

Teoria da Resposta ao Item para analisar atitudes e concepções de acadêmicos na Disciplina de Estatística

Item Response Theory to analyze academic attitudes and conceptions in the Statistics Discipline

Rosebel Trindade Cunha Prates^I, Fernando de Jesus Moreira Junior^{II},
Angela Isabel dos Santos Dullius^{III}, Angela Pelegrin Ansuji^{IV},
Rafaella Trindade Cunha Prates^V, Maria Emília Camargo^{VI}

RESUMO

A pesquisa teve como objetivo analisar as atitudes e a concepção de alunos dos cursos de graduação da Universidade Federal de Santa Maria em relação à disciplina de Estatística. Foram entrevistados 256 acadêmicos de diversos cursos de graduação que cursaram a disciplina de Estatística. O questionário aplicado adotou a escala de Likert, composto de 20 afirmações a respeito da Estatística, onde 10 eram afirmações positivas e 10 afirmações negativas, e para análise dos dados utilizou o Modelo Logístico de Dois Parâmetros da Teoria da Resposta ao Item. Os resultados apresentaram que todos os 20 itens possuem uma boa qualidade e conseguem estimar adequadamente o nível de sentimento dos alunos que possuem sentimento mediano em relação à Estatística, entre mais ou menos 1,5 desvio padrão em torno da média. No entanto, há carência de itens para medir adequadamente aqueles alunos que possuem sentimentos extremos em relação à Estatística, ou seja, aqueles que estão posicionados 1,5 desvio padrão acima ou abaixo da média.

Palavras-chave: Escala Likert; Modelo Logístico de Dois Parâmetros; Concepções; Atitudes; Sentimentos.

ABSTRACT

The research aimed to analyze the attitudes and the conception of undergraduate students of the Federal University of Santa Maria in relation to the discipline of Statistics. Were interviewed 256 academics from various undergraduate courses who attended the discipline of Statistics. The applied questionnaire adopted the Likert scale, composed of 20 statements about Statistics, where 10 were positive statements and 10 negative statements, and for data analysis used the Two Parameter Logistic Model of Item Response Theory. The results showed that all 20 items have a good quality and can adequately estimate the level of feeling of students who have a medium feeling in relation to Statistics, between about 1.5 standard deviation around the average. However, there is a lack of items to adequately measure those students who have extreme feelings about Statistics, that is, those who are positioned 1.5 standard deviation above or below average.

Keywords: Likert Scale; Two-parameter Logistic Model; Conceptions; Attitudes; Feelings.

I Universidade Estadual do Oeste do Paraná, Francisco Beltrão, Brasil. E-mail: rosebel.prates@gmail.com.

II Universidade Federal de Santa Maria, Santa Maria, Brasil. E-mail: fmjunior777@yahoo.com.br.

III Universidade Federal de Santa Maria, Santa Maria, Brasil. E-mail: angeladullius@gmail.com.

IV Universidade Federal de Santa Maria, Santa Maria, Brasil. E-mail: angelaansuj@yahoo.com.

V Universidade Tecnológica Federal do Paraná, Ponta Grossa, Brasil. E-mail: rafa191982@hotmail.com.

VI Universidade de Caxias do Sul, Caxias do Sul, Brasil. E-mail: mariaemiliappga@gmail.com.



1 INTRODUÇÃO

Atitude é uma disposição pessoal, indiossincrática, presente em todos os indivíduos, dirigida a objetos, eventos ou pessoas, que assumem diferentes direções e intensidade de acordo com as experiências do indivíduo (Brito, 1996). Tesser e Shaffer (1990), definem atitude como um tripé, formado pelos componentes do domínio afetivo, cognitivo e motor. Conforme Ragazzi (1976) apud Silva (1999), atitude é a prontidão de uma pessoa para responder a determinado objeto de maneira favorável ou desfavorável. Para Zanna & Rempel (1988) atitudes são avaliações baseadas em crenças, sentimentos e/ou comportamento passado, de tal forma que alguém pode ter diferentes atitudes (avaliações) sobre o mesmo objeto ao mesmo tempo. Segundo Klausmeier (1977) apud Moron (1999), o indivíduo forma suas atitudes de acordo com as experiências de aprendizagem e do seu padrão de desenvolvimento.

Concepção é essencialmente o ato de identificar duas ou mais entidades ou eventos dentro de uma relação experimentalmente demarcada. As relações que ocorrem nas concepções podem, mas não é necessário, ter uma conotação avaliativa. (TESSER; SHAFFER, 1990). Para Ponte (1992), conhecer as concepções não é tarefa fácil, pois estas não se revelam facilmente por meio de comportamentos observáveis, pois elas possuem uma natureza essencialmente cognitiva, associadas ao pensar, que atuam como um filtro, dando sentido às coisas ou atuando como um elemento bloqueador para novas situações.

Uma das formas de medir atitudes e concepções é através da Teoria da Resposta ao Item (TRI). A TRI é uma metodologia que propõe uma maneira de representar a relação entre a probabilidade de uma determinada resposta a um item, o traço latente, e as características dos itens, por meio de modelos de probabilidade (ANDRADE; TAVARES; VALLE, 2000; BAKER, 2001; DE AYALA, 2009; EMBRETSON; REISE, 2000; HAMBLETON; SWAMINATHAM, 1985; RECKASE, 2009). Em geral, tem-se: (1) um conjunto de itens, (2) uma amostra de respondentes e (3) um modelo de probabilidade associado, também chamado de Modelo de Resposta ao Item (MRI). Em relação à avaliação tradicional da Teoria Clássica dos Testes (TCT) ou Teoria Clássica da Medida (TCM), onde o traço latente é mensurado por meio de um somatório ponderado, ou

não, dos pontos atribuídos à cada questão do questionário, a TRI apresenta alguns avanços (EMBRETSON; REISE, 2000). Os principais avanços são os procedimentos de equalização e a criação de escalas interpretáveis.

Os procedimentos de equalização foram um avanço para a avaliação educacional, uma vez que permite que indivíduos avaliados por instrumentos de avaliação parcialmente diferentes (com alguns itens em comum) sejam colocados numa mesma escala, o que permite compará-los e acompanhar a sua evolução ao longo do tempo (ANDRADE; TAVARES; VALLE, 2000; EMBRETSON; REISE, 2000). A equalização pode ser feita por meio do método dos Grupos Múltiplos (BOCK; ZIMOWSKI, 1997), se o software permitir, ou por meio de métodos de equalização a posteriori (ANDRADE; TAVARES; VALLE, 2000).

O outro avanço da TRI é a possibilidade de criar escalas onde itens e respondentes são posicionados simultaneamente, o que permite a interpretação dessa escala (ANDRADE; TAVARES; VALLE, 2000). Além desses avanços, a TRI possui outras vantagens, tais como, a possibilidade de calcular a probabilidade de dar certa resposta a um determinado item que o indivíduo não tenha respondido (VENDRAMINI; SILVA; CANALE, 2004) e, na avaliação educacional, a possibilidade de verificar o acerto casual (ANDRADE; TAVARES; VALLE, 2000).

A TRI foi utilizada pela primeira vez no Brasil em 1995 no SAEB (Sistema de Avaliação da Educação Básica), para montagem de instrumentos, tratamento de dados e construção de escalas a partir de resultados apresentados por alunos em provas de rendimento (SOUZA, 2005), o que permitiu que alunos de diferentes séries pudessem ser comparados e colocados na mesma escala, que não era possível fazer através da TCT. A maioria das aplicações tem sido na avaliação educacional (MOREIRA JUNIOR, 2010), onde ENEM (Exame Nacional do Ensino Médio) tem sido, atualmente, o maior exemplo que mostra os benefícios da TRI. Tanto para estimação dos parâmetros dos itens (procedimento conhecido como “calibração dos itens”) quanto para a estimação do traço latente, há vários métodos estatísticos sofisticados que podem ser utilizados, tais como o método de Máxima Verossimilhança Marginal (MVM), o método bayesiano da Moda a Posteriori (MAP), o método bayesiano da Média a Posteriori (EAP) e o

método da Máxima Verossimilhança Conjunta (MVC). Esses métodos não possuem solução explícita, o que torna necessária a utilização de algum método numérico iterativo, como o Algoritmo Newton-Raphson (ISSAC; KELLER, 1966), o Método Scoring de Fisher (RAO, 1973) e o Algoritmo EM (DEMPSTER; LAIRD; RUBIN, 1977). Essas soluções envolvem cálculos bastante complexos e, conseqüentemente, necessitam de programas computacionais específicos.

Esse estudo analisou os itens do questionário “atitudes e concepções dos alunos em relação à Disciplina de Estatística” ministrada em diversos cursos de graduação da Universidade Federal de Santa Maria (UFSM) por meio do Modelo Logístico de Dois Parâmetros (ML2) da Teoria da Resposta ao Item (TRI).

2 METODOLOGIA

A amostra do estudo foi composta por 256 alunos da UFSM que cursaram pelo menos alguma disciplina de Estatística. O questionário foi composto por 20 itens que foram avaliados em relação à concordância em uma escala *likert* de quatro pontos: concordo totalmente, concordo, discordo e discordo totalmente. Metade dos itens do questionário foi aplicada de forma invertida, ou seja, entre as 20 afirmações a respeito da Estatística, 10 eram afirmações positivas e 10 eram afirmações negativas. Para o tratamento dos dados, as respostas dos itens com afirmações negativas foram invertidas. Dessa forma, a escala mede o sentimento crescente de negativo para positivo em relação à Ciência Estatística.

Os dados foram analisados por meio do Modelo Logístico de Dois Parâmetros (ML2) da TRI (Birnbau, 1968) disponível no pacote *irt* do *Software R*. Para essa análise, os dados foram dicotomizados em: (1) concordo (categorias concordo totalmente e concordo) e (0) discordo (categorias discordo e discordo totalmente).

2.1 Modelo Logístico de Dois Parâmetros

Existem vários modelos matemáticos utilizados na TRI, diferentes quanto à sua função e à quantidade de parâmetros, e cada um deles é específico para uma (ou mais) situação. Esses modelos podem ser classificados quanto à sua dimensão (unidimensionais ou multidimensionais), quanto ao tipo de traço latente (cumulativo

ou não cumulativo), quanto ao tipo de item (dicotômico ou politômico) e quanto ao número de populações envolvidas (MOREIRA JUNIOR, 2011). Nesse trabalho será dada ênfase ao Modelo Logístico Unidimensional de Dois Parâmetros (ML2), que é um modelo unidimensional, dicotômico e para uma única população.

O ML2, desenvolvido por Birnbaum (1968), é representado pela seguinte equação:

$$P(U_{ij} = 1/\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_i)}},$$

onde:

a_i é o parâmetro de discriminação do item i , proporcional à inclinação da curva no ponto b_i , onde está localizado o item i ;

b_i é o parâmetro de dificuldade ou de posição do item i , e representa o valor do traço latente θ (no caso, nível de satisfação do aluno) na escala onde há 0,50 de probabilidade do indivíduo j (no caso, o aluno j) responder positivamente ao item;

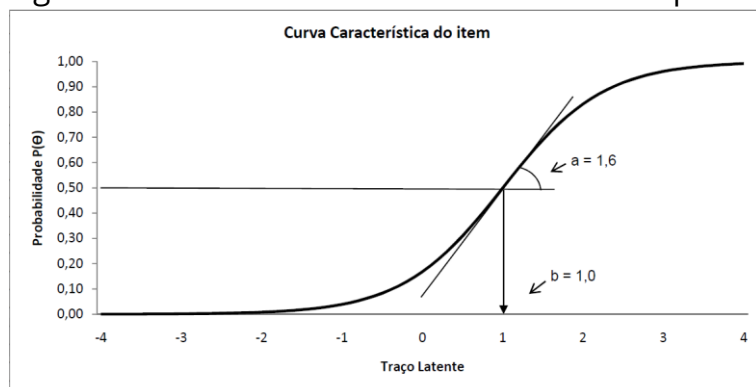
U é a resposta ao item, que pode ser positiva ($U=1$), no caso, se o aluno considera o item satisfatório, ou negativa ($U=0$), caso contrário;

θ_j representa o valor do traço latente (no caso, nível de satisfação do aluno).

A relação entre a resposta prevista ao item e o traço latente do indivíduo é conhecido através da Curva Característica do Item (CCI). A Figura 1 apresenta um exemplo de uma CCI de um ML2 e a influência do parâmetro a_i (poder de discriminação) na inclinação da curva, e do parâmetro b_i (posição do item na escala). O eixo X indica o valor do traço latente (habilidade), considerando uma escala com média 0 e desvio padrão 1, enquanto que o eixo Y indica a probabilidade do indivíduo j responder positivamente o item dicotômico i , dado que ele tem habilidade θ_j .

Em uma comparação com a TCT, o parâmetro de dificuldade da TRI (b_i) está relacionado ao parâmetro de dificuldade da TCT (proporção de respostas positivas); o parâmetro de discriminação da TRI (a_i) está relacionado ao coeficiente de correlação Ponto Bisserial da TCT; e o parâmetro do traço latente (θ_j) está associado ao escore (contagem de respostas positivas) do indivíduo na TCT.

Figura 1 – Curva Característica de um item hipotético



No caso de medição da satisfação do aluno, o eixo x, da Figura 1, representa o nível de satisfação no qual é possível posicionar os itens e os alunos, o que possibilita verificar quais os itens que são satisfatórios para determinado aluno.

A contribuição de um item é a informação que ele fornece para o teste, e está diretamente relacionada com o quadrado da discriminação do item, ou seja, quanto maior a discriminação, mais informação o item fornece. No caso no MLU2P, abordado nesse trabalho, a informação do item é obtida pela expressão:

$$I_i(\theta) = a_i^2 P_i(\theta) Q_i(\theta)$$

A Função de Informação do Teste (FIT) é simplesmente a soma das informações fornecidas por cada item respondido pelo respondente, ou seja:

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

Segundo Andrade, Tavares e Valle (2000), os valores da escala do traço latente (ou habilidade) podem assumir teoricamente qualquer valor real entre $-\infty$ e $+\infty$, diferentemente da TCT, onde a escala geralmente varia entre 0 e a quantidade total de questões do teste. No entanto, o importante nessa escala não é a sua magnitude, mas as relações de ordem existentes. Dessa forma, é preciso estabelecer valores para a média e para o desvio padrão que representam a escala de habilidade dos indivíduos na população. A calibração dos itens, devido à facilidade computacional, geralmente é feita na escala (0,1), ou seja, numa escala com média igual a zero e desvio padrão igual a 1, onde os parâmetros b mais adequados devem variar entre -3 e 3 e os parâmetros a devem ser maior que 0,7. Após essa etapa, é realizada a construção da escala do traço

latente, que é colocada na mesma métrica dos itens. Conforme, Fontanive, Elliot e Klein (2007), as escalas de habilidade ordenam o desempenho dos indivíduos do menor para o maior de forma contínua e são cumulativas, isto é, os indivíduos que situam-se em um determinado nível da escala são capazes de demonstrar as habilidades descritas nesse nível e nos níveis anteriores dessa escala. No entanto, na prática, muitas pessoas não habituadas com a questão da mensuração matemática possuem dificuldades em compreender os valores negativos e decimais que existem na escala (0,1), onde muitas vezes os valores das proficiências são interpretados inadequadamente (VALLE, 2001). Nesse sentido, é bastante comum fazer uma transformação linear em todos os parâmetros envolvidos antes da construção das escalas, fazendo com que todos os valores fiquem positivos e inteiros, de forma a representar a habilidade em valores numéricos que tornem mais fácil o entendimento.

A construção da escala de habilidade é efetuada após a calibração (e equalização, se necessário) dos itens, com o objetivo de encontrar uma interpretação qualitativa dos valores obtidos pela aplicação do modelo da TRI, possibilitando assim, a interpretação pedagógica dos valores das habilidades. Nesse sentido, surge a ideia dos níveis âncoras e a técnica conhecida como ancoragem (BEATON; ALLEN, 1992). Segundo, Andrade, Tavares e Valle (2000) definem níveis âncora como pontos selecionados pelo analista na escala da habilidade para serem interpretados pedagogicamente. Valle (2001) ressalta que esses níveis âncoras não podem ser muito próximos nem muito distantes, podendo-se tomar como base a média e o desvio padrão. Usualmente, os níveis âncoras são estabelecidos em função do desvio padrão em relação à média, obtendo-se, normalmente, cerca de até 7 níveis âncoras ($-3\sigma + \mu$, $-2\sigma + \mu$, $-1\sigma + \mu$, μ , $1\sigma + \mu$, $2\sigma + \mu$, $3\sigma + \mu$). Outras configurações de níveis âncoras podem ser definidas (por exemplo, KLEIN; FONTANIVE; ELLIOT, 2007), porém não são muito utilizadas na prática. No entanto, não se sabe a priori quantos níveis âncoras serão representados, já que isso depende da quantidade de itens âncoras.

Kolen e Brennan (1995) definem item âncora da seguinte forma: considere dois níveis âncora consecutivos Y e Z sendo que $Y < Z$. Um determinado item é âncora para o nível Z se e somente se as 3 condições abaixo forem satisfeitas simultaneamente:

$$\begin{aligned} P(U = 1 | \theta = Z) &\geq 0,65, \\ P(U = 1 | \theta = Y) &\leq 0,50 \text{ e} \\ P(U = 1 | \theta = Z) - P(U = 1 | \theta = Y) &\geq 0,30. \end{aligned}$$

Isso significa que, para um item ser considerado âncora em um determinado nível âncora, ele precisa ser respondido positivamente por um grande percentual de indivíduos (pelo menos 65%) com este nível de habilidade e por um percentual menor de indivíduos (no máximo 50%) com o nível de habilidade imediatamente anterior. Além disso, a diferença entre essas duas proporções deve ser de pelo menos 0,30. Assim, para um item ser âncora em um nível, ele deve ser um item “típico” desse nível, ou seja, bastante respondido positivamente por indivíduos com aquele nível de habilidade e pouco respondido positivamente por indivíduos com um nível de habilidade imediatamente inferior (ANDRADE; TAVARES; VALLE, 2000). Na prática, às vezes um item se caracteriza por ser “quase âncora”, ao violar “levemente” uma das três condições necessárias. Nessas situações, muitos autores consideram esse item como sendo âncora, se ele for importante ou se existirem poucos itens no instrumento de pesquisa. Valle (2001) salienta que alguns níveis âncoras extremos podem ser mal caracterizados por serem definidos por itens muito fáceis ou muito difíceis, os quais geralmente são poucos.

Depois de identificados os itens âncoras de cada nível âncora, os especialistas no traço latente estudado devem caracterizar os níveis âncoras, segundo o conteúdo abordado no conjunto de itens que compõem cada nível. Após essa etapa, a escala está pronta para ser utilizada, por exemplo, para o posicionamento das populações ou dos indivíduos a fim de verificar os conteúdos dominados ou para identificar o percentual de indivíduos em cada nível de habilidade (VALLE, 2001). Assim, um sujeito que se situa acima de um determinado nível dominará os conteúdos caracterizados por esse nível e pelos níveis inferiores à esse, porém não dominará os assuntos dos níveis superiores a esse.

3 RESULTADOS E COMENTÁRIOS

Os parâmetros dos itens foram estimados por meio do ML2 em uma escala com média 0 (zero) e desvio padrão 1 (um). Todos os itens mostraram um comportamento adequado, como pode-se observar nas Curvas Características dos Itens (CCI's) da Figura 2a. Além disso, os itens possuem boa qualidade, como mostram as Funções de Informação dos Itens (FII's) da Figura 2b.

A Tabela 1 apresenta a descrição dos itens e os respectivos parâmetros de discriminação (a) e de dificuldade (b) estimados pelo ML2. Os itens invertidos estão sinalizados por (*). A Figura 2 apresenta as CCI's dos itens na ordem em que aparecem na Tabela 1. Pode-se observar que o item que possui maior probabilidade de concordância foi o Item 1 ($b = -0,9911$), ou seja, a grande maioria dos alunos não fica sob tensão em uma aula de Estatística. Por outro lado, observou-se que o item que possui menor probabilidade de concordância foi o Item 18 ($b = 2,2034$), ou seja, a aula de Estatística, em relação às outras disciplinas, não é a aula que deixa os alunos mais felizes.

Figura 2 – Curvas Características e Funções de Informação dos Itens

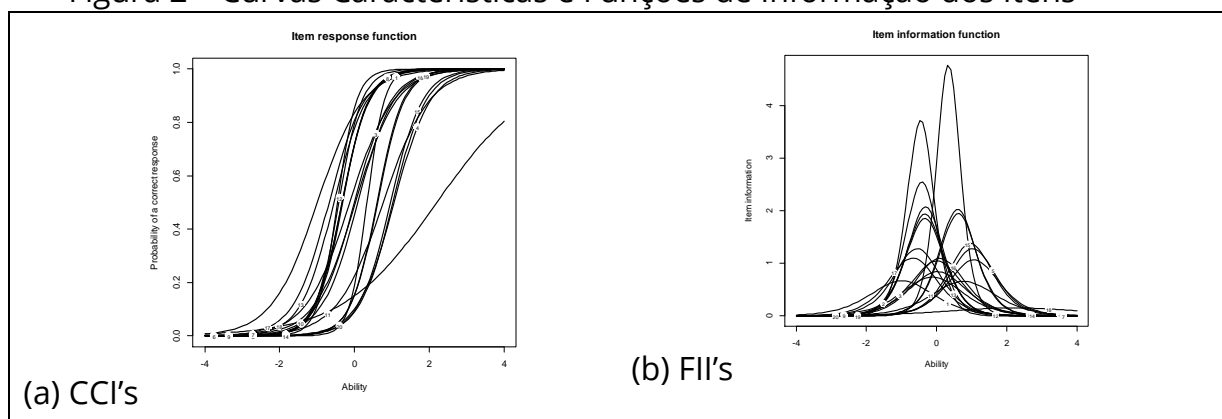


Tabela 1 – Descrição dos itens e os parâmetros estimados do ML2

Descrição dos Itens	Parâmetro a	Parâmetro b
1. Eu fico sob uma terrível tensão na aula de Estatística. *	1,6378	-0,9911
2. Eu não gosto de Estatística e me assusta ter que fazer essa matéria. *	3,8632	-0,4569
3. Eu acho a Estatística muito interessante e gosto das aulas de Estatística.	2,0475	0,0003
4. A Estatística é fascinante e divertida.	2,0657	1,0602
5. A Estatística me faz sentir seguro(a) e é, ao mesmo tempo, estimulante.	2,2648	1,0273
6. "Dá um branco" na minha cabeça e não consigo pensar claramente quando estudo Estatística. *	1,7231	-0,1252
7. Eu tenho a sensação de insegurança quando me esforço em Estatística. *	1,8360	0,0293
8. A Estatística me deixa inquieto(a), descontente, irritado(a) e impaciente. *	2,7876	-0,3322
9. O sentimento que eu tenho com relação à Estatística é bom.	2,8827	-0,3177
10. A Estatística me faz sentir como se estivesse perdido(a) em uma selva de números e sem encontrar saída. *	2,7256	-0,3337
11. A Estatística é algo que eu aprecio grandemente	1,6243	0,7675
12. Quando eu ouço a palavra Estatística, eu tenho um sentimento de aversão. *	3,1960	-0,4213
13. Eu encaro a Estatística com um sentimento de indecisão, que é resultado do medo de não ser capaz em Estatística. *	2,2637	-0,5507
14. Eu gosto realmente de Estatística	2,8528	0,5997
15. A Estatística é uma das matérias que eu realmente gosto de estudar na universidade	2,3492	0,9634
16. Pensar sobre a obrigação de resolver um problema estatístico me deixa nervoso(a). *	2,0940	0,0921
17. Eu nunca gostei de Estatística e é a matéria que me dá mais medo. *	2,0982	-0,6793
18. Eu fico mais feliz na aula de Estatística que na aula de qualquer outra matéria	0,7868	2,2034
19. Eu me sinto tranquilo(a) em Estatística e gosto muito dessa matéria.	2,7916	0,6171
20. Eu tenho uma reação definitivamente positiva com relação a Estatística. Eu gosto e aprendo essa matéria.	4,3696	0,3367

(*) Itens Invertidos

A Figura 4 apresenta a Informação Total do Teste – FIT (Figura 4a) e o Histograma da distribuição do traço latente estimado (Figura 4b), ou seja, do nível de sentimento estimado do aluno numa escala com média zero e desvio padrão 1. O traço latente foi estimado por meio do método bayesiano da Esperança a Posteriori (EAP). Nota-se, na Figura 4a, que a informação do teste concentra-se em torno de zero e havendo pouca informação fora do intervalo entre -1,5 e 1,5 da escala. Isso significa que o questionário carece de itens para avaliar com precisão os indivíduos que estão posicionados fora desse intervalo, como observa-se no Histograma da Figura 4b. No entanto, o questionário consegue discriminar indivíduos que têm sentimentos positivos em relação à estatística daqueles que têm sentimentos negativos, principalmente entre aqueles posicionados na região central da escala. Observa-se que alguns itens parecem medir a mesma coisa, como, por exemplo, os itens 8 e 10, que possuem quase os

mesmos valores para os parâmetros estimados (vide Tabela 1). Nesse sentido, o sentimento negativo de “estar perdido” parece estar bastante relacionado com os sentimentos de inquietude, descontentamento, irritação e impaciência.

Figura 3 – Curvas Características dos Itens

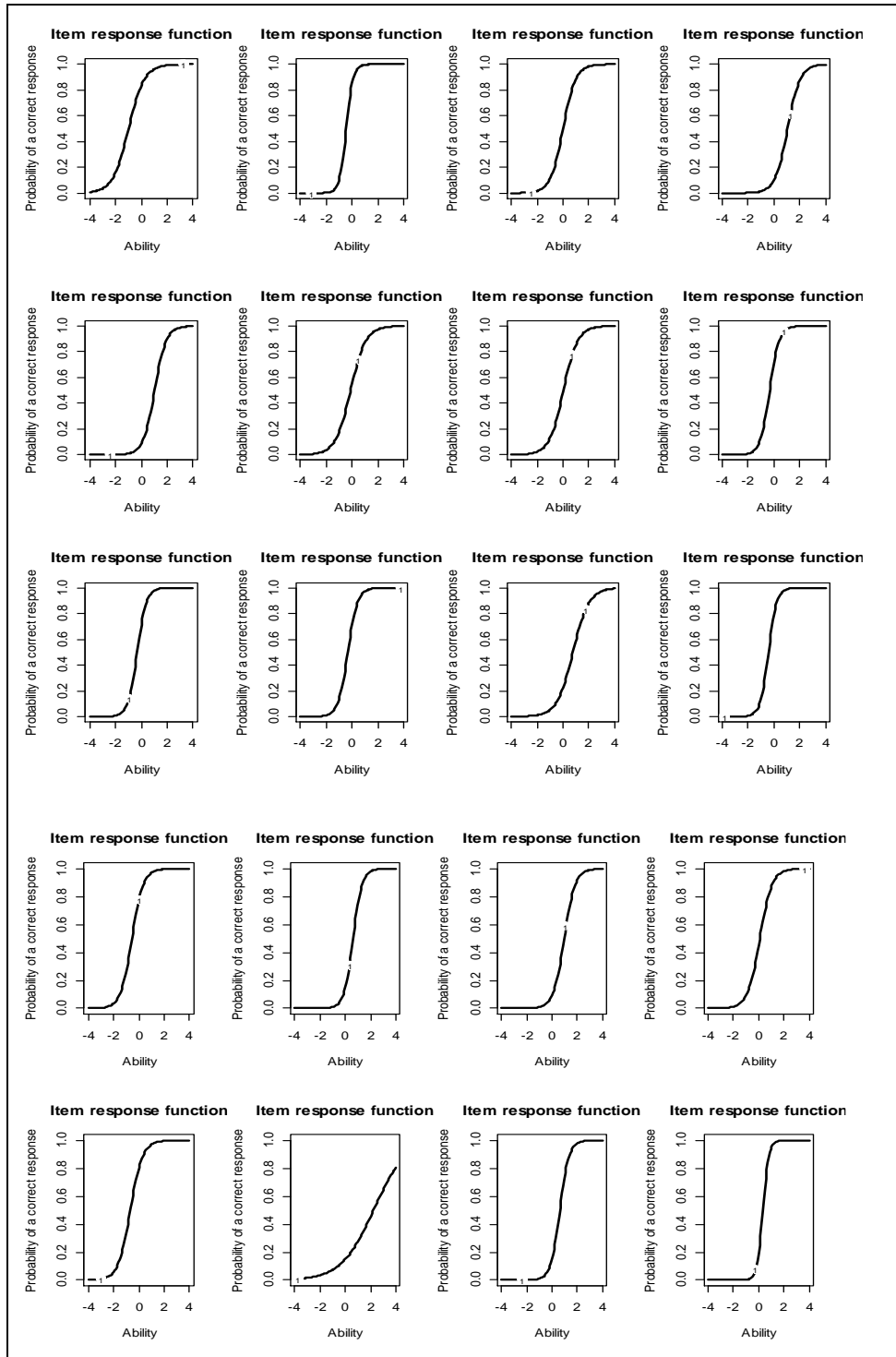
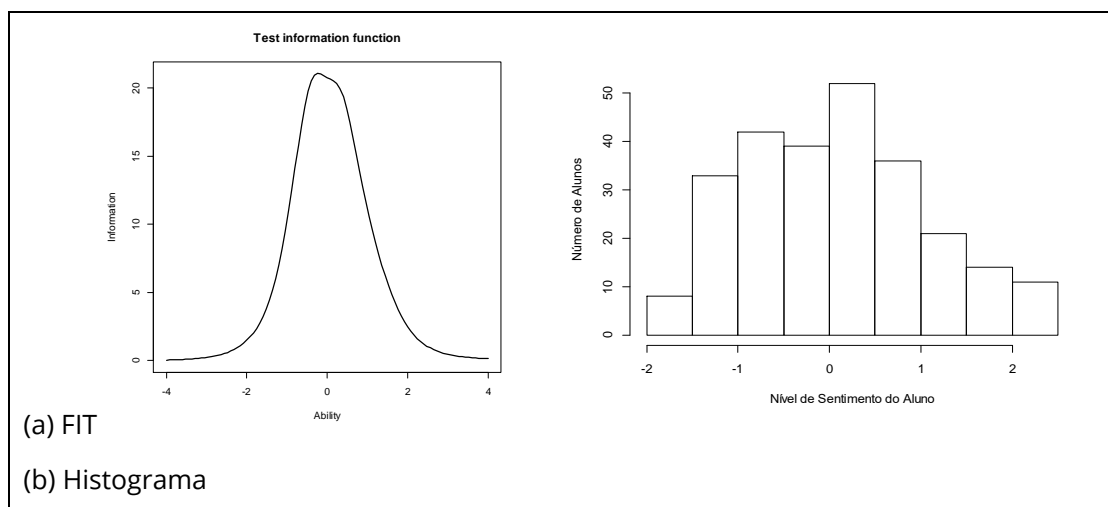


Figura 4 – Informação Total do Teste e Histograma do Traço Latente



Na identificação dos itens âncoras, para fins de interpretação da escala, quase todos os itens foram identificados como itens âncoras, com exceção dos itens 6, 11, 15 e 18 identificados como itens quase âncoras. No entanto, todos eles foram considerados na interpretação da escala, apresentada na Tabela 2. Nessa tabela, para cada nível identificado, foi assinalado "SIM" se o indivíduo que está naquele nível possui aquela característica e "NÃO" se o indivíduo não apresenta aquela característica.

Nota-se que os itens invertidos apresentam uma sentença de sentimento negativo, dessa forma, quanto mais acima na escala o indivíduo estiver, menos sentimentos negativos ele apresentará e mais sentimentos positivos ele apresentará, e vice-versa.

Foram caracterizados 4 níveis com diferentes comportamentos de sentimentos com relação à Estatística, os quais foram descritos abaixo:

- Nível -1 (Aversão à Estatística): Os indivíduos que estão no nível -1 ou abaixo dele apresentam todos os sentimentos negativos e nenhum sentimento positivo com relação à Estatística. Esses indivíduos têm total aversão à Estatística.

- **Nível 0 (Indiferença à Estatística):** Os indivíduos que estão no nível 0 apresentam apenas dois sentimentos negativos mas apenas 1 sentimento positivo com relação à Estatística. Esses indivíduos são indiferentes à estatística, pois praticamente quase não apresentam sentimentos negativos ou positivos com relação à estatística.

- **Nível 1 (Aceitação moderada à Estatística):** Já os indivíduos que estão no nível 1 não apresentam nenhum sentimento negativo e 6 sentimentos positivos com relação à Estatística. A partir desse nível, os alunos já não apresentam sentimentos negativos com relação à Estatística.

- **Nível 2 (Gosto pela Estatística):** Os indivíduos que estão no nível 2 não apresentam nenhum sentimento negativo e ainda apresentam 9 sentimentos positivos com relação à Estatística. Esses indivíduos gostam muito de Estatística.

- **Nível 3 (Paixão pela Estatística):** Os indivíduos que estão no nível 3 ou acima apresentam nenhum sentimento negativo e todos os sentimentos positivos com relação à Estatística, resultado totalmente oposto dos indivíduos que estão no nível -1 ou abaixo. No entanto, não foi identificado nenhum indivíduo nesse nível nessa amostra, como pode-se observar no histograma da Figura 4b. Ainda, esse nível não apresenta grandes diferenças com relação ao nível anterior, uma vez que há mudança de sentimento em apenas um item do instrumento (Eu fico mais feliz na aula de Estatística que na aula de qualquer outra matéria). Na amostra analisada, a aula de Estatística não parece ser a aula preferida dos alunos.

Tabela 2 – Interpretação da escala

Descrição dos Itens	-1 ou -	0	1	2	3 ou +
1. Eu fico sob uma terrível tensão na aula de Estatística. *	Sim	Não	Não	Não	Não
2. Eu não gosto de Estatística e me assusta ter que fazer essa matéria. *	Sim	Não	Não	Não	Não
3. Eu acho a Estatística muito interessante e gosto das aulas de Estatística.	Não	Não	Sim	Sim	Sim
4. A Estatística é fascinante e divertida.	Não	Não	Não	Sim	Sim
5. A Estatística me faz sentir seguro(a) e é, ao mesmo tempo, estimulante.	Não	Não	Não	Sim	Sim
6. "Dá um branco" na minha cabeça e não consigo pensar claramente quando estudo Estatística. *	Sim	Não	Não	Não	Não
7. Eu tenho a sensação de insegurança quando me esforço em Estatística. *	Sim	Sim	Não	Não	Não
8. A Estatística me deixa inquieto(a), descontente, irritado(a) e impaciente. *	Sim	Não	Não	Não	Não
9. O sentimento que eu tenho com relação à Estatística é bom.	Não	Sim	Sim	Sim	Sim
10. A Estatística me faz sentir como se estivesse perdido(a) em uma selva de números e sem encontrar saída. *	Sim	Não	Não	Não	Não
11. A Estatística é algo que eu aprecio grandemente	Não	Não	Sim	Sim	Sim
12. Quando eu ouço a palavra Estatística, eu tenho um sentimento de aversão. *	Sim	Não	Não	Não	Não
13. Eu encaro a Estatística com um sentimento de indecisão, que é resultado do medo de não ser capaz em Estatística. *	Sim	Não	Não	Não	Não
14. Eu gosto realmente de Estatística	Não	Não	Sim	Sim	Sim
15. A Estatística é uma das matérias que eu realmente gosto de estudar na universidade	Não	Não	Não	Sim	Sim
16. Pensar sobre a obrigação de resolver um problema estatístico me deixa nervoso(a). *	Sim	Sim	Não	Não	Não
17. Eu nunca gostei de Estatística e é a matéria que me dá mais medo. *	Sim	Não	Não	Não	Não
18. Eu fico mais feliz na aula de Estatística que na aula de qualquer outra matéria	Não	Não	Não	Não	Sim
19. Eu me sinto tranqüilo(a) em Estatística e gosto muito dessa matéria.	Não	Não	Sim	Sim	Sim
20. Eu tenho uma reação definitivamente positiva com relação a Estatística. Eu gosto e aprendo essa matéria.	Não	Não	Sim	Sim	Sim

(*) Itens Invertidos

4 CONCLUSÕES

Os resultados mostraram que todos os itens apresentaram um comportamento adequado, em relação aos parâmetros estimados por meio do ML2. No entanto, o questionário precisaria ser composto por mais itens que envolvam atitudes e concepções extremas em relação à Estatística (aqueles que gostam muito de Estatística e aqueles que odeiam) a fim de estimar melhor o nível de sentimento dos alunos que estão posicionados numa distância de 1,5 desvio padrão em relação à média.

Na identificação dos itens âncoras, para fins de interpretação da escala, quase todos os itens foram identificados como itens âncoras, com exceção de 4 itens identificados como itens quase âncoras. No entanto, todos eles foram considerados na interpretação da escala, a qual caracterizou 4 níveis com diferentes comportamentos de sentimentos com relação à Estatística.

REFERÊNCIAS

- ANDRADE DF, TAVARES HR, VALLE RC. **Teoria da resposta ao item: conceitos e aplicações**. São Paulo: ABE - Associação Brasileira de Estatística, 2000.
- BAKER FB. **The Basics of Item Response Theory**. 2 ed. USA: ERIC Clearinghouse on Assessment and Evaluation, 2001.
- BEATON AE, ALLEN NL. Interpreting Scales through Scale Anchoring. **Journal of Educational Statistics**, n. 17, p. 191-204, 1992.
- BIRNBAUM A. Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In: LORD FM, NOVICK, MR. **Statistical Theories of Mental Test Scores**. Reading, MA: Addison-Wesley, 1978.
- BOCK RD. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. **Psychometrika**, 1972;37:29-51.
- BRITO MRF. **Um Estudo sobre as Atitudes em Relação à Matemática em estudantes de 1º e 2º graus**. FE-UNICAMP. Campinas-SP. Tese de Livre Docência, 1996.
- DE AYALA RJ. **The Theory and Practice of Item Response Theory**. New York, USA: The Guilford Press, 2008.
- DEMPSTER AP, LAIRD NM, RUBIN DB. Maximum likelihood from incomplete data via the EM algorithm (with discussion). **Journal of the Royal Statistical Society**. 1977; 39 Series B:1-38.
- EMBRETSON SE, REISE SP. **Item Response Theory for Psychologists**. New Jersey, USA: Lawrence Erlbaum Associates, 2000.
- HAMBLETON RK, SWAMINATHAN H. **Item Response Theory: Principles and Applications**. Boston: Kluwer Nijoff; 1985.
- KLEIN R, FONTANIVE NS, ELLIOT LG. O Exame Nacional do Ensino Médio – Tecnologia e Principais Resultados Em 2005. **REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**. 2007;5(2e):116-131.
- KOLEN MJ, BRENNAN RL. **Test Equating - Methods and Practices**. New York, USA: Springer, 1995.
- ISSAC E, KELLER HB. **Analysis of Numerical Methods**. New York: Wiley & Sons, 1966.
- MOREIRA JUNIOR FJ. Aplicações da Teoria da Resposta ao Item (TRI) no Brasil. **Revista Brasileira de Biometria**. 2010;28(4):137-170.

MOREIRA JUNIOR, FJ. **Sistemática para a Implantação de Testes Adaptativos Informatizados baseados na Teoria da Resposta ao Item**. 2011. 334 f. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2011.

MORON, C. F. **As Atitudes e as concepções dos professores de educação infantil com relação à matemática**. Zetetiké –CEMPEM – FE/UNICAMP. 1999;7(1);87-102.

PONTE JP. Concepções dos professores de matemática e processo de formação. In: Brown, M, et al. **Educação Matemática: Temas de Investigação**. Lisboa: Instituto de Inovação Nacional e Secção de Educação e Matemática, 1992.

RAO CR. **Linear Statistical Inference and Its Applications**. New York: Wiley & Sons, 1973.

RECKASE MD. Multidimensional Item Response Theory. Statistical for Social and Behavioral Sciences. **New York: Springer Science+Business Media, LLC**, 2009.

SILVA CB, CAZORLA IM, BRITTO MRF. Concepções e atitudes em relação à Estatística. **Anais da Conferência Internacional Experiências e Perspectivas do Ensino de Estatística: desafios para o século XXI**. 1999; 18-29. 2006 jul 17-21; Florianópolis, SC,

SOUZA SZ. 40 Anos de Contribuição à Avaliação Educacional. **Estudos em Avaliação Educacional**. 2005;16(31):7-36.

TESSER A, SHAFFER DR. **Attitudes Change. Annual Review of Psychology**. 1990;41:479-523.

VALLE RCA. Construção e a Interpretação de Escalas de Conhecimento – Considerações Gerais e uma Visão do que vem sendo feito no SARESP. **Estudos em Avaliação Educacional**. 2001(23):71-92.

VENDRAMINI CMM, SILVA MC, CANEL M. Análise de Itens de uma Prova de Raciocínio Estatístico. **Psicologia em Estudo**. 2004;9(3):487-498.

ZANNA MP, REMPEL JK. Attitudes: a new look at an old concept. In The Social Psychology of Knowledge, ed. **D. Bar-Taj, A. Kruglanski**. New York: Cambridge Univ. Press. 1988.