



**ISMAEL SIMÃO**

**TRI APLICADA À ANÁLISE DE ITENS PARA  
O ENSINO DE ESTATÍSTICA**

**LAVRAS - MG  
2016**

**ISMAEL SIMÃO**

**TRI APLICADA À ANÁLISE DE ITENS PARA O ENSINO  
DE ESTATÍSTICA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Orientador

Dr. Marcelo Silva de Oliveira

Co-orientador

Dr. Paulo Henrique Sales Guimarães

**LAVRAS - MG**

**2016**

**Ficha catalográfica elaborada pelo Sistema de Geração de  
Ficha Catalográfica da Biblioteca Universitária da UFLA,  
com dados informados pelo(a) próprio(a) autor(a).**

Simão, Ismael.

Tri Aplicada a Análise de Itens para o Ensino de Estatística /  
Ismael Simão. - Lavras: UFLA, 2016.

53 p.: il.

Dissertação (mestrado acadêmico) – Universidade Federal de  
Lavras, 2016.

Orientador: Marcelo Silva de Oliveira.

Bibliografia.

1. TRI. 2. Estatística Educacional. 3. Banco de itens I.  
Universidade Federal de Lavras. II. Título.

**ISMAEL SIMÃO**

**TRI APLICADA À ANÁLISE DE ITENS PARA O ENSINO  
DE ESTATÍSTICA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 18 de abril de 2016.

Dr. Ulisses Azevedo Leitão

UFLA

Dr. Eric Batista Ferreira

UNIFAL

Dr. Marcelo Silva de Oliveira  
Orientador

**LAVRAS - MG  
2016**

## AGRADECIMENTOS

Primeiramente a Deus, por me orientar, dar persistência, renovar minhas forças, me suprir em todas as minhas necessidades e que se mostra presente nos momentos difíceis.

Ao meu orientador Professor Dr. Marcelo Silva de Oliveira por acreditar em mim, me mostrar a direção para atingir os meus objetivos nesse trabalho. Ao co-orientador Professor Dr. Paulo Henrique Sales Guimarães que não mediu esforços em colaborar para a conclusão deste trabalho.

Aos professores do Departamento de Ciências Exatas e do programa, em especial ao professor Dr. Ulisses Azevedo Leitão, que prestaram, com dedicação, serviços de excelência. Aos colegas que juntos fizemos este caminho e àqueles que não chegaram conosco, mas que de alguma forma me prestaram auxílio.

Aos meus filhos, os quais amo muito, em especial ao Lucas Caixeta Simão agradeço pelo carinho, paciência e incentivos os quais foram fundamentais nesta jornada.

Meus agradecimentos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq pelo financiamento desta pós-graduação.

## RESUMO

A proposta desta pesquisa é a formação e o armazenamento de um repositório de itens selecionados que objetiva a elaboração de instrumentos de avaliação que podem auxiliar o educador da disciplina de Estatística Básica. Para tanto, foram usados os conceitos e fundamentos da Teoria de Resposta ao Item (TRI) na tarefa de elaboração dos instrumentos que contribuam para uma avaliação no processo de ensino-aprendizagem. A arquitetura de um banco de itens calibrados pela TRI depende do parâmetro do modelo adotado. Nesta dissertação será adotado o modelo logístico unidimensional de três parâmetros, que enfatiza as características inerentes ao indivíduo, seu traço latente e as características apresentadas do item: os parâmetros de discriminação ( $a$ ), a dificuldade ( $b$ ) e acerto ao acaso ( $c$ ) e as curvas características do item (CCI). Foi aplicada a mesma em duas turmas diferentes para construir o banco de itens. Esse banco de itens foi obtido através de uma avaliação aplicada ao ensino de nivelamento inerente aos alunos de todos os cursos da UFLA, que é oferecido pelo Departamento de Exatas da mesma. O teste foi aplicado em um total de 50 alunos, divididos em duas turmas, sendo uma com 24 alunos e outra com 26 alunos. As análises foram feitas utilizando o software estatístico livre ICL.

Palavras-chave: TRI, Estatística Educacional, Banco de itens.

## ABSTRACT

The purpose of this research is the development and storage of a collection of selected items that aims to design assessment tools, which may help the educator of the Basic Statistics subject. Therefore, concepts and fundamentals of the Item Response Theory (IRT) were used in order to prepare instruments that contribute to an assessment of the teaching-learning process. The architecture of a collection of items calibrated by IRT depends on the model parameter adopted. On this thesis, the three parameter logistic unidimensional model will be adopted, which emphasizes the characteristics inherent to the individual, his or her latent trait and the characteristics: the discrimination (a), difficulty (b) and pseudo-guessing (c) parameters and the item characteristic curves (ICC). The same curve was used in two different classes to build the collection of items. This collection of items was obtained through a test applied to the students from all courses of the Federal University of Lavras (UFLA) taking the leveling teaching subjects, which is offered by the Department of Exact Sciences. The test was taken by a total of 50 students, divided into two groups, one with 24 students and another with 26. The analyses were performed using the open source statistical software called IRT Command Language.

Keywords: IRT; Educational statistics; Bank of items.

## LISTA DE FIGURAS

1	Curva característica do item (CCI). . . . .	19
2	Ementa da disciplina Estatística. . . . .	25
3	Gráfico de barras da proporção do nível de dificuldades das provas 1 e 2. . . . .	33
4	Curva CCI do item 1 da prova 1 ( $a = 0,682128, b = -3,075034,$ $c = 0,228190$ ). . . . .	34
5	Curva CCI do item 24 da prova 1 ( $a = 0,296090, b = -2,908130,$ $c = 0,229041$ ). . . . .	36
6	Curva CCI do item 8 da prova 1 ( $a = 0,959081, b = -1,100258,$ $c = -0,212807$ ). . . . .	37
7	Curva CCI do item 8 da prova 2 ( $a = 0,880697, b = -5,493616,$ $c = 0,227457$ ). . . . .	38
8	Curva CCI do item 1 da prova 2 ( $a = 0,880697, b = -5,493616,$ $c = 0,227457$ ). . . . .	39

## LISTA DE TABELAS

1	Estimativa dos parâmetros dos itens análise do ICL da prova 1.	29
2	Estimativa dos parâmetros dos itens análise do ICL da prova 2.	30
3	Escala de dificuldade utilizada como referência no trabalho. .	31
4	Dados prova 1. . . . .	49
5	Dados prova 2. . . . .	50

## SUMÁRIO

1	INTRODUÇÃO . . . . .	9
2	REVISÃO DE LITERATURA . . . . .	11
2.1	O ensino da Estatística. . . . .	11
2.2	Itens de múltipla escolha. . . . .	13
2.3	Teoria de resposta ao item - TRI . . . . .	14
2.3.1	Modelos da TRI. . . . .	16
2.3.2	Modelos unidimensionais dicotômicos . . . . .	17
2.3.3	Modelo logístico de três parâmetros. . . . .	17
2.3.4	Função de informação do item -FII . . . . .	21
2.4	Escala de habilidades na TRI. . . . .	22
2.5	Estimação dos parâmetros . . . . .	22
3	MATERIAL E MÉTODOS. . . . .	24
3.1	Material . . . . .	24
3.2	ICL, estimação . . . . .	26
3.3	Métodos . . . . .	27
4	RESULTADOS E DISCUSSÃO. . . . .	28
4.1	Resultados dos alunos em cada teste . . . . .	28
4.2	Estimação dos parâmetros dos itens da prova 1 pelo o <i>software</i> ICL . . . . .	28
4.3	Estimação dos parâmetros dos itens da prova 2 pelo o <i>soft-</i> <i>ware</i> ICL . . . . .	29
4.4	Discussão dos resultados . . . . .	31
5	CONCLUSÕES . . . . .	41
	REFERÊNCIAS . . . . .	42
	APÊNDICE . . . . .	46

# 1 INTRODUÇÃO

Atualmente, um grande desafio tem sido imposto aos docentes em relação à forma de avaliação na educação. Esse desafio se faz necessário devido à preocupação em avaliar, de forma eficaz, as habilidades e a compreensão dos estudantes em uma determinada área do conhecimento.

Um instrumento de medida muito utilizado é o teste de avaliação (prova), o qual tem a finalidade de medir o desempenho dos participantes em determinadas situações para realizar inferências sobre o processo educacional em desenvolvimento.

Uma atividade útil para que tais avaliações aconteçam é a manutenção de um repositório de itens que ofereça subsídio para construir os instrumentos de medida, ou seja, uma coleção de itens de testes de natureza específica, organizada segundo determinados critérios, disponíveis para a construção de instrumentos de avaliação.

Tradicionalmente, as avaliações, principalmente no processo de ensino e seleção, têm sido aplicadas utilizando um método chamado Teoria Clássica das Medidas - TCM, onde a mensuração do resultado de cada estudante é dada pela soma dos resultados de sucessos do conjunto de itens que compõe a avaliação (ANDRADRE; TAVARES; VALLE, 2000).

Hoje em dia, métodos que apresentam instrumentos mais precisos de medição, de forma que o discente possa ser avaliado efetivamente em sua capacidade, estão sendo estudados. Um desses métodos é a Teoria de Resposta ao Item - TRI.

Segundo Andrade, Tavares e Valle (2000), a TRI é uma metodologia que sugere formas de representar a relação entre a probabilidade de um indivíduo dar certa resposta a um item e seus traços latentes, que são características do indivíduo não podendo ser observadas diretamente. A TRI surgiu da Psicometria, uma área da Psicologia, que tem por finalidade a busca de um instrumento de avaliação que seja capaz de medir satisfatoriamente os traços latentes de uma determinada pessoa.

Assim, a TRI pode ser entendida como um conjunto de modelos destinados a mensurar a probabilidade de uma pessoa responder, de forma correta,

um dado item (também conhecido como questão), tendo como parâmetros os traços latentes (habilidade) e as características do item avaliado.

Nesse método destaca-se um modelo específico, o modelo de três parâmetros, que mensura o poder de discriminação do item, a dificuldade imposta ao discente pelo item e a sua probabilidade de acerto casual, ou seja, acertar o item sem que o indivíduo tenha a habilidade necessária.

O objetivo deste trabalho está relacionado diretamente à avaliação no processo de ensino e aprendizagem de conteúdos relacionados a disciplina de Estatística Básica nos cursos presenciais de graduação e pós-graduação da Universidade Federal de Lavras - UFLA. Esse objetivo geral foi apresentar uma proposta de construção de um Banco de Itens para o ensino da Estatística (BIE). Como objetivos específicos, buscou-se:

- i. elaborar itens para um pré-teste;
- ii. estimar os parâmetros dos itens;
- iii. selecionar os itens para estruturar o BIE.

## 2 REVISÃO DE LITERATURA

Apresentam-se nesse capítulo algumas informações referentes à importância do ensino da Estatística e à Teoria de Resposta ao Item. Ambos os temas se mostram relevantes para a compreensão deste trabalho.

### 2.1 O ensino da Estatística

O ensino de Estatística nas instituições de ensino superior vem ganhando destaque, em face de sua importância na formação científica e ética do cidadão e seu crescente uso em uma sociedade cada vez mais informatizada. Indispensável no planejamento, na coleta, organização, análise de dados e interpretação, a Estatística é hoje uma ferramenta necessária para qualquer profissional ou para um simples participante desta sociedade moderna que vem passando por grandes transformações tecnológicas e científicas (BARROSO, 2010).

Segundo Cazorla (2002), é crescente a importância da Estatística na formação do cidadão, na medida em que este fica exposto a informações estatísticas veiculadas pela mídia, jornais e revistas. Muitas vezes, pela ausência de conhecimento na área, essas informações são consumidas sem filtragem, tornando o cidadão vulnerável a interpretações que não correspondem à realidade.

Devido às grandes transformações tecnológicas e científicas, o ensino de Estatística se realiza em novo contexto social, buscando desenvolver competências que facilitem novas alternativas para resolver diferentes situações, que interpretem a realidade atual.

O volume de informação disponível é muito grande e isso torna necessário aos indivíduos qualificar, selecionar, analisar e contextualizar essas informações, de modo que possam ser incorporadas às suas próprias experiências (WODEWOTZKI; JACOBINI, 2004).

Segundo Batanero (2001) é indiscutível que o século XX tem sido o século da Estatística, que passou a ser considerada uma das ciências metodológicas fundamentais e base do método científico experimental.

Crianças, jovens ou adultos, que estão diante de dados estatísticos, precisam desenvolver a sua capacidade crítica e de autonomia em relação à informação disponível, a fim de que tenham melhores condições para refletir, emitir opiniões e tomar decisões em sua vida cotidiana (BARROSO, 2010). Nos Parâmetros Curriculares Nacionais do Ensino Fundamental, o ensino da Estatística Básica, que por hora é também denominada Estatística e Probabilidade, aparece no currículo de Matemática, no bloco de conteúdo denominado Tratamento das Informações, acenando para a necessidade de o estudante compreender as informações veiculadas, tomar decisões e fazer previsões que influenciam sua vida pessoal e em comunidade (BRASIL, 1998).

Ministrada, hoje em dia, na formação de profissionais e pesquisadores de várias áreas do conhecimento científico, em quase todos os cursos do ensino superior, tanto em nível de graduação como de pós-graduação, a Estatística na maior parte das vezes é vista pelos alunos do curso superior como uma ciência que está ligada à Matemática.

Pesquisas demonstram que os alunos estão chegando, na maioria das vezes, aos cursos superiores com pouco ou nenhum conhecimento sobre o que é Estatística. De acordo com Cazorla (2002), estas pesquisas revelam que os alunos mostram ter dificuldades mesmo na utilização das ferramentas de análise de dados mais simples como, por exemplo, a representação de dados em gráficos.

Ainda de acordo com Cazorla (2002) existem algumas dificuldades restritas à formação de usuários em Estatística no Ensino Superior e que interferem no processo de ensino-aprendizagem, tais como:

1. a natureza da Estatística: trabalha experimentos aleatórios, com conceitos de natureza probabilística, obrigando o usuário a tomar decisões em meio à incerteza;
2. a fundamentação teórica da Estatística: está na Matemática, campo de dificuldade para muitos estudantes. Há, portanto, alguma transferência deste sentimento;

3. a carga horária da disciplina Estatística: é reduzida, assim abordam-se mais os cálculos e procedimentos, deixando de lado o desenvolvimento do pensamento estatístico.

Assim sendo, professores de Estatística devem refletir sobre o objetivo principal desta disciplina e desenvolver seu trabalho em torno do compreender e do fazer estatístico, trabalhando as atitudes e a motivação dos alunos a fim de que os mesmos possam formar o pensamento estatístico. O primeiro passo para o ensino da estatística é trabalhar a base da mesma, o que é feito em um curso introdutório nas instituições superiores. Tal curso aparece nessas instituições em uma disciplina cujo nome atribuído Estatística Básica, abrange noções de Probabilidade, Estatística Descritiva e noções de Inferência Estatística.

Outros cursos que demandam um conhecimento maior de Estatística, oferecem uma disciplina com maior enfoque em inferência estatística e a área de aplicação do conhecimento específico. Apesar da importância do pensamento estatístico na abordagem dos conteúdos programáticos, a sua presença raramente é percebida.

Segundo Guimarães (2008), a ciência Estatística é aplicável a qualquer área do conhecimento onde se manipulam dados experimentais.

## 2.2 Itens de múltipla escolha

Os testes de avaliação de aproveitamento desempenham uma função importante em sala de aula, uma vez que são frequentemente usados para compor as notas dos alunos. O Banco de itens para o ensino de Estatística (BIE) é composto por itens de múltipla escolha. Esse tipo de questão apresenta várias vantagens, pois são fáceis de marcar, práticos para exames de larga escala e adaptáveis a vários níveis de operações mentais (CATALANI, 2013).

Entretanto, sua elaboração requer bastante cuidado, deve ser bem definido, resultado de aprendizagem, ter uma linguagem direta e adequada em sua formulação. Sempre evitar termos como “exceto”, “não”, “nunca”, “errado”, “somente”, “sempre”, “exclusivamente” (FERNANDES, 2015). Se-

gundo Haladyna, Downing e Rodriguez (2002), um item é composto pelas seguintes partes:

- Enunciado: é usado para estimular o estudante a mobilizar recursos cognitivos para solucionar a situação-problema proposta. O estímulo pode conter um texto, uma imagem ou outros recursos que recebem o nome de comando para resposta.
- Comando para a resposta: pode ser dado sob a forma de complementação ou de interrogação. Ele tem que ser preciso e deve estar totalmente ligado à habilidade que se pretende avaliar.
- Alternativas de resposta: são apresentadas sobre a forma de quatro ou cinco opções, sendo somente uma correta, que é o gabarito. As alternativas que não contemplam a resposta são chamadas de distratores.

Os distratores indicam as alternativas incorretas à resolução da situação-problema proposta. Além disso, essas respostas devem ser plausíveis, isto é, devem parecer corretas para aqueles participantes do teste que não desenvolveram a habilidade em questão (HALADYNA; DOWNING; RODRIGUEZ, 2002).

O que torna os distratores mais plausíveis é o fato de que, em sua elaboração, devem-se utilizar erros comuns observados em situações de ensino-aprendizagem.

De acordo com São Paulo (2008), um cuidado que deve ser tomado é não tornar o distrator uma opção, a qual é chamado de “pegadinha”, ou seja, induzir o aluno ao erro.

### **2.3 Teoria de resposta ao item - TRI**

A Teoria de Resposta ao Item (TRI) surgiu na década de 1950 sendo, no entanto, consolidada no final da década de 1970, com o desenvolvimento de computadores e softwares capazes de realizar análises complexas de dados. A TRI se apresenta como um instrumento de análise formado por um conjunto de métodos estatísticos que, por meio de processos de estimação

de características de cada item aplicado e da aptidão do indivíduo avaliada em um teste, descreve o desempenho e quantifica em uma escala métrica, a habilidade desenvolvida deste indivíduo, sua competência e, portanto, seu escore (SOARES, 2014).

Alguns pesquisadores seguem em seus estudos para aplicações dos conceitos fundamentados por Lazarsfeld (1950) até então conhecido como Teoria do Traço Latente. Tais estudos tiveram dentre seus precursores Lord (1952a,1952b,1953), apresentado na literatura como um dos primeiros pesquisadores que iniciaram os estudos do que fundamentou o que hoje, chama-se Teoria de Resposta ao Item.

Outros trabalhos bastante citados na literatura foram apresentados por Birnbaum (1968), Rasch (1960), Samejima (1969, 1972) e Wright (1968), sendo Birbaum e Wright responsáveis por mudanças significativas nos modelos de análises dos dados. Rasch (1960) estudou o modelo logístico de um parâmetro e Samejima (1969, 1972) estudou a estimação das habilidades.

Em 1995, com a retomada do Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (SARESP), do Governo de São Paulo e do Sistema de Avaliação da Educação Básica (SAEB) do INEP/MEC, do Governo Federal, a TRI começa a ter maior desenvolvimento em pesquisas e aplicações no Brasil.

As primeiras aplicações da TRI no Brasil, segundo Gatti (1996), são referenciadas ao Programa de Avaliação da Secretaria de Estado de Educação do Estado de São Paulo em 1993, mas com limitações na aplicação.

De acordo com Moreira Junior (2010) a TRI é uma poderosa ferramenta estatística que surgiu para suprir as necessidades decorrentes das limitações da Teoria Clássica da Medida (TCM) ou Teoria Clássica do Teste (TCT), teoria que tradicionalmente era, e ainda é, utilizada nas avaliações. Neste contexto, as habilidades do indivíduo são características próprias, independentem do tipo de avaliação, da mesma forma os parâmetros ou características do item também não dependem do indivíduo respondente. Tais propriedades são definidas como Invariância dos Parâmetros do item em uma determinada população, conforme Baker (2001).

Para aplicação de modelos da TRI deve-se considerar aspectos relevantes. Couto e Primi (2011) afirmam que alguns critérios são fundamentais, entre eles a unidimensionalidade e a independência local.

A unidimensionalidade se refere à circunstância em que a avaliação de uma característica psicológica depende unicamente desta característica, no caso de itens, a resposta correta depende exclusivamente da habilidade exigida.

Já a independência local acontece quando a resposta dada a um item por um indivíduo não causa influência na resposta dada a outros itens.

Para Hambleton, Swaminathan e Rogers (1991), juntamente com Pasquali e Primi (2003), há grandes inovações que foram trazidas na TRI em relação aos modelos já existentes e, destas, três são de maior relevância:

- I . a medida da aptidão do indivíduo independe da amostra do item utilizado;
- II . o cálculo dos parâmetros de dificuldade e discriminação do item independem da amostra de indivíduos ou da população utilizada;
- III . a possibilidade de discriminar itens adequados ao nível de aptidão do indivíduo.

### **2.3.1 Modelos da TRI**

Os modelos da TRI adotados neste trabalho foram aqueles em que a probabilidade de resposta certa dada por um indivíduo ao item dependerá do modo crescente do traço latente, e dependerá também dos parâmetros que caracterizam o item.

Nestes modelos serão considerados somente os modelos acumulativos, em que Andrade, Tavares e Valle (2000) enfatizam três fatores principais destes na TRI:

- I. a natureza do item - dicotômicos ou não dicotômicos;
- II. o número de populações - apenas uma ou mais de uma;

- III. a quantidade de traços latentes ou habilidades a serem avaliadas - apenas um ou mais de um.

No desenvolvimento deste trabalho, serão discutidos apenas os modelos relacionados com avaliação de apenas um traço latente, unidimensionais e dicotômicos.

### **2.3.2 Modelos unidimensionais dicotômicos**

Diz-se que um modelo é dicotômico ou dicotomizado quando é aplicado a itens que tenham como resultado de correção somente as opções certo ou errado. Segundo Andrade, Tavares e Valle (2000), tais modelos podem ser utilizados tanto para análise de itens de múltipla escolha, quanto para análise de itens abertos, desde que em ambas as situações, os itens sejam avaliados de forma dicotomizada, isto é, em cada item seja avaliada a resposta dada apenas como certa ou errada.

A classificação destes modelos está relacionada a quantidade de parâmetros envolvidos nos itens em análise:

- I. modelo de um parâmetro - somente o parâmetro de dificuldade;
- II. modelo de dois parâmetros - apresentam parâmetros de dificuldade e discriminação;
- III modelo de três parâmetros - apresentam tanto parâmetros de dificuldades, discriminação e também a possibilidade de acerto casual.

Dos modelos descritos, somente o modelo com três parâmetros será aplicado neste trabalho, uma vez que além de preservar as características dos modelos de um e dois parâmetros, acrescenta o parâmetro do acerto ao acaso, o que torna mais adequado a este estudo.

### **2.3.3 Modelo logístico de três parâmetros**

Segundo Baker (2001), o modelo logístico de três parâmetros é dado por:

$$P(U_{ij} = 1; a, b, c, \theta_j) = c + (1 - c) \frac{1}{1 + e^{-Da_i(\theta - b_i)}} \quad (1)$$

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

Em que:

$I$  representa a quantidade total de ítems,  $n$  representa o total de indivíduos analisados,  $U_{ij}$  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, assumindo valores em toda a reta real;

$a_i$  é o parâmetro do item que representa a discriminação, que também assume valores em toda a reta real positiva;

$b_i$  é o parâmetro do item que representa a dificuldade, podendo assumir valores em toda a reta real. Nesse trabalho, esse parâmetro assume tipicamente valores no intervalo de -4 a +4;

$c_i$  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). Por se tratar de probabilidade, esse parâmetro assume valores em uma escala dentro dos números reais que varia de 0 a 1;

$D$  é um fator de escala. De acordo com Birnbaum (1968), o valor de 1,7 costuma ser utilizado quando se deseja que a função logística forneça resultados semelhantes ao da ogiva normal;

$P(U_{ij} = 1|\theta_j)$  é a probabilidade de um indivíduo  $j$  com habilidade  $j$  responder corretamente o item  $i$  e é chamada de Função de Resposta do Item - FRI.

Ainda, segundo Baker (2001), os parâmetros da TRI e a Curva Característica do Item (CCI - é o gráfico da FRI) podem ser mais bem interpretados como descrito a seguir. Discriminação, indicado em geral por  $a$ , representa o quanto o item pode apresentar característica discriminativa, isto é, a clareza do nível de conhecimento apropriado que o indivíduo necessita para

resolver o item.

Dificuldade, usualmente indicado por  $b$ , é expresso na mesma escala do traço latente  $\theta$ , representa a habilidade adequada para o indivíduo responder com conhecimento necessário e de forma correta o item avaliado. Vejamos.

Na curva CCI na Figura 1,  $b$  é a abscissa do ponto de inflexão, então a melhor discriminação do item ocorre em intervalo mínimo em que o valor de  $b$  é o ponto médio, ou seja, o intervalo no qual o item é mais discriminativo.

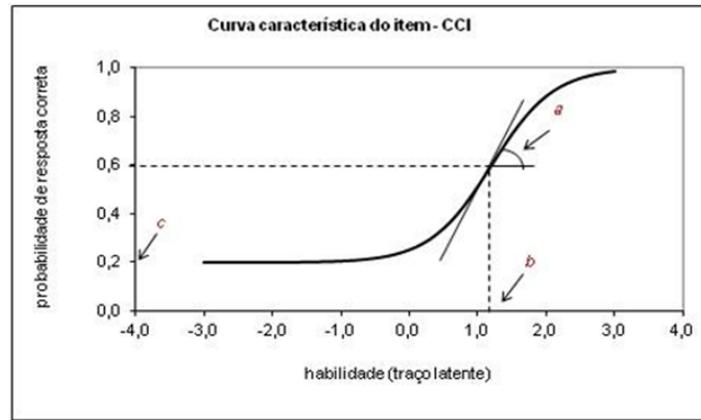


Figura 1 Curva característica do item (CCI).

De fato, seja

$$P(U_{ij} = 1; a, b, c, \theta_j) = c + (1 - c) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}$$

,

tem-se que

$$\frac{dP}{d\theta} = (1 - c_i) \frac{Da_i e^{-Da_i(\theta_j - b_i)}}{(1 + e^{-Da_i(\theta_j - b_i)})^2}$$

$$\frac{dP^2}{d\theta^2} = D^2 a_i^2 e^{-Da_i(\theta_j - b_i)} (1 - c_i) \frac{e^{-Da_i(\theta_j - b_i)} - 1}{(1 + e^{-Da_i(\theta_j - b_i)})^3}$$

Tomando  $\frac{dP^2}{d\theta^2} = 0$ , obtém-se que

$$D^2 a_i^2 e^{-Da_i(\theta_j - b_i)} (1 - c_i) \frac{e^{-Da_i(\theta_j - b_i)} - 1}{(1 + e^{-Da_i(\theta_j - b_i)})^3} = 0$$

Como  $D$ ,  $a_i$  são reais positivos e  $c_i \in ]0, 1[$ , então:

$$\begin{aligned} \implies e^{-Da_i(\theta_j - b_i)} - 1 = 0 &\implies e^{Da_i(\theta_j - b_i)} = 1 \\ \implies -Da_i(\theta_j - b_i) = 0 &\implies \theta_j - b_i = 0 \implies \theta_j = b_i \end{aligned}$$

Assim, o parâmetro  $b$  representa, quando  $\theta_j = b_i$ , a habilidade exigida para que um estudante responda de forma correta o item, tendo como probabilidade de acerto o valor  $P(U_{ij} = 1 | \theta_j = b_i) = \frac{1+c}{2}$ .

O parâmetro de acerto ao acaso, indicado por  $c$ , representa a probabilidade do indivíduo com habilidade mais baixa que a exigida pelo item, responder de forma correta o item avaliado.

Para mostrar matematicamente o efeito dos parâmetros utiliza-se a CCI já definida acima, a qual descreve a relação entre os parâmetros do item no modelo e as probabilidades de acerto. A CCI ideal é uma representação não linear e monótona não decrescente que descreve a relação entre a habilidade de um indivíduo e a probabilidade de acerto ao item proposto.

Pode-se considerar que um item apresenta características importantes de discriminação quando o traço da CCI assume a forma de S. Quando um item apresenta CCI decrescente, há necessidade de reavaliação do item ou descartá-lo como instrumento de avaliação.

Para Couto e Primi (2011), a curva característica do item CCI, conceitualmente, é uma propriedade da TRI e representa a forma como os parâmetros dos itens são apresentados pelas curvas no gráfico, conforme Figura 1.

Um ponto que pode ser notado na CCI é quando a discriminação é alta, a curva tem formato em S e com um crescimento abrupto na região mediana. Além disso, pode-se verificar que a inclinação é mais acentuada na região da escala ao redor da dificuldade do item.

### 2.3.4 Função de informação do item -FII

Na TRI, denomina-se Função de Informação do Item, a função que estima a variabilidade do traço latente em relação a quantidade de informação que o item apresenta. Segundo Baker (2001) e Couto e Primi (2011), tal curva é individual, pois depende do valor do traço latente e, portanto não se pode definir um desvio padrão para todas escalas, uma vez que depende do traço latente de cada indivíduo avaliado.

De acordo com Baker (2001) e Couto e Primi (2011), a função de informação diz o quão melhor cada nível de habilidade pode ser estimado, usando-se um conjunto de itens específico isto é, a função de informação é um indicador de precisão.

Para Couto e Primi (2011), a função de informação do item representa um importante instrumento descritivo para um item, pois possibilita a determinação para qual valor de  $\theta$  o item apresenta maior quantidade de informações, ou ainda, a quantidade de informações que um item apresenta em determinado valor de  $\theta$ .

Para Baker (2001) a função de informação pelo item  $i$  no nível de habilidade  $\theta_j$  é dada por:

$$I_i(\theta) = D^2 a^2 \frac{1 - P(\theta)}{P(\theta)} \left[ \frac{P(\theta) - c_i}{1 - c_i} \right]^2 \quad (2)$$

$P_i(\theta_j) = P(X_y = 1/\theta_j)$  e  $1 - P_i(\theta)$  é a probabilidade de um indivíduo  $j$  com habilidade  $j$  responder incorretamente o item  $i$ .

Em (2), cada item está associado a um intervalo na escala da habilidade no qual o item tem maior poder de discriminação e as funções de informação dos itens são mais precisas, definido em torno do valor do parâmetro  $b_i$  e corresponde à  $(b_i - D; b_i + D)$  (GUEWHER, 2007).

De acordo com Pasquali e Primi (2003) os modelos de um e dois parâmetros atingem sua capacidade máxima de informação quando for igual à dificuldade dos itens ( $\theta = b$ ) e os valores da função de informação são simétricos para cada valor acima ou abaixo do valor de  $\theta = b$ , conforme

demonstrado acima.

Neste ponto o parâmetro de dificuldade, chamado de  $b_i$ , é o ponto na escala de aptidão no qual a probabilidade de uma resposta correta é de 50% (ou seja, 0,5). No modelo de três parâmetros a quantidade de informação será sempre menor que nos dois outros modelos devido à influência do parâmetro probabilidade de acerto ao acaso  $c$ .

Andrade, Tavares e Valle (2000) destacam que um item apresenta maior quantidade de informações quando tem um alto índice discriminativo e baixo índice de acerto ao acaso.

## 2.4 Escala de habilidades na TRI

De acordo com Guewher (2007) a escala da habilidade (traço latente) é uma escala arbitrária, podendo teoricamente assumir qualquer valor real entre  $-\infty$  e  $+\infty$ . Tais valores são atribuídos de modo a possuir um valor médio e um valor de desvio padrão considerados adequados para a população em estudo. A primeira escolha na teoria da TRI é media zero e desvio padrão 1, mas tem-se a possibilidade de se adotar outros valores que tornassem a escala mais da próxima da finalidade do público a que se destina. Por exemplo, no ENEM (Exame Nacional do Ensino Médio) a média adotada é 500 e o desvio padrão é 100, fazendo os valores típicos de habilidades variarem folgadoamente no intervalo de 0 a 1000.

## 2.5 Estimação dos parâmetros

A estimação do parâmetro que representa a habilidade do indivíduo respondente e dos parâmetros do item aplicado a este indivíduo, que na TRI é chamada de calibração dos parâmetros do item, representa uma das etapas determinantes no processo de análise e interpretação dos resultados.

Andrade, Tavares e Valle (2000) destacam que nesta etapa há dois conjuntos de parâmetros a serem analisados: as habilidades do indivíduo respondente e os parâmetros do item aplicado e destaca três situações que ocorrem no processo de estimação:

- Conhecidas as habilidades do indivíduo respondente, tem-se que esti-

mar os parâmetros do item.

- Conhecidos os parâmetros do item, tem-se que estimar as habilidades do indivíduo.
- Conjuntamente tem-se que estimar as habilidades do indivíduo e os parâmetros do item.

Em qualquer uma das situações citadas acima, geralmente a estimação é feita pelo Método da Máxima Verossimilhança (MOOD; GRAYBILL; BOES, 1974) por meio da aplicação de algum processo iterativo, como o algoritmo Newton-Raphson ou Scoring de Fisher. Alguns procedimentos bayesianos também são aplicados com muita frequência. Na situação em que se deseja estimar tanto os parâmetros dos itens quanto as habilidades há duas abordagens usuais: estimação conjunta, parâmetros dos itens e habilidades, ou em duas etapas, primeiro a estimação dos parâmetros dos itens e, posteriormente, das habilidades (ANDRADE; TAVARES; VALLE, 2000).

### 3 MATERIAL E MÉTODOS

Os dados utilizados neste trabalho são referentes a dois testes diferentes compostos por 25 itens em cada teste, aplicados a uma turma piloto de 50 estudantes do curso presencial de nivelamento em Estatística Básica, oferecido pelo DEX-UFLA em Julho de 2015.

A estimação tanto dos parâmetros dos itens quanto da habilidade do indivíduo respondente foi feita por meio do *software IRT Command Language (ICL)* (HANSON, 2002), que usa estimativas de máxima verossimilhança e Bayesianas, sendo utilizado o algoritmo EM para computar os parâmetros do modelo utilizado.

<sup>1</sup>.

#### 3.1 Material

Para estruturar o BIE, com o objetivo de focar nos resultados de aprendizagem e organizar o que será avaliado, tomou-se como referência o conteúdo programático da disciplina de Estatística (GEX-112) do Departamento de Ciências Exatas da Universidade Federal de Lavras - UFLA.

Em síntese o conteúdo é composto de: Introdução. Estatística Descritiva, Amostragem, Probabilidade, Distribuições Discretas e Contínuas de Probabilidades, Distribuições de Amostragem, Teoria de Estimação, Teoria de Decisão, Regressão e Correlação.

Na Universidade Federal de Lavras a disciplina fica a cargo do Departamento de Ciências Exatas (DEX-UFLA), e o conteúdo nela trabalhado é avaliado e aprovado em assembleia do Departamento como se pode observar na Figura 2.

---

<sup>1</sup>O algoritmo EM utilizado pelo ICL pode ser encontrado em Hanson (1998) e Woodruff e Hanson (1997)

<b>MINISTÉRIO DA EDUCAÇÃO</b>					
<b>UNIVERSIDADE FEDERAL DE LAVRAS</b>					
<b>PRÓ-REITORIA DE GRADUAÇÃO</b>					
<b>EMENTA E CONTEÚDO PROGRAMÁTICO</b>					
Código	Denominação	Créditos <sup>(*)</sup>	Carga horária		
			Teórica	Prática	Total
GEX 112	Estatística	04	68	0	68
<b>Departamento</b>		<b>Professores</b>			
Ciências Exatas					
EMENTA: (Síntese do Conteúdo)					
Introdução. Estatística Descritiva. Amostragem. Probabilidade. Distribuições Discretas e Contínuas de Probabilidades. Distribuições de Amostragem. Teoria de Estimação. Teoria de Decisão. Regressão e Correlação.					
ASSINATURAS: _____					
_____					
_____					
_____					
Aprovado na Assembleia Departamental em ____/____/____					_____
Lavras, ____/____/____					Chefe do Departamento
* 17 horas/aula teóricas = 1 crédito					
17 horas/aula práticas = 1 crédito					

Figura 2 Ementa da disciplina Estatística.

Aplicou-se dois testes, de forma simultânea, com 25 itens cada teste, abrangendo todo o conteúdo programático da ementa da disciplina de Estatística básica, durante um período de quatro horas. A habilidade requerida era o conhecimento em: Estatística Descritiva, Amostragem, Probabilidade, Distribuição discreta e Contínua de Probabilidade, Distribuição de amostragem, Teoria de Estimação, Teoria de Decisão, Regressão Linear.

As respostas dos itens foram compostas de quatro alternativas, *a*, *b*, *c*, *d*, sendo apenas uma resposta correta e três distratores. Para calibração

(ou estimação) dos itens foi utilizado o *software* ICL.

### 3.2 ICL, estimação

O ICL (*IRT Command Language*) é um *software* criado por Brad Hanson, que pode estimar os parâmetros dos modelos logísticos de 1, 2 e 3 parâmetros para itens dicotômicos (ou resposta dada está correta ou errada), os parâmetros dos modelos de crédito parcial e crédito parcial generalizado para itens politômicos (mais de duas alternativas). Em todos os modelos pode-se considerar um ou mais grupos de examinados.

Existem versões do *software* para Linux, Windows e Macintosh. O ICL é um *software* livre que faz parte do projeto Ciências Sociais de Medição e está hospedado no site da *Source Forge*, sendo que seu desenvolvimento é administrado pelo site e permite a participação de qualquer pessoa interessada no seu desenvolvimento. O ICL é escrito na linguagem C++ e utiliza o *Estimation Toolkit for Item Response Models* (ETIRM) que é uma ferramenta para a estimativa de modelos da teoria de resposta ao item.

Segundo Pasquali e Primi (2003), os primeiros *softwares* a serem utilizados na TRI foram BICAL, LOGIST e BILOG, todos datando do final da década de 1970 e início da década de 1980. Atualmente diversos softwares de licenças restritas ou livres podem ser utilizadas em aplicações da TRI.

Mead, Morris e Blitz (2007) fizeram uma comparação entre o BILOG-MG (um *software* aplicado para estimar os parâmetros da TRI) e o ICL. Os pesquisadores compararam diversas características dos *softwares* como: manual, métodos de estimação, opções de estimação, possibilidade de escolha das distribuições *a priori*, informações produzidas pelo *software* como o ajuste do modelo, dentre outras. Mead, Morris e Blitz (2007) concluíram que o BILOG-MG é mais vantajoso que o ICL em alguns casos como, por exemplo, o BILOG apresenta um interface mais amigável, enquanto o ICL é mais linguagem de comando. Porém na precisão das estimações dos parâmetros, ambos os *softwares* apresentaram resultados semelhantes.

Jurich e Goodman (2009) também fizeram uma comparação entre o ICL e o *software* PARSCALE e concluíram que os dois *softwares* estimam os

parâmetros da TRI de forma muito próxima.

O ICL é muito flexível com relação aos critérios utilizados para realizar as estimações, como a escolha de em quantos intervalos será dividida a variável latente para a discretização, as distribuições das priores, os critérios de convergência, a quantidade de grupos respondentes e o modelo utilizado (MENDONÇA, 2012).

### 3.3 Métodos

Com base na literatura, para a elaboração da prova foram selecionados alguns atributos do discente julgados de maior relevância, pois tratavam de características que influenciam na aprendizagem do mesmo. Ou seja, dessa forma a prova foi desenvolvida para avaliar adequadamente o desempenho do aluno, com ênfase no raciocínio, na descoberta e na reprodução do conhecimento.

Nesse trabalho aplicou-se duas provas (1 e 2) com 25 itens cada, para 24 e 26 discentes da turma da Disciplina de Estatística Básica para o curso de nivelamento da Pós-graduação da UFLA no segundo semestre do período letivo do ano de 2015. As respostas foram transformadas 0 e 1 (0 para resposta errada e 1 para correta), processado no *software* ICL, produzindo-se os valores das estimativas de cada parâmetro em cada um dos itens. A partir desses resultados analisou-se o desempenho de cada item.

Por meio do parâmetro dificuldade ( $b$ ) classificou-se os itens de cada prova em muito fácil (MF), fácil (F), difícil (D) e muito difícil (MD). Com isso realizou-se a classificação observando os quantis de 25%, 50% e 75% na curva logística padrão. A partir desta classificação foram selecionados os itens que melhor se ajustam ao BIE.

Os gráficos da CCI foram elaborados por meio do *software* R (R CORE TEAM, 2016).

## 4 RESULTADOS E DISCUSSÃO

Os testes aplicados à turma de nivelamento foram elaborados com temas desenvolvidos durante o curso, os quais se fundamentam na ementa aprovada pelo Departamento de Ciências Exatas da Universidade Federal de Lavras (DEX), como já citado, cujo conteúdo programático se encontra no Apêndice.

### 4.1 Resultados dos alunos em cada teste

Os estudantes em número total de 50, divididos em dois grupos, sendo o grupo 1 composto de 24 alunos e o grupo 2 composto de 26 alunos, realizaram testes com vinte e cinco questões em cada teste. Estes testes foram corrigidos e transformados em informações de certo e errado. O número 1 corresponde à resposta correta e o número 0 corresponde à resposta errada.

Estes dados compõem as Tabelas 4 e 5 que se encontram no Apêndice deste trabalho. Nessas Tabelas pode-se verificar que os itens compõem as colunas e nas linhas os números correspondente a cada respondente, ou seja, os alunos.

### 4.2 Estimação dos parâmetros dos itens da prova 1 pelo o *software ICL*

Os testes codificados em 0 e 1, como podem ser visto na Tabela 4, foram inseridos no software ICL. A partir desse software foram obtidas as estimativas para cada item, no caso, as estimativas para  $a$ ,  $b$  e  $c$ . E o resultado dessa estimação pode ser verificada no Tabela 1.

O acerto casual (parâmetro  $c$ ) varia em uma escala de zero a um. Por meio da Tabela 1, pode-se verificar que os limites para a exclusão do item estão entre -4 e 4, ou seja, quando o item possui um valor maior que no intervalo será excluído do banco de itens. No caso, tem-se 25 itens, no qual, 8 itens (na Tabela 1 coloridos de vermelho) foram excluídos, pois ultrapassaram o limite definido neste trabalho. Ou seja, para a primeira prova foram considerados 17 itens bons. Para os oitos itens excluídos deve-se

Tabela 1 Estimativa dos parâmetros dos itens análise do ICL da prova 1.

Item	Discriminação ( $a$ )	Dificuldade ( $b$ )	Acerto casual ( $c$ )
1	0,682128	-3,075034	0,22819
2	0,859838	5,443654	0,179302
3	0,795129	0,149795	0,220007
4	1,359263	0,148105	0,226293
5	0,264677	1,064932	0,236629
6	0,871095	5,486771	0,126418
7	1,356247	-1,16794	0,217508
8	0,959081	-1,100258	0,212807
9	1,277914	0,009373	0,195593
10	0,849937	4,934898	0,152246
11	1,335143	-0,300615	0,198037
12	1,057194	0,309525	0,219785
13	0,866677	5,461528	0,179329
14	0,866677	5,483737	0,179379
15	0,231444	0,678638	0,238109
16	1,469599	-0,341745	0,184849
17	0,988101	0,300578	0,248671
18	0,264133	-1,368488	0,235243
19	0,893315	5,609357	0,179623
20	1,289256	1,226913	0,212353
21	1,091161	2,05196	0,162526
22	0,524242	1,458806	0,229627
23	0,888589	-5,578003	0,227494
24	0,29609	-2,90813	0,229041
25	0,840443	5,270111	0,152792

fazer um estudo para excluí-los ou reformulá-los e posteriormente testá-los.

### 4.3 Estimação dos parâmetros dos itens da prova 2 pelo o *software* ICL

Os testes codificados em 0 e 1, que compõem a Tabela 5 (Apêndice), foram inseridos no *software* ICL. A partir desse *software* foram obtidas as estimativas para cada item, no caso, as estimativas para  $a$ ,  $b$  e  $c$ . Este *software* gerou os resultados das análise dos parâmetros de cada item, de-

monstrados na Tabela 2.

Tabela 2 Estimativa dos parâmetros dos itens análise do ICL da prova 2.

Item	Discriminação ( <i>a</i> )	Dificuldade ( <i>b</i> )	Acerto casual ( <i>c</i> )
1	0,880697	-5,493616	0,227457
2	0,683362	-2,163112	0,229181
3	0,281786	-2,778351	0,229669
4	0,515037	-1,262397	0,232761
5	1,028224	0,012500	0,242348
6	0,885578	5,571754	0,105067
7	1,024761	0,169445	0,236706
8	0,880697	-5,493616	0,227457
9	1,169369	-0,592838	0,229174
10	1,073942	-1,33289	0,226143
11	0,847558	-0,333992	0,212211
12	1,002493	-0,837901	0,231157
13	0,348373	-2,922677	0,229238
14	1,055130	0,344332	0,234855
15	0,904001	2,120892	0,22951
16	0,523248	-0,658856	0,226832
17	0,551272	0,250050	0,235014
18	0,847256	1,329482	0,280777
19	0,124692	-1,103686	0,234884
20	0,830882	5,097401	0,240126
21	0,446784	1,936221	0,246136
22	0,597731	0,302846	0,243732
23	0,285789	3,287752	0,261231
24	0,955727	-2,31519	0,227353
25	0,450660	0,234471	0,228431

No caso, tem-se 25 itens, no qual, quatro itens (na Tabela 2 coloridos de vermelho) foram excluídos, pois ultrapassaram o limite definido neste trabalho. Ou seja, para a primeira prova foram considerados 21 itens bons. Para os oito itens excluídos deve-se fazer um estudo para excluí-los ou reformulá-los e posteriormente testá-los.

#### 4.4 Discussão dos resultados

De acordo com as Tabelas 1 e 2, observa-se a coluna que corresponde ao parâmetro dificuldade, ou seja,  $b$ , que está intrinsicamente ligado a escala da habilidade. Assim, ao definir o parâmetro “ $b$ ” tem-se uma probabilidade mínima de que o indivíduo acerte o referido item analisado.

Segundo Cunha (2014) é plausível pensarmos que quanto maior a dificuldade de uma questão maior deve ser a habilidade de um aluno para respondê-la corretamente. Como a dificuldade do item é fixa, ocorre uma variação no nível de proficiência e, logo, determinará a probabilidade de acerto da questão.

Nesse sentido, saber o nível do item é relevante, pois a partir desse nível pode-se ocorrer uma classificação do nível de dificuldade. A classificação do item utilizada nesse trabalho foi definida como sendo, muito fácil (MF), fácil (F), difícil (D) e muito difícil (MD) baseados nos quantis.

Para uma análise precisa das informações apresentadas neste trabalho, foi criada uma escala de dificuldades tomando como referência a distribuição logística com parâmetros de dificuldades igual a zero e discriminação 1 (logística padrão). Neste sentido foram considerados os quantis a 25%, 50% e 75%. A Tabela 3 ilustra a escala dos valores utilizados para classificação dos itens em relação ao grau de dificuldade utilizados nesta pesquisa.

Tabela 3 Escala de dificuldade utilizada como referência no trabalho.

Valor do parâmetro $b$ (dificuldade)	Grau de dificuldade
$b < -1,0986$	Muito Fácil
$-1,0986 \leq b < 0$	Fácil
$0 \leq b < 1,0986$	Difícil
$b \geq 1,0986$	Muito difícil

Em relação ao parâmetro  $c$ , definido como a probabilidade de um aluno com habilidade indefinidamente baixa acertar uma questão, tem-se que, ao realizar a prova existem duas possibilidades: o candidato “sabe que não sabe” e “chuta” ou ele “pensa que sabe” e não “chuta”, mas erra com alta probabilidade (pois tem baixa habilidade). Neste trabalho, cada item pos-

sui 4 alternativas, portanto a probabilidade de acerto casual foi em torno de 0,25. Valores maiores que esse são ruins pois indicam questões mal formuladas. Valores muito abaixo de 0,25 indicam que a probabilidade do aluno acertar aleatoriamente é muito baixa. Isso pode estar acontecendo devido ao fato de que o candidato pensa que sabe mas não sabe. Ele tem baixa habilidade, mas crê que entendeu o item. Consequentemente, terá a tendência de assinalar uma alternativa errada achando que acertou, induzido por algum conhecimento superficial do assunto. Daí essa probabilidade de acertar essa questão é baixa.

A probabilidade de acerto por indivíduo com baixa habilidade está relacionada com o nível de habilidade do candidato. Se ele “chuta” sabendo que não sabe tem mais chance de acertar do que se ele “chuta” por achar que sabe mas não sabe. Assim sendo, pode-se ter evidência de que uma prova considerada fácil possui valores baixos para o parâmetro  $c$ .

O resultado da estimação dos parâmetros dos itens, Tabelas 1 e 2, apresentam para cada item aplicado no teste as informações dos parâmetros de discriminação, dificuldade e acerto ao acaso, com erro inferior a  $10^{-6}$ , representando um alto grau de precisão. Com esses dados houve a possibilidade de uma descrição de cada item e quais serão considerados para a composição do banco de itens.

A partir do parâmetro de dificuldade, o item será avaliado para a exclusão do BIE. No caso deste trabalho, considerou-se o intervalo de  $] - 4, 4[$  para o valor observado do parâmetro  $b$  (dificuldade). Ou seja, os itens foram eliminados baseado nesse intervalo. Na prova 1, os itens 2; 6; 10; 13; 14; 19; 23; 25 foram eliminados, pois estavam fora do intervalo predeterminado para este trabalho e, dessa forma, tem-se que dos 25 itens foram eliminados 8, e assim ficando 17 itens considerados bons.

Já para a prova 2, os itens 1; 6; 8; 20 não fizeram parte do BIE, pois da mesma forma da prova 1 não estão no limite do intervalo. Nessa prova tinha-se 25 itens e foram eliminados 4, sendo assim, 21 itens foram considerados bons para compor o BIE. Um outro item que foi eliminado, mas não baseado no intervalo predeterminado, esse item é o 1, a sua eliminação é

justificada pelo fato de todos os alunos terem acertado, ou seja, o item ficou não informativo.

Na Figura 3 tem-se a classificação dos itens ao se comparar com os resultados da Tabela 3.

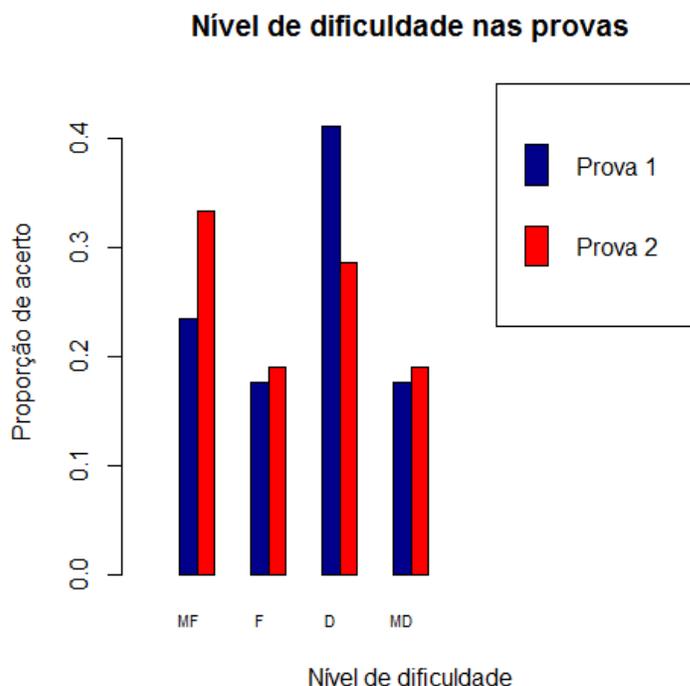


Figura 3 Gráfico de barras da proporção do nível de dificuldades das provas 1 e 2.

Por meio da Figura 3 pode-se concluir que o comportamento de cada item para a prova 1 apresentou uma tendência de classificação de D, pois a quantidade de itens classificado como D é maior do que as demais classificações. Já para a prova 2 pode-se verificar que a maioria dos itens possuem uma tendência de classificação MF, e nos demais houve um equilíbrio entre a quantidade de itens em cada classificação.

Para um aprimoramento da análise observou-se cada item da prova 1 e da prova 2, pois na Tabela 3 tem-se uma classificação mais geral para os

itens. Nas Tabelas 4 e 5 verificou-se que alguns itens demonstram serem poucos eficientes como instrumento avaliativo.

Dessa maneira, foi realizada uma análise com os itens que tiveram maior número de acerto por todos os alunos, o nível de dificuldade baixo e que não foram muito discriminativo.

Na Prova 1, os itens 1 e 24 foram considerados muito fáceis (observar parâmetro grau de dificuldade conforme a Tabela 1 e classificação do grau de dificuldade na Tabela 3), principalmente para o item 1, todos os alunos acertaram este item (observar Tabela 4 - coluna 1). Apesar do item 1 apresentar um parâmetro de discriminação ( $a = 0,682128$ ) relevante, o parâmetro de dificuldade é muito baixo ( $b = -3,075034$ ), logo não exige habilidade do conteúdo para resolver a questão, portanto foi descartado do BIE. A Figura 4 ilustra a curva CCI do item 1 da prova 1.

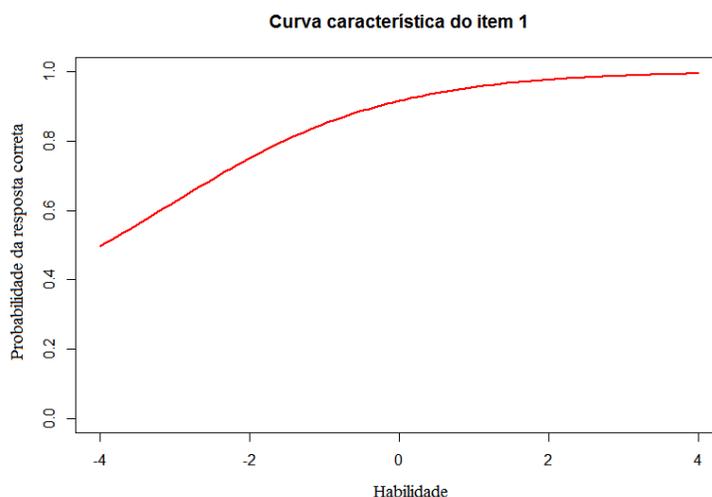


Figura 4 Curva CCI do item 1 da prova 1 ( $a = 0,682128$ ,  $b = -3,075034$ ,  $c = 0,228190$ ).

De forma análoga, pode ser feita a mesma análise para os demais itens, sendo retiradas do BIE as questões que apresentam baixa discriminação e baixo grau de dificuldade.

A Questão 1 está alinhada com o conteúdo do curso ministrado no tópico

de estatística descritiva e versa sobre classificação de variáveis.

(**Questão 01**) Considere as seguintes variáveis:

- I-** Receita do Imposto Sobre Serviços no município de Lavras na década de 90;
- II-** Números de filhos dos funcionários da UFLA;
- III-** Tipo de serviços prestados por empresas terceirizadas da câmara municipal de certo município;
- IV-** Número de aposentados do estado de Minas Gerais;
- V-** Grau de instrução de funcionários de uma repartição pública.

São exemplos de variáveis qualitativas e quantitativas os itens:

- a) Variáveis qualitativas: III, V; variáveis quantitativas: I, II, IV.**
- b) Variáveis qualitativas: I, II, V; variáveis quantitativas: II, III.**
- c) Variáveis qualitativas: I, V; variáveis quantitativas: II, III, IV.**
- d) Variáveis qualitativas: II, V; variáveis quantitativas: I, III, IV.**

A Figura 5 ilustra a curva CCI do item 24 da prova 1. A Questão 24, apresentou tanto o parâmetro de discriminação ( $a = 0,296090$ ) quanto o parâmetro de dificuldade ( $b = -2,908130$ ) muito baixo. Assim, o item 24, nem discrimina e nem exige habilidade para resolução da questão. Portanto, estes itens foram descartados do BIE.

A Questão 24 está alinhada com o conteúdo do curso ministrado no tópico Regressão e Correlação e versa sobre coeficiente de correlação. E neste caso, não deve ser destacada, pois a mesma é apenas uma questão muito fácil. O Quadro 6 ilustra-se esta questão.

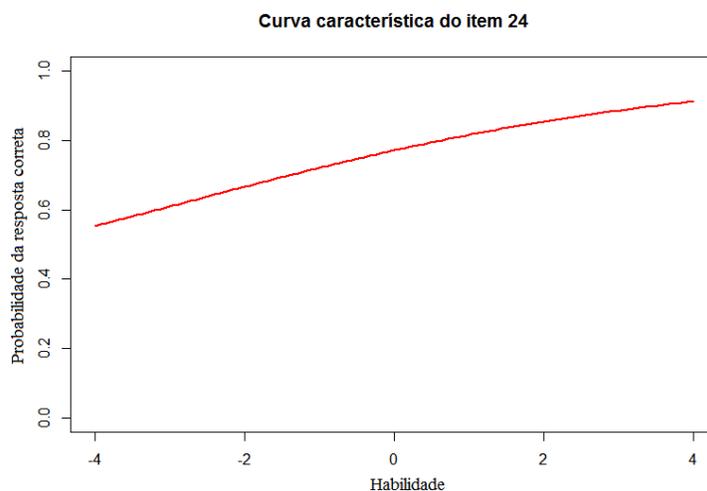


Figura 5 Curva CCI do item 24 da prova 1 ( $a = 0,296090$ ,  $b = -2,908130$ ,  $c = 0,229041$ ).

(**Questão 24**) Os dados de produção de leite de um grupo homogêneo de vacas holandesas tratadas com diferentes níveis de proteínas (%) na ração são:

X	10	12	14	16	18	20	22
Y	11,8	10,2	12,1	13,2	15,1	15,4	15,6

O valor do coeficiente de correlação entre X e Y é:

- (a) **0,9149**    (b) 0,8372    (c) 0,0850    (d)  $-0,9149$

A Questão 1 (observar Tabela 1) não foi bem elaborada, pois aborda conceitos muito básicos da estatística apresentando um grau de dificuldade baixo, exigindo pouca habilidade do estudante para responder a questão. Já a questão 24 trata-se de uma aplicação de fórmula, como as fórmulas e tabelas são fornecidas em folha anexa à prova, os alunos não tiveram tanta

dificuldade para responder de forma correta a questão. Sendo assim, houve a exclusão desses itens no BIE.

Por outro lado, na Questão 8 (observar Tabela 1) todos os estudantes acertaram a questão, porém o item 8 não foi eliminado, pois apresenta um grau de dificuldade ( $b = -1,100258$ ) e alto grau de discriminação ( $a = 0,959081$ ). Assim, apesar do item ser fácil é discriminativo, ou seja, se o estudante possuir maior habilidade a probabilidade de acerto do item é maior. A Figura 6 ilustra a curva CCI da Questão 8.

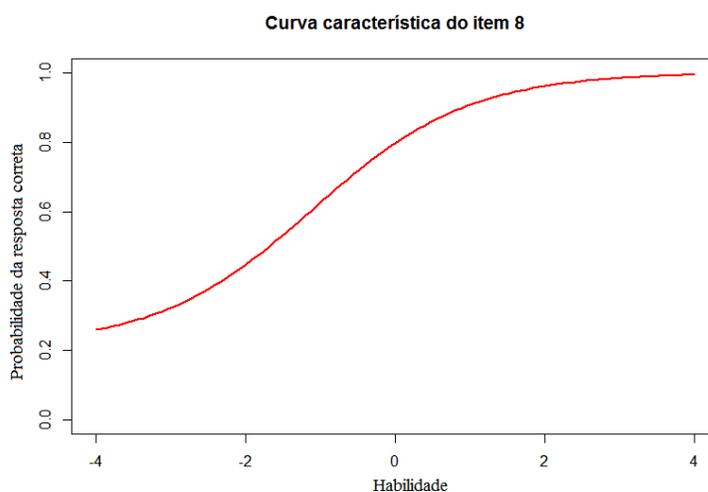


Figura 6 Curva CCI do item 8 da prova 1 ( $a = 0,959081$ ,  $b = -1,100258$ ,  $c = -0,212807$ ).

De forma análoga, pode ser feita a mesma análise para os demais itens, sendo retiradas do BIE as questões que apresentam baixa discriminação e baixo grau de dificuldade.

Na Prova 2, os itens 1 e 8 foram considerados muito fáceis (observar parâmetro grau de dificuldade na Tabela 2 e classificação do grau de dificuldade na Tabela 3), tanto o item 1 quanto item 8, quase todos os alunos acertaram estes itens (observar Quadro 2 - coluna 1 e coluna 8). Ambos os itens apresentarem um parâmetro de discriminação ( $a = ,880697$ ) relevantes, o parâmetro de dificuldade é muito baixo ( $b = -5,493616$ ), logo

não exigem habilidade do conteúdo para resolver a questão, portanto será descartado. A Figura 7 ilustra a curva CCI do item 8 da prova 2.

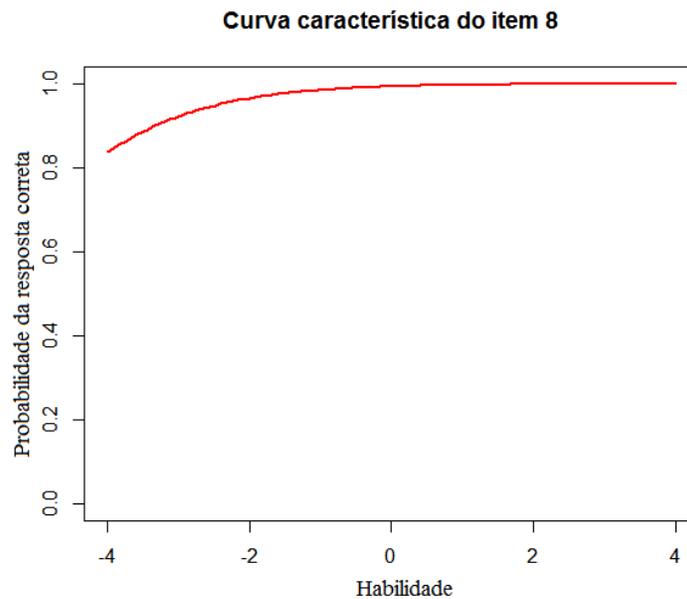


Figura 7 Curva CCI do item 8 da prova 2 ( $a = 0,880697$ ,  $b = -5,493616$ ,  $c = 0,227457$ ).

**(Questão 08)** O pessoal de inspeção de qualidade afirma que os rolos de fita magnética apresentam em média uma emenda a cada 50 metros. Admitindo-se que a probabilidade do número de emendas é dada pela distribuição de Poisson. Sendo assim qual é a probabilidade de ocorrer no mínimo uma emenda em um rolo de 100 metros.

- a) 85,46%
- b) 86,47%**
- c) 84,67%
- d) 83,46%

Para resolver a questão 8, o aluno deve ter conhecimento de distribuição de probabilidade. A questão tem discriminação alta, mas baixo grau de

dificuldade encontrando-se fora do intervalo considerado neste trabalho.

A Figura 8 ilustra a curva CCI do item 20 da prova 2.

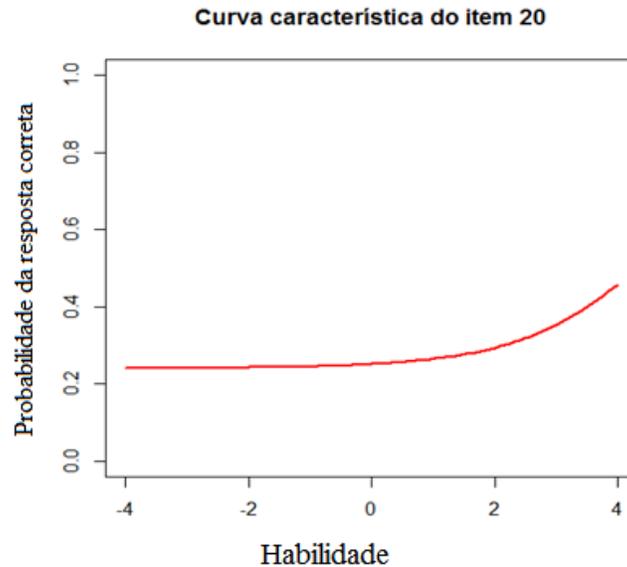


Figura 8 Curva CCI do item 1 da prova 2 ( $a = 0,880697$ ,  $b = -5,493616$ ,  $c = 0,227457$ ).

(**Questão 20**) Um pesquisador avaliou se a pressão sanguínea dos candidatos do último concurso para Técnico Administrativo da UFLA se alterava o início da prova. Em condições normais, sem stress, os candidatos entre 18 e 32 anos apresentaram uma pressão sistólica média de 120 mm de Hg. Após medir a pressão de 36 candidatos a cinco minutos do início da prova, foi encontrada a pressão sistólica média de 125,2 mm de Hg com desvio padrão amostral de 12 mm Hg. Deve-se testar:  $H_0: \mu = 120$ ,  $\mu > 120$  O valor calculado da estatística t é:

- a) **2,60**
- b) 0,43
- c) -0,43
- d) -2,60

A resolução do item 20 exige conhecimento de teste de hipóteses. Como pode ser visto pela Tabela 2, verifica-se que a questão 20 obteve valor de discriminação muito alto e além disso, o seu nível de dificuldade obtido está fora do intervalo considerado neste trabalho.

## 5 CONCLUSÕES

A avaliação deve ser o instrumento que permita ao professor a percepção e descrição das habilidades apreendidas pelo aluno no conteúdo trabalhado.

A análise dos resultados concluiu que o uso da TRI foi de suma importância, pois a mesma permitiu a construção e seleção de itens e além disso, permitindo também a identificação, de forma segura, do nível de habilidade alcançado pelo aluno em cada item.

A construção do banco de itens foi dada por meio da TRI e com o uso do *software* ICL. Assim, verificou-se que o uso dessa técnica estatística foi consistente para a obtenção de itens, no qual, os mesmos serão utilizados de modo preciso em avaliação de Estatística.

A partir da elaboração de testes aplicados em uma amostra de 50 alunos do curso de nivelamento em Estatística Básica, com o resultado da aplicação, utilizando a TRI, foi realizada uma classificação dos itens, que fazem parte do BIE que permitirá ao professor verificar a habilidade dos alunos.

Como trabalho futuro pode-se haver a inserção de novos itens no banco de itens para a construção de um BIE mais significativo. Além disso, utilizar outros modelos, como modelo multinível, entre outros para melhorias na metodologia.

## REFERÊNCIAS

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. Teria de resposta ao item: conceito e ampliações. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 14., 2000, Caxambu. **Anais...** Caxambu: SINAPE, 2000. 1 CD-ROM.

BAKER, F. B. **The basics of item response theory**. Washington: ERIC, 2001. 186 p.

BARROSO, J. C. D. G. A. Teria de resposta ao item: conceito e ampliações. In: ENCONTRO NACIONAL DE EDUCAÇÃO MATEMÁTICA, CULTURA E DIVERSIDADE, 5., 2010, Salvador. **Anais...** Salvador: PUC, 2010. 1 CD-ROM.

BATANERO, C. **Didáctica de la estadística**. Granada: Grupo de Investigación em Educación Estadística, 2001. 219 p.

BIRNBAUM, A. Some latent trait models and their use in inferring an examinee's ability. In: LORD, F. M.; NOVICK, M. R. (Ed.). **Statistical theories of mental test scores**. Oxford: Reading, 1968. p. 397-479.

BRASIL. **Decreto nº 2.494**, de 10 de fevereiro de 1998. Regulamenta o Art. 80 da LDB (Lei nº 9.394/96). Brasília, 1998. Disponível em: <<http://www.mec.gov.br/seed/tvescola/ftp/leis/D2494.doc>>. Acesso em: 5 fev. 2016.

CATALANI, E. M. T. **Avaliação e análise de itens pela Teoria Clássica dos Testes (TCT)**. 2013. Curso de curta duração ministrado. Disponível em: <<http://revistaeducacao.uol.com.br/textos/0/10-dicas-de-como-laborar-bons-itens-de-multipla-escolha-335338-1.asp>>. Acesso em: 7 fev. 2016.

CAZORLA, I. M. **A relação entre a habilidade viso-pictórica e o domínio dos conceitos estatísticos na leitura de gráficos**. 2002. 315 f. Tese (Doutorado em Educação) - Universidade Estadual de Campinas, Campinas, 2002.

COUTO, G.; PRIMI, R. Teoria de Resposta ao Item (TRI): conceitos elementares dos modelos para itens dicotômicos. **Boletim de Psicologia**, São Paulo, v. 61, n. 134, p. 1-15, 2011.

CUNHA, D. d'A. da. **A teoria de resposta ao item na avaliação em larga escala**: um estudo sobre o Exame Nacional de Acesso do Mestrado Profissional em Matemática em Rede Nacional- PROFMAT. 2014. 74 p. Dissertação (Mestrado Profissional em Matemática) - Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 2014.

FERNANDES, I.S. **Manual de orientações para a elaboração de itens**: questões de múltipla escolha: orientações técnico pedagógicas para o elaborador. Belo Horizonte: Faculdades Promove, 2015. 9 p.

GATTI, B.A. Os professores e suas identidades: o desenvolvimento da heterogeneidade. **Caderno de Pesquisa**, São Paulo, n. 98, p. 85-90, ago. 1996.

GUEWEHR, K. **Teoria da resposta ao item na avaliação de qualidade de vida de idosos**. 2007. 179 f. Dissertação (Mestrado em Epidemiologia) - Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

HALADYNA, T. M.; DOWNING, S. M.; RODRIGUEZ, M. C. A review of multiple-choice item-writing guidelines for classroom assessment. **Applied Measurement in Education**, Oxford, v. 15, p. 309-334, 2002.

HAMBLETON, R. K.; SWAMINATHAN, H.; ROGERS, H. J. **Fundamentals of item response theory**. Newbury Park: SAGE, 1991. 173 p.

HANSON, B. A. **IRT Command Language (ICL)**. 2002. Software. Disponível em: <<http://www.b-ah.com/software/irt/icl/index.html>>. Acesso em: 25 fev. 2016.

HANSON, B. A. *IRT parameter estimation using the EM algorithm*. 1998. Disponível em: <<http://www.b-ah.com/papers/note9801.html>>. Acesso em: 24 maio 2016.

JURICH, D.; GOODMAN, A. **Comparison of IRT parameter recovery in mixed format examinations using PARSELE and ICL.** Illinois: Illinois Institute of Technology, 2009. 28 p.

LAZARSELD, P. F. The logical and mathematical foundation of latent structure analysis. In: STAUFFER, S. A. et al. (Ed.) **Measurement and prediction.** Princeton: Princeton University Press, 1950. v. 4, p. 366-411.

LORD, F. M. The relation of test score to the trait underlying the test. **Educational and Psychological Measurement**, Thousand Oaks, v. 13, n. 1, p. 517-549, 1953.

LORD, F. M. The relation of the reliability of multiple-choice tests to the distribution of item difficulties. **Psychometrika**, Williamsburg, v. 17, n. 1, p. 181-194, 1952a.

LORD, F. M. **A theory of test scores.** Iowa City: Psychometric Society, 1952b. 83 p.

MEAD, A. D.; MORRIS, S. B.; BLITZ, D. L. **Open-source IRT: a comparison of BILOG-MG and ICL features and item parameter recovery.** Illinois: Institute of Psychology, 2007.

MENDONÇA, J. D. **Análise da eficiência de parâmetros da Teoria de Resposta ao Item pelo Software ICL.** 2012. 127 f. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2012.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics.** 3rd ed. New York: McGraw Hill, 1974. 564 p.

MOREIRA JUNIOR, F. J. Aplicações na teoria de resposta ao item (TRI) no Brasil. **Revista Brasileira de Biometria**, São Paulo, v. 28, n. 4, p. 137-170, 2010.

PASQUALI, L.; PRIMI, R. Fundamentos da teoria da resposta ao item: TRI. **Avaliação Psicológica**, Porto Alegre, v. 2, n. 2, p. 99-110, 2003.

R CORE TEAM.**R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2016. Disponível em: <<https://www.r-project.org/>>. Acesso em: 20 fev. 2016

RASCH, G.**Probabilistic models for some intelligence and attainment tests**. Chicago: MESA Press, 1960. 199 p.

SAMEJIMA, F.**A general model for free-response data**. Richmond: Psychometric Society, 1972. (Psychometric Monograph, 18).

SAMEJIMA, F. Graded response model of the latent theory and tailored testing. **Psychometrika Monograph Supplement**, New York, v. 34, n. 4, p. 15-29, 1969.

SÃO PAULO (Estado). Secretaria da Educação. Coordenadoria de Estudos e Normas Pedagógicas. **Legislação de ensino fundamental e médio**. São Paulo, 2008.

SOARES, M. S.**Proposta de um software de banco de itens calibrados pela teoria da resposta ao item (TRI), para uso de professores de matemática da educação básico**. 2014. 127 f. Dissertação (Mestrado em Ciências) - Universidade Federal do Acre, Rio Branco, 2014.

WODEWOTZKI, M. L. L.; JACOBINI, O. R. O Ensino de estatística no contexto da educação matemática. In: BICUDO, M. A. V.; BORBA, M. de C. (Org.). **Educação matemática: pesquisa em movimento**. São Paulo: Cortez, 2004. p. 232-249.

WOODRUFF, D. J.; HANSON, B. A. Estimation of item response models using the EM algorithm for finite mixtures. In: ANNUAL MEETING OF THE PSYCHOMETRIC SOCIETY, 1997, Gatlinburg. **Proceedings...** Gatlinburg, 1997. Disponível em: <<http://www.b-ah.com/papers/paper970.html>>. Acesso em: 28 mar. 2016.

WRIGHT, G. H. V. **An essay in denotic logic and the general theory of action**. Amsterdam: North-Holland, 1968. 77 p.

## APÊNDICE

Conteúdo programático

1 Introdução à Disciplina

1.1 Apresentação do Professor e Estudantes.

1.2 Apresentação do Plano de Curso.

1.3 Metodologia de Ensino-Aprendizagem e Avaliação.

1.4 A Disciplina nos Currículos e Sua Integração com Outros Estudos.

1.5 A Disciplina de Formação do Profissional e da Pessoa.

2 Introdução à Estatística

2.1 Definições Básicas, Motivação e Objetivos.

2.2 Tipos de Dados.

2.3 População e Amostras.

2.4 Amostras Aleatórias.

3 Estatística Descritiva

3.1 Importância da Estatística Descritiva.

3.2 Coleta, organização e apresentação de dados. Distribuições de Frequências.

3.3 Medidas de Posição.

3.4 Medidas de Variabilidade.

3.5 Medidas Descritivas da Distribuição: Coeficientes de Curtose e Assimetria.

3.6 Tópicos em Estatística Descritiva.

4 Amostragem

4.1 Importância da Amostragem.

4.2 Amostra e População. Amostragem Probabilística e Não-Probabilística.

4.3 Amostragem Simples ao Acaso, Estratificada, por Conglomerados e Sistemática.

4.4 Tópicos em Amostragem.

5 Cálculo de Probabilidades

5.1 O conceito de Probabilidade e sua importância.

5.2 Definição Não-Axiomática de Probabilidade.

- 5.3 Espaço Amostral e Eventos.
- 5.4 Definição Axiomática de Probabilidade.
- 5.5 Propriedades de Probabilidade, Probabilidades condicionais e Independência. Teorema de Bayes.
- 5.6 Tópicos em Cálculo de Probabilidades.
- 6 Variáveis Aleatórias e Distribuição de Probabilidades
  - 6.1 Definição de Variável Aleatória, Classes: Discretas e Contínuas.
  - 6.2 Função de Distribuição, Função Densidade de Probabilidade, Função de Probabilidade.
  - 6.3 Esperança Matemática e suas leis. Momentos Estatísticos.
  - 6.4 Distribuições conjunta, marginal, e condicional, de variáveis aleatórias. Independência de variáveis aleatórias. Funções de variáveis aleatórias.
  - 6.5 Distribuição de Probabilidades Discretas: Bernoulli, Binomial, Poisson, Geométrica.
  - 6.6 Distribuição de Probabilidades Contínuas: Uniforme, Exponencial, Normal, Aproximação Normal da Binomial e Poisson.
  - 6.7 Tópicos em Distribuição de Probabilidades.
- 7 Distribuições de Amostragem
  - 7.1 Importância do Estudo das Distribuições de Amostragem.
  - 7.2 Distribuição de Amostragem das Médias.
  - 7.3 Distribuições de Amostragem de Proporções.
  - 7.4 Distribuições  $t$ ,  $2$  e  $F$ .
  - 7.5 Tópicos em Distribuições de Amostragem.
- 8 Teoria da Estimação
  - 8.1 Importância do Estudo da Teoria da Estimação.
  - 8.2 Estimação por Ponto e por Intervalo. Propriedades dos Estimadores.
  - 8.3 Métodos de Estimação: Momentos, Máxima Verossimilhança e Mínimos Quadrados.
  - 8.4 Estimação por Intervalo das Médias, Variâncias e Proporções.
  - 8.5 Dimensionamento de Amostras.
  - 8.6 Estimação por Intervalo de Diferenças de Médias e Proporções.
  - 8.7 Tópicos em Teoria de Estimação.

## 9 Teoria da Decisão

9.1 Importância do Estudo da Teoria da Tomada da Decisões.

9.2 Hipótese Estatística. Erros Envolvidos em um Processo de Decisão.

9.3 Construção de uma Regra de Decisão e Mecânica Operacional de Aplicação de Testes.

9.4 Testes para Médias e Proporções Binomiais.

9.5 Testes para Diferenças de Médias e Proporções.

9.6 Testes de Independência, Aderência e Comprovação de Leis.

9.7 Tópicos em Teoria de Decisão.

## 10 Regressão e Correlação

10.1 Importância da Regressão e Correlação.

10.2 Regressão Linear Simples.

10.3 Correlação Linear (Correlação de Pearson).

10.4 Tópicos em Regressão e Correlação.

## 11 Avaliação da Disciplina

11.1 Avaliação do Conteúdo do Curso.

11.2 Avaliação de Atuação de Aluno.

11.3 Avaliação da Atuação do Professor.

11.4 Avaliação das Condições Materiais e Físicas em que se Desenvolve o Curso.

Tabela 4 Dados prova 1.

Alunos	Itens																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1	1	1	1	1	0	0	1	1	0	0	0	1	1	1	1	0	1	1
2	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	0
3	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0	0	0	0	1
4	1	1	1	1	1	0	0	1	0	1	1	0	0	0	0	0	1	0	1
5	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1
6	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	0	0	1
7	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0
8	1	0	0	1	0	0	1	1	0	1	1	1	0	0	0	0	1	1	1
9	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	1	1	0	0
10	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	1	1	0	0
11	1	1	1	1	0	0	0	1	0	1	1	1	1	0	1	1	1	1	1
12	1	1	1	1	0	0	0	1	1	1	1	1	1	0	0	1	1	1	0
13	1	1	1	0	1	0	1	1	0	1	0	0	1	0	0	0	1	0	1
14	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	0	1	1	0	1	1	1	1	1	1	1	0	0	1	1	0	0
16	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1	1
17	1	1	1	0	0	1	0	1	0	1	1	1	1	0	0	1	1	1	1
18	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1	0	1
19	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	1	0	0	1
20	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1
21	1	1	0	1	0	0	1	1	1	0	0	0	1	1	0	1	1	1	1
22	1	1	1	1	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0
23	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	1
24	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0

Alunos	Itens					
	20	21	22	23	24	25
1	0	0	1	1	0	0
2	0	1	1	1	1	0
3	0	1	0	0	1	1
4	0	1	1	1	1	0
5	0	0	1	0	1	1
6	0	1	1	0	1	1
7	0	0	1	0	1	1
8	0	1	0	1	1	0
9	0	0	0	0	1	1
10	0	0	0	0	1	1
11	1	1	0	0	1	1
12	1	1	1	0	1	1
13	0	0	0	0	1	0
14	0	0	1	0	1	0
15	0	0	0	1	1	1
16	1	0	1	1	1	1
17	1	1	1	1	1	1
18	0	0	0	1	1	0
19	0	1	0	1	1	0
20	1	0	1	0	1	0
21	1	0	1	1	1	1
22	0	0	0	0	0	1
23	0	0	1	0	1	0
24	0	1	1	0	1	1

Tabela 5 Dados prova 2.

Alunos	Itens																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1	1	1	0	1	0	1	1	0	0	1	0	0	0	0	1	1	1	0
2	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0
3	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0
4	1	0	1	1	1	0	1	1	1	0	1	1	0	0	1	1	1	1	0
5	1	0	0	0	0	0	1	1	1	0	1	0	0	0	1	1	1	0	0
6	1	1	0	0	1	0	1	1	0	0	1	1	0	0	0	0	0	0	0
7	1	0	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	0
8	1	0	0	0	1	0	1	1	1	0	0	0	0	0	1	1	1	1	0
9	1	0	1	0	1	0	1	1	1	0	1	1	0	1	1	0	0	1	0
10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1
11	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	0	0	0
12	1	0	0	1	0	0	1	1	0	1	1	1	0	1	0	1	0	1	0
13	1	1	1	1	1	1	1	1	0	1	1	0	1	0	1	1	1	0	0
14	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	1	1
15	1	0	1	1	0	0	1	1	0	0	0	1	0	0	1	1	1	0	0
16	1	0	1	1	1	0	1	0	1	0	1	1	1	0	1	1	1	1	0
17	1	0	1	1	0	0	1	1	1	0	1	1	0	0	0	1	1	1	0
18	1	1	1	1	0	0	1	1	1	0	1	0	1	0	1	1	0	1	0
19	0	0	0	0	0	1	1	1	0	0	1	0	0	1	1	0	1	0	1
20	1	0	0	0	1	0	1	1	0	0	1	1	0	0	1	1	0	1	0
21	1	0	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0
22	1	0	1	1	0	0	0	1	1	0	0	0	1	0	0	0	1	1	0
23	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1
24	1	0	0	1	0	0	1	1	1	0	1	1	0	0	1	1	1	1	0
25	1	0	1	1	1	0	1	1	1	0	1	1	0	0	0	1	1	1	0
26	1	0	1	1	0	0	1	1	1	0	1	0	0	0	0	1	1	1	0

Alunos	Itens					
	20	21	22	23	24	25
1	1	0	1	1	0	0
2	0	0	1	1	0	1
3	0	0	0	1	1	1
4	1	0	1	1	0	0
5	0	0	1	1	1	0
6	0	1	1	1	1	0
7	1	0	1	1	1	1
8	1	0	0	1	0	0
9	0	0	0	1	1	0
10	0	0	0	1	0	0
11	0	0	0	1	1	0
12	0	0	1	1	1	0
13	0	0	1	1	1	1
14	0	0	0	1	1	0
15	0	0	0	1	1	0
16	0	0	0	1	1	0
17	1	1	0	1	1	0
18	1	0	0	1	1	1
19	1	0	0	1	1	0
20	0	0	0	1	1	0
21	0	0	1	1	1	0
22	1	1	0	1	1	0
23	0	0	1	1	1	0
24	1	1	1	1	1	0
25	1	1	0	1	1	0
26	0	0	1	1	1	1