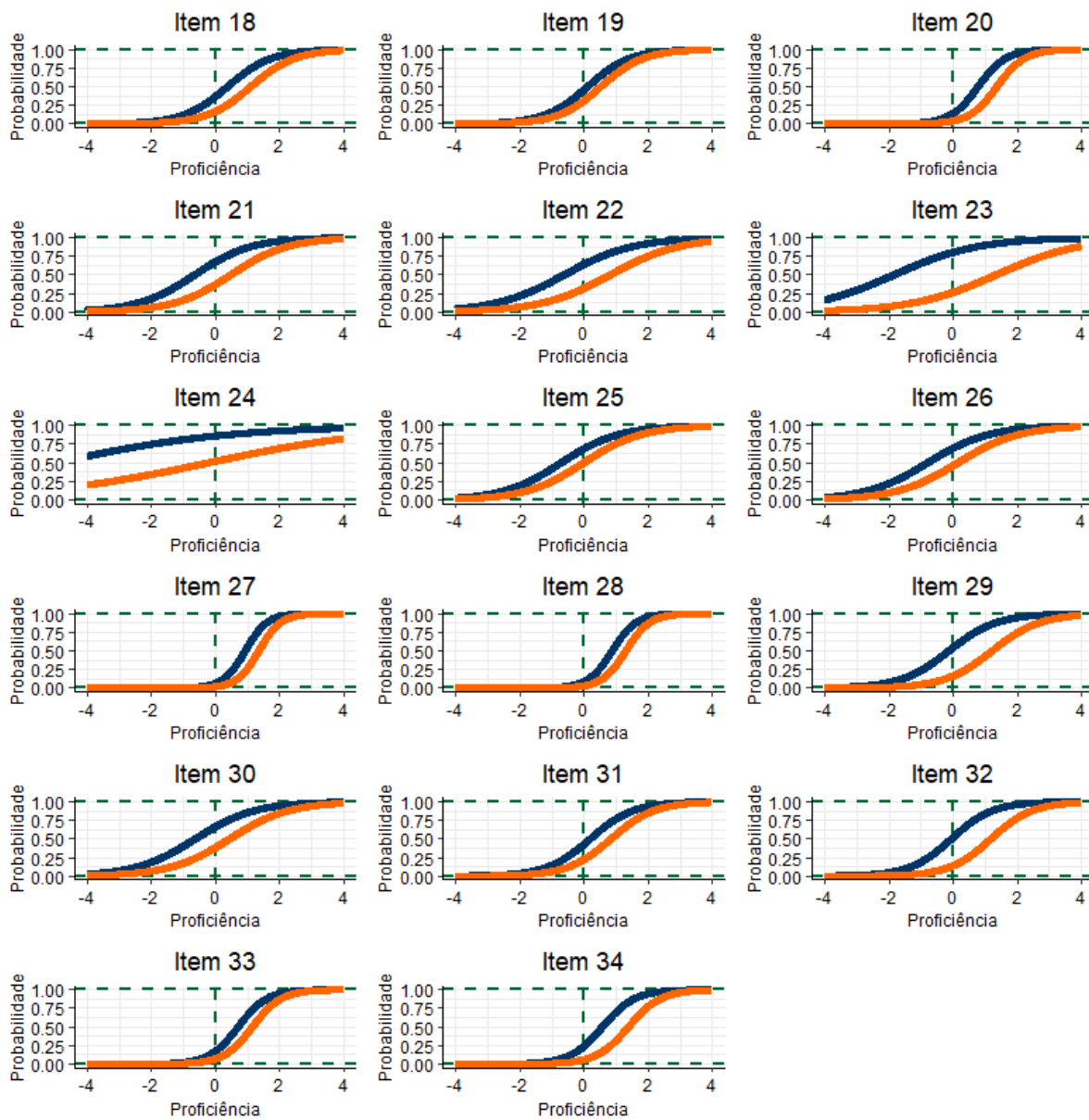


Figura 17: CCI do MRG para os itens 18 a 34

## 4 Conclusão

Os resultados indicam, que os indivíduos com maior conhecimento tendem a responder as questões, correndo o risco de eventualmente darem uma resposta incorreta para um item, enquanto que indivíduos com menor conhecimento tendem a não responder os itens, devido a penalização adotada no cálculo da nota do vestibular.

Tomando o parâmetro de discriminação como fator decisivo para a presença ou não dos itens em futuras provas do vestibular, sugere-se que os itens 5, 13 e 24, não sejam utilizados, pois apresentam baixo poder de discriminação. Tal fato é um problema, pois itens com baixo poder de discriminação não diferenciam candidatos com níveis de habilidades diferentes.

## Referências

- Andrade, D.; Tavares, H.; Valle, R. *Teoria de Resposta ao Item: Conceitos e Aplicações*. [S.l.]: São Paulo, ABE, 2000.
- Anjos, A.; Andrade, D. *Teoria de Resposta ao Item com o uso do R*. [S.l.]: João Pessoa, PB, 2012.
- Baker, F. B.; Kim, S.-H. *The Basics of Item Response Theory Using R*. [S.l.]: Springer International Publishing, 2017.
- Pasquali, L. *TRI - Teoria de Resposta ao Item: teoria, procedimentos e aplicações*. [S.l.]: Brasília: LabPAM/UnB, 2009.
- Samejima, F. *Estimation of latent ability using a response pattern of graded scores*. [S.l.]: Richmond, VA: Psychometric Society, 1969.